Scale aware modeling and monitoring of the urban energy chain

Zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

von der KIT-Fakultät für

Bauingenieur-, Geo- und Umweltwissenschaften

des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von MSc.

Alexandru NICHERSU

aus Tulcea, Rumänien

Tag der mündlichen Prüfung: 31.05.2022

Referent: Prof. Dr.-Ing. habil. Stefan HINZ

Institut für Photogrammetrie und Fernerkundung (IPF)

Karlsruher Institut für Technologie (KIT)

Korreferent: Prof. Dr. Martin Breunig

Geodätischen Institut (GIK)

Karlsruher Institut für Technologie (KIT)

Korreferent: assist. Prof. Giorgio Agugiaro, PhD

Department of Urbanism, Faculty of Architecture and the Built Environment

Delft University of Technology (TU Delft)

Karlsruhe 2022

Declaration of Authorship

I, Alexandru NICHERSU, declare that this thesis titled, "Scale aware modeling and monitoring of the urban energy chain" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly indicated.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

"If you really want to affect environmental outcomes it's not shaping a building that matters, it's shaping a community"

Peter Calthorpe (Architect and urban planner)

"We ourselves feel that what we are doing is just a drop in the ocean. But the ocean would be less because of that missing drop."

Mother Teresa (Nun)

"Sustainable development is the central challenge of our times"

Ban Ki-Moon (Diplomat)

"There is no such thing as a free lunch"

Kevin McKoen (Physicist)

"... it's always a team sport"

Diana Nyad (Long-distance swimmer)

KARLSRUHE INSTITUTE OF TECHNOLOGY

Abstract

Department of Civil Engineering, Geo and Environmental Sciences

Doctor of Engineering

Scale aware modeling and monitoring of the urban energy chain

by Alexandru NICHERSU

Keywords: smart cities, 3D semantical city models, spatial aware energy data, interoperability, UBEM, IoT, spatial actors, spatial scales, spatial standards, urban energy management, urban planning, city-wide energy chain, spatial awareness

With energy modeling at different complexity levels for smart cities and the concurrent data availability revolution from connected devices, a steady surge in demand for spatial knowledge has been observed in the energy sector. This transformation occurs in population centers focused on efficient energy use and quality of life. Energy-related services play an essential role in this mix, as they facilitate or interact with all other city services. This trend is primarily driven by the current age of the *Ger.: Energiewende* or energy transition, a worldwide push towards renewable energy sources, increased energy use efficiency, and local energy production that requires precise estimates of local energy demand and production. This shift in the energy market occurs as the world becomes aware of human-induced climate change, to which the building stock has a significant contribution (40% in the European Union [71]). At the current rate of refurbishment and building replacement, of the buildings existing in 2050 in the European Union, 75% would not be classified as energy-efficient, [70]. That means that substantial structural change in the built environment and the energy chain is required to achieve EU-wide goals concerning environmental and energy policy. These objectives provide strong motivation for this thesis' work and are generally made possible by energy monitoring and modeling activities that estimate the urban energy needs and quantify the impact of refurbishment measures.

To this end, a modeling library called aEneAs was developed in the scope of this thesis that can perform city-wide building energy modeling. The library performs its tasks at the level of a single building and was a first in its field, using standardized spatial energy data structures that allow for portability from one city to another. For data input, extensive use was made of digital twins provided from CAD, BIM, GIS, architectural models, and a plethora of energy data sources. The library first quantifies primary thermal energy demand and then the impact of refurbishment measures. Lastly, it estimates the potential of renewable energy production from solar radiation. aEneAs also includes network modeling components that consider energy distribution in the given context, showing a path toward data modeling and

simulation required for distributed energy production at the neighborhood and district level.

In order to validate modeling activities in solar radiation and green façade and roof installations, six spatial models were coupled with sensor installations. These digital twins are included in three experiments that highlight this monitoring side of the energy chain and portray energy-related use cases that utilize the spatially enabled web services SOS-SES-WNS, SensorThingsAPI, and FIWARE. To this author's knowledge, this is the first work that surveys the capabilities of these three solutions in a unifying context, each having its specific design mindset.

The modeling and monitoring activity and their corresponding literature review indicated gaps in scientific knowledge concerning data science in urban energy modeling. First, a lack of standardization regarding the spatial scales at which data is stored and used in urban energy modeling was observed. In order to identify the appropriate spatial levels for modeling and data aggregation, scale is explored in-depth in the given context and defined as a byproduct of resolution and extent, with ranges provided for both parameters. To that end, a survey of the encountered spatial scales and actors in six different geographical and cultural settings was performed. The information from this survey was used to put forth a standardized spatial scales definition and create a scale-dependent ontology for use in urban energy modeling. The ontology also provides spatially enabled persistent identifiers that resolve issues encountered with object relationships in modeling for inheritance, dependency, and association. The same survey also reveals two significant issues with data in urban energy modeling. These are data consistency across spatial scales and urban fabric contiguity. The impact of these issues and different solutions such as data generalization are explored in the thesis.

Further advancement of scientific knowledge is provided specifically with spatial standards and spatial data infrastructure in urban energy modeling. A review of use cases in the urban energy chain and a taxonomy of the standards were carried out. These provide fundamental input for another piece of this thesis: inclusive software architecture methods that promote data integration and allow for external connectivity to modern and legacy systems. In order to reduce time-costly extraction, transformation, and load processes, databases and web services to ferry data to and from separate data sources were used. As a result, the spatial models become central linking elements of the different types of energy-related data in a novel perspective that differs from the traditional one, where spatial data tends to be non-interoperable / not linked with other data types. These distinct data fusion approaches provide flexibility in an energy chain environment with inconsistent data structures and software. Furthermore, the knowledge gathered from the experiments presented in this thesis is provided as a synopsis of good practices.

KARLSRUHER INSTITUT FÜR TECHNOLOGIE

Zusammenfassung

Fakultät für Bauingenieur-, Geo- und Umweltwissenschaften

Doktor-Ingineurs

Skalenbewusste Modellierung und Monitoring der städtischen Energiekette

by Alexandru NICHERSU

Stichworte: Smart Cities, semantische 3D-Stadtmodelle, ortsbezogene Energiedaten, Interoperabilität, UBEM, IoT, räumliche Akteure, räumliche Maßstäbe, räumliche Standards, energetisches Stadtmanagement, Raumplanung, stadtweite Energiekette, Raumbewusstsein

Einhergehend mit der Energiemodellierung auf verschiedenen Komplexitätsebenen für smart cities und der gleichzeitigen Revolution der Datenverfügbarkeit durch vernetzte Geräte ist im Energiesektor ein stetiger Anstieg der Nachfrage nach räumlichem Wissen zu beobachten. Dieser Wandel vollzieht sich in Bevölkerungszentren, die auf effiziente Energienutzung und Lebensqualität ausgerichtet sind. Energiebezogene Dienstleistungen spielen in diesem Mix eine wesentliche Rolle, da sie alle anderen städtischen Dienstleistungen erleichtern oder mit ihnen interagieren. Dieser Trend wird vor allem durch die Entwicklungen der Energiewende angetrieben, einem weltweiten Vorstoß zum Ausbau erneuerbarer Energiequellen, effizienterer Energienutzung und lokaler Energieerzeugung. Dieser Wandel des Energiemarktes, wird durch eine wachsende Wahrnehmung des vom Menschen verursachten Klimawandels befeuert, zu dem der Gebäudebestand einen erheblichen Beitrag leistet (40 % in der Europäischen Union [71]). Bei der derzeitigen Sanierungs- und Erneuerungsrate würden von den im Jahr 2050 in der Europäischen Union bestehenden Gebäuden 75 % nicht als energieeffizient eingestuft werden, [70]. Das bedeutet, dass ein erheblicher Strukturwandel in der bebauten Umwelt und der Energiekette erforderlich ist, um die EU-weiten Ziele in der Umwelt- und Energiepolitik zu erreichen. Diese Ziele stellen eine starke Motivation für diese Arbeit dar und werden im Allgemeinen durch Energiemonitoring und Modellierungsaktivitäten ermöglicht.

Zu diesem Zweck wurde im Rahmen dieser Arbeit eine Modellierungsbibliothek namens aEneAs entwickelt, mit deren Hilfe stadtweite Gebäudemodellierung durchgeführt werden kann. Die Bibliothek repräsentiert die Elemente der Energiekette auf der Ebene einzelner Gebäude und war ein Novum in ihrem Bereich, da sie standardisierte räumliche Energiedatenstrukturen verwendet, die eine Übertragbarkeit von einer Stadt zur anderen ermöglichen. Für die Dateneingabe wurden digitale Zwillinge aus CAD, BIM, GIS, Architekturmodellen und einer Reihe von Energiedatenquellen verwendet. Die Bibliothek quantifiziert den primären Wärmeenergiebedarf

und die Auswirkungen von Sanierungsmaßnahmen. aEneAs enthält auch Komponenten zur Netzwerkmodellierung, die die Energieverteilung im gegebenen Kontext berücksichtigen und einen Weg zur Datenmodellierung und Simulation aufzeigen, die für die verteilte Energieerzeugung auf Nachbarschafts- und Stadtteilebene erforderlich sind.

Um die Modellierungsaktivitäten im Bereich der Sonneneinstrahlung und der Fassaden- und Dachbegrünung zu validieren, wurden sechs räumliche Modelle mit Sensorinstallationen gekoppelt. Diese digitalen Zwillinge werden in drei Experimenten genutzt, die diese Überwachungsseite der Energiekette beleuchten und energiebezogene Anwendungsfälle, welche die raumbezogenen Webdienste SOS-SES-WNS, SensorThings API und FIWARE nutzen, darstellen. Soweit dem Autor bekannt ist, ist dies die erste Arbeit, die die Fähigkeiten dieser drei Webdienste in einem vereinheitlichenden Kontext untersucht, wobei jeder Dienst seine eigene Design-Mentalität hat.

Im Rahmen der Modellierungs- und Überwachungsaktivitäten und der dazugehörigen Literaturrecherche zeigten sich Lücken in den datenwissenschaftlichen Methoden im Bereich der städtischen Energiemodellierung. Erstens wurde ein Mangel an Standardisierung in Bezug auf die räumlichen Maßstäbe festgestellt, auf denen Daten in der städtischen Energiemodellierung gespeichert und verwendet werden. Um die geeigneten räumlichen Ebenen für die Modellierung und Datenaggregation zu ermitteln, wurde der Maßstab im gegebenen Kontext eingehend untersucht und als Nebenprodukt von Auflösung und Umfang definiert, wobei für beide Parameter Gültigkeitsbereiche angegeben werden. Zu diesem Zweck wurde eine Erhebung über die angetroffenen räumlichen Maßstäbe und Akteure in sechs verschiedenen geografischen und kulturellen Umgebungen durchgeführt. Die Informationen aus dieser Umfrage wurden verwendet, um eine standardisierte Definition von räumlichen Maßstäben zu erstellen und eine maßstabsabhängige Ontologie für die Verwendung in der städtischen Energiemodellierung zu entwickeln. Diese Ontologie bietet auch ortsbezogene persistente Identifikatoren, die Probleme mit Objektbeziehungen bei der Modellierung. Die gleiche Umfrage zeigt auch zwei wichtige Probleme mit Daten in der städtischen Energiemodellierung. Dabei handelt es sich um die Datenkonsistenz über räumliche Maßstäbe hinweg und die Kontiguität der städtischen Struktur.

Weitere wissenschaftliche Fortschritte konnten insbesondere durch Standards für räumliche Daten und eine ortsbezogene Dateninfrastruktur für die Modellierung städtischer Energie erreicht werden. Es wurde ein Überblick über die Anwendungsfälle in der urbanen Energiekette und eine Taxonomie der Standards erstellt. Diese liefern den grundlegenden Input für einen weiteren Teil dieser Arbeit: integrative Softwarearchitekturmethoden, die die Datenintegration fördern und die externe Konnektivität zu modernen und bestehenden Systemen ermöglichen. Um zeit- und kostenintensive Extraktions-, Transformations- und Ladeprozesse zu reduzieren, wurden Datenbanken und Webservices eingesetzt, um Daten zu und von separaten Datenquellen zu transportieren. Infolgedessen werden die räumlichen Modelle zu zentralen Verbindungselementen der verschiedenen Arten von energierelevanten Daten in einem neuartigen Ansatz, der sich von dem traditionellen unterscheidet, in welchem räumliche Daten in der Regel nicht interoperabel sind. Diese unterschiedlichen Datenfusionsansätze bieten Flexibilität in einem Ökosystem von Energieketten mit inkonsistenten Datenstrukturen und Software.

Acknowledgements

I want to acknowledge my family's unwavering support throughout the entire life of this project. You are my lighthouse; I'd be lost without you. Thank you for guiding my voyages.

The most significant influence on my development has come from my parents, my mom, a mechanical engineer, and my dad, a Ph.D. in geodesy. Both provided me with endless love and persistent stability. Since the very start, they have been a source of strength and inspiration in my life. They have taught me priceless lessons, from decency to modesty, geometry to French, and the value of teamwork to the significance of family. In addition, my friend, teammate, ally, collaborator, companion, and twin sister, who has known me ever since I was smaller than a raindrop, was always there, guided by her outstanding determination. She offered support and different perspectives, as twins always do, and her career as a Ph.D. Civil Engineer gifted me with another path of inspiration.

My dear sweetheart, partner, love, companion, and beloved wife provided me with a constant source of support and encouragement even as she moved to another country, grappled with recognizing her titles to be allowed to practice Orthodontics in a new language (not a feat for the fainthearted) and pursued, parallel to my own, her Ph.D. in Medicine. To myself and all who know her, she is a fountain of perseverance and integrity and played a large part in bringing this thesis' to the printer.

We traveled the world, had two kids, and completed our PhDs.

For the duration of my employment at EIFER, I have actively participated in constructing a city simulation platform with a broad goal of urban simulation. The project, named CURTIS (Coupled URban SimulaTIon), financed by the EDF (Électricité de France) group, successfully built and delivered a city simulation platform. This has supported activities (meetings, conferences, publications, and student thesis supervision) in this Ph.D. For that, the author is ever grateful to the EIFER management: Dr. Jean Copreaux, Dr. Roman Zorn, Dr. Nurten Avci and Ludmila Gautier, the EDF management: Fabrice Casciani and Maxime Cassat, the project management: Dr. Kevin McKoen, David Blin and Dr. Alberto Pasanisi, my project colleagues: Alexander Simons, Dr. Jochen Wendel, Dr. Sebastien Cajot, Dr. Samuel Thiriot, Dr. Daniel Fehrenbach, Dr. Jonathan van der Kamp, Francisco Marzabal, Isaac Boates, Jason Yip, Manfred Wieland, Dr. Maria Sipowicz, Dr. Atom Mirakyan, Dr. Syed Monjur Murshed, Omar Benhamid, Saed Muhamad and Wanji Zhu, the partners at ULPGC: Pablo Fernández and Jaisiel Santana and my students: Aleksandra Gabryjalowicz, Chenfeng Liu, Yao Jiacheng and Thibault Morin.

Sincere thanks go to the thesis supervisors for their assistance. The GRACE school system had a system of three persons helping and guiding, each with their fair share of work and associate contribution. Prof. Dr. Stefan Hinz, head of the IPF (Institut für Photogrammetrie und Fernerkundung), is the supervisor in charge and the main guarantor responsible for the work presented here. On behalf of EIFER, my former group manager from EIFER, Dr.-Ing. Andreas Koch made this thesis possible from an administrative and managerial perspective. Both Dr. Koch and Dr. Hinz saw the potential of collaboration between IPF and EIFER and helped create the structure and support that this Ph.D. candidate benefited from in his work. For this, I am very much grateful.

The third supervisor, Dr.-Ing. Sven Wursthorn, my direct/personal supervisor, assisted me in the continuous struggle that a Ph.D. is. He constantly provided different perspectives that I would have otherwise overlooked and was actively involved in the entire process of the thesis conception, development, and writing. At times he motivated me and acted as a beacon of hope. In addition to this, there was infrastructure (servers), expertise (geomatics), and human compassion that Dr. Wursthorn never hesitated to show, for which I am deeply grateful.

Other individuals also assisted this construct. These are Prof. Dr. Ute Karl (EIFER's scientific director) and Dr.-Ing. Andreas Schenk (GRACE's scientific coordinator). Both Prof. Karl and Dr. Schenk supported this work from the very beginning. I recognize and thank both of you for your guidance and support.

This thesis life ran parallel with standardization efforts in the semantic city models field. The author was very fortunate to be involved with the Energy ADE and the UtilityNetwork ADE development efforts. The experience gained helping these standardization efforts greatly facilitated the work within this manuscript. For that, I would like to thank Dr. Joachim Benner, whose data modeling skills helped finish the first version and then further improve the Energy ADE to version 2.0, Romain Nouvel for his vision and work on the Energy ADE extension, Dr. Giorgio Agugiaro for creating database tools that facilitate work with the Energy ADE and the UtilityNetwork ADE, Dr. Tatjana Kutzner and Prof. Thomas Kolbe for their work on both CityGML and the UtilityNetwork ADE.

Extensive assistance was provided in proofreading by Ursula Antoni and Dr. Franz Symalla. This support is highly cherished and greatly improved the present thesis.

Furthermore, this doctoral candidate was supported by many individuals and entities. The author would like to mention that many other silent contributors and colleagues helped the thesis with advice, administratively, and motivation. That is why: it is teamwork.

Lastly, I would like to thank all my teachers and professors for their passion, inspiration, and precious time. I am but a drop in your ocean of teaching.

Contents

List of Figures

List of Tables

List of Abbreviations

ACH	**A**ir ex**C**hange **R**ate
ADE	**A**pplication **D**omain **E**xtension
API	**A**pplication **P**rogramming **I**nterface
ANN	**A**rtifical **N**eural **N**etworks
BESTEST	**B**uilding **E**nergy **S**imulation **TEST**
BCF	**B**IM **C**ollaboration **F**ormat
BCF-API	**B**IM **C**ollaboration **F**ormat - **A**pplication **P**rogramming **I**nterface
BEM	**B**uilding **E**nergy **M**odeling
BIM	**B**uilding **I**nformation **M**odeling
BKI	**B**au**K**osten*i*nformationszentrum Deutscher Architektenkammern
CAD	**C**omputer-**A**ided **D**esign
CAPEX	**CA**pital **EX**penditure
CEP	**C**omplex **E**vent **P**rocessing
CIM	**C**ity **I**nformation **M**odel
CityGML	**City** **G**eographic **M**arkup **L**anguage
CHP	**C**ombined **H**eat and **P**ower
CCHP	**C**ombined **C**ooling **H**eating and **P**ower
CO	**C**itizen **O**bservatories
CSV	**C**omma **S**eparated and **V**alue
DaaS	**D**ata **as** **a** **S**ervice
DB	**D**ata **B**ase
DCP	**D**istributed **C**omputing **P**latform
DG	**D**irect **G**eneration
DHN	**D**istrict **H**eating **N**etworks
DWG	**D**ra**W**in**G**
EIFER	**E**uropean **I**nstitute **F**or **E**nergy **R**esearch
EDF	**É**lectricité **D**e **F**rance
EPFL	**É**cole **P**olytechnique **F**édérale de **L**ausanne
ETL	**E**xtract **T**ransform **L**oad
EU	**E**uropean **U**nion
FID	**F**eature **ID**entification
FIWARE	**F**uture **I**nternet platform
FM	**F**acility **M**anagement
GEri	**GE**ne**ri**c enablers
GIS	**G**eographical **I**nformation **S**ystems
glTF	**g**raphics **l**anguage **T**ransmission **F**ormat
GML	**G**eographic **M**arkup **L**anguage
GHG	**G**reen **H**ouse **G**as
GWPV	**G**lobal **W**arming **P**otential **V**alues
GRACE	**GRA**duate school for **C**limate and **E**nvironment
GUID	**G**lobally **U**nique **ID**entifier
GWPV	**G**lobal **W**arming **P**otential **V**alues

HDD	**H**eating **D**egree **D**ay
HTTP	**H**yper**T**ext**T**ransfer **P**rotocol
ICT	**I**nformation and **C**ommunications **T**echnology
IAI	**I**nstitute for **A**utomation and Applied **I**nformatics
IMK	**I**nstitut für **M**eteorologie und **K**limaforschung
IoT	**I**nternet **o**f **T**hings
IPF	**I**nstitut für **P**hotogrammetrie und **F**ernerkundung
ISO	**I**nternational **O**rganization for **S**tandardization
IT	**I**nformation **T**echnology
IWU	**I**nstitut für **W**ohnen und **U**mwelt
KIT	**K**arlsruhe **I**nstitute of **T**echnology
KVP	**K**ey **V**alue **P**air
KPI	**K**ey **P**erformance **I**ndicator
LOD	**L**evel **O**f **D**etail
NGO	**N**on **G**overnmental **O**rganisation
OECD	**O**rganisation for **E**conomic **C**o-operation and **D**evelopment
OID	**O**bject **ID**dentification
OGC	**O**pen **G**eospatial **C**onsortium
O&M	**O**bservations & **M**easurements
OPEX	**OP**erating **EX**pense
OSF	**O**pen **S**oftware **F**oundation
OSM	**O**pen **S**treet **M**ap
OSS	**O**pen-**S**ource **S**oftware
OWS	**OGCW**eb **S**ervices Common Standard
PCA	**P**rincipal **C**omponent **A**nalysis
PID	**P**ersistent **ID**entifier
PV	**P**hoto**V**oltaic
REST	**RE**presentational **S**tate **T**ransfer
QoL	**Q**uality **o**f **L**ife
SAR	**S**ynthetic-**A**perture **R**adar
SAS	**S**ensor **A**llert **S**ervice
SCIS	**S**mart **C**ities **I**nformation **S**ystem
SCP	**S**mart **C**ity **P**lanning
SDI	**S**patial **D**ata **I**nfrastructure
SensorThingsAPI	**S**ensor **T**hings
SES	**S**ensor **E**vent **S**ervice
SensorML	**S**ensor **M**odeling **L**anguage
SHP	**SH**a**P**efile
SOAP	**S**imple **O**bject **A**ccess **P**rotocol
SOM	**S**elf **O**rganizing **M**aps
SOS	**S**ensor **O**bservation **S**ervice
sqkm	**sq**uare **k**ilo**m**eters
SQL	**S**tructured **Q**uerry **L**anguage
STEP	**S**tandard for the **E**xchange of **P**roduct **D**ata
SWE	**S**ensor **W**eb **E**nablement
UM	**U**rban **M**odeling
UBEM	**U**rban **B**uilding **E**nergy **M**odeling
UESM	**U**rban **E**nergy **S**ystem **M**odeling
ULPGC	**U**niversidad de **L**as **P**almas de **G**ran **C**anaria
UML	**U**nified **M**odeling **L**anguage
URL	**U**niform **R**esource **L**ocator

URI	**U**niform **R**esource **I**dentifier
UUID	**U**niversal **U**nique **ID**dentifier
VBK	**V**erkehr**B**etriebe **K**arlsruhe
VCS WFS	**V**irtual **C**ity **S**ystems **W**eb **F**eature **S**ervice
WCS	**W**eb **C**overage **S**ervice
WFS	**W**eb **F**eature **S**ervice
WMS	**W**eb **M**ap **S**ervice
WNS	**W**eb **N**otification **S**ervice
W3C	**W**orld Wide Web **C**onsortium
XML	e**X**tensible **M**arkup **L**anguage
XML-RPC	e**X**tensible **M**arkup **L**anguage-**R**emote **P**rocedure **C**all

List of Symbols

A_i	is the component surface area	$\mathrm{W\,m^{-2}}$
A_P	area represented by each surface point	$\mathrm{m^2}$
ED_i	Annual heat energy demand for building i	kWh
B_{Pi}	Beam and diffuse components of solar irradiance for a point P at time step i	$\mathrm{W\,m^{-2}}$
C_{air}	heat capacity of air	$\mathrm{J\,K^{-1}}$
c_{pstr}	Specific heat of structural material	J/kg K
c_{pwater}	Specific heat of water	J/kg K
c_{sub}	Specific heat of substrate material	J/kg K
C_{vw}	Convection to/from the vegetated façade	W/m2
CO_2Ee	Energy related GHG emissions in Carbon Dioxide equivalent	(tonsCO_2e
d	distance	m
D	Subtract thickness	m
D_M	number of days per month	
D_{Pi}	Beam and diffuse components of solar irradiance for a point P at time step i	$\mathrm{W\,m^{-2}}$
ED_{hl}	energy demand corresponding to the heat loss of the building	W h
EF_i	Emission Factor for heating source of building i	kg/KWh
f_t	factor for the temperature zone	-
G_B	global radiation on a building	W h
G_P	global radiation on a point	W h
G_S	global radiation on a surface	W h
H_D	heating hours a day	
HDD_a	Heating Degree Days per year	K h
L	structural thickness	m
LR_{vw}	Longwave radiation to/from the vegetated façade	W/m2
n_A	building air exchange rate	ACH
NF_i	Number of floors for building i	-
P_{hl}	building heat loss	$\mathrm{W\,K^{-1}}$
P	power	$\mathrm{W\,J\,s^{-1}}$
RC_i	Refurbishment cost for building i	€
Q_{vw}	Heat conduction through the wall behind the vegetated façade	W/m2
S_{vw}	Heat stored in the wall behind the vegetated façade	W/m2
SR_{vw}	Shortwave radiation to/from the vegetated façade	W/m2
t	Time	s
T_0	daily average temperature per month	°C
T_i	set temperature inside the building	°C
U	thermal transmittance	$\mathrm{W/(m^2\,K)}$
UA_i	Usable area for building i	m^2
V_A	inner building volume	$\mathrm{m^3}$
XR_{vw}	Radiative exchange between the bare façade/soil and plant layer	W/m2

Δt	time step length	h
ω	angular frequency	rad
ρ_{pstr}	Structural material density	kg/m^3
ρ_{pwater}	Water density in substrate layer	kg/m^3
ρ_{sub}	Substrate material density	kg/m^3

Dedicated to those who made it all possible: my Family and my Teachers

Chapter 1

Introduction

From Kupe (Polynesian/Maori) to Zheng He (Chinese), Ahmad ibn Majid (Omani), Ferdinand Magellan (Portuguese) or James Cook (English), all pioneers understood, used and respected maps. With their maps they planned the course of their lives and of millions of others that followed in their steps. Today the questions we want maps to answer are ever more complicated and have evolved, from finding the shortest route to spice producing lands to calculating solar radiation potential on vertical surfaces or finding the best location for electric car charging stations inside cities. To handle this changing environment, map makers and data providers have to provide an evolution of the backbone that maps are built on. This thesis provides an in-depth look at how spatial data models associated with an entire infrastructure chain can support energy city planning and management activities by means of 3D city digital twins. The cyber versions of cities are linked to near-real live data and showcase a *living, breathing* (by way of sensors and things that sense) digital copy of the built environment that allows for modeling and simulation of the energy chain in cities. This work allows for shared immersive decision making to happen by accelerating data sharing among stakeholders.

1.1 Motivation

Cities are great centers of human capital concentration. They have always been places that attract, gather, facilitate and drive innovation. To do this they help people congregate and foster companies development. This happens as long as a sum of factors transpire simultaneously. These factors are generally related to either geography or governance. A successful city is one that provides good quality city services that convince people to move in and stay. According to [175], one century ago 10% of all people used to live in cities, where as today, that ratio has exceeded 50%. This trend, slowly decelerated by today's pandemic, will nevertheless accelerate towards the mid of this century and is projected to stabilize in 2100 at around 85%. The previously cited report argues that cities have always existed and expanded because of the economic growth and services that benefit their residents. These services provide for a better quality of life at a lower economic cost due to economies of scale. These can only be achieved in urban areas with a corresponding human settlement density. On the other side, large numbers of humans bundled up in cramped population centers bring about issues of its own. These are related to quality of life, city services and space sufficiency which urban planners and managers have to deal with. The indicators, [175], used for the comparison of cities are:

- Livability – clean, healthy living, digital infrastructure for city services

- Workability – enabling infrastructure: energy, connectivity, computing, essential services
- Sustainability – future generations
- Learnability – quality learning for all

From a mayors perspective, [190], cities are built on three pillars: general public services, citizens satisfaction (it represents a measure of quantifying the first pillar) and narrative storytelling. The first pillar is the foundation pillar, as it allows for the others to exist and includes public utilities, transportation, hospitals, schools and cleaning. This thesis focuses specifically on public utilities and services related to energy in cities. It aims to use data modeling in energy chains in cities and follows energy from production to transportation of the commodity and finally to consumption.

In the attempt of providing an ever better quality of life, the built environment (human-made) has become the world's largest energy consuming sector, constituting 64% of global primary energy, and producing 70% of the world's carbon dioxide (CO_2) emissions, [106]. To alleviate these issues, significant innovation and transformation of the way in which energy demand and consumption happens in urban environments is required, a statement also made in the preamble of the New Green Deal pushed by the European Commission, [71]. This thesis aims to facilitate this change by improving the handling of energy-related data. This was done with the development of models and standards, essential tools that support these changes.

Smart cities

Most often these changes in city energy data management and data handling are identified as a primordial requirement, which is part of the new generation cities (often also referred to as smart cities) revolution. In interviews with different city stakeholders (the people responsible for planning and management of cities), [104], three depths of the smart city trend setters were identified:

- Tools-economic dimension
- Administrative-procedural dimension
- Governance dimension

From the city perspective, for change to happen there is a clear need to improve and coordinate the collection, processing and connectivity of existing (and/or available) information in order to achieve efficient resource management (tool/instrumental-economic dimension). Also required are technical specifications for objectives and actions for city development (administrative-procedural dimension) and new city stakeholder interaction constellations (governance dimension). This thesis tackles parts of the dimensions mentioned for next generation cities trying to offer an integrated perspective from both a spatial but also an energy thematic viewpoint of the city wide energy chain.

Both scientists and practitioners of the urban management have shown heightened interest in the topic of smart cities, being very visible in standardization organizations. This standardization work is being carried by the OGC (Open Geospatial Consortium) - with the Smart Cities WG (Working Group) and the ISO (International Organization for Standardization) [111] with the Smart community infrastructures

- Specification of multi-source urban data integration for smart city planning (SCP) group. The work pushed by these communities is also joined by volunteer based committees that work on standardization within the intertwining geospatial and energy domain, namely the Energy ADE, [22] and the Utility Network ADE, [1]. A definition of the term smart city is provided in Chapter 2 as part of the trends that ever changing cities are currently experiencing.

GIS and the new generation cities evolution

To bring further clarity to the contribution, the question of what does GIS (Geographical Information Systems) provide with in energy urban planning and urban management is asked. And further, what motivates state entities in building extensive GIS systems? Historically, it appears that the original reason was financing. Every budget starts from land taxation, or, in other terms, cadastre. It is the basis for any budget planning as it is the safest form of income (land does not grow or shrink in size). In cities, GIS is needed to simplify spatial complexity because it presents information in situational context which most often explains complex entity relationships. These systems have evolved from taxation to providing spatial repositories of multiple types of information. Once such a system is built it needs to be maintained and constantly refreshed so it stays up to date. It means that cities most often have a department with the sole focus on maintaining such spatial data for use in city management and urban planning. GIS repositories often contain information on the building stock but also for near real time event monitoring. This allows for trends to be observed and measures to be implemented when they are needed.

According to [182], quality GIS data systems and tools provide cities and specifically their municipal governance with increased governance capacities. Also, latest technology delivers two-, three-, four-, ..., n-dimensional management and visualization of cities and their incremental change over different periods of time. Change detection and change impact functions aid architects, urban designers, and public and private stakeholders in seeing ways to think beyond their individual projects and understand the impact on an entire district, city, or region. This also transforms the change process, as it tends to make it more democratic and exposes it to a larger cross section of city residents.

Energy, infrastructure and the city wide energy chain

Energy is a primary commodity which impacts real and perceived life quality. Most often it also impacts decisions taken by people and companies of whether to move to or away from a city or region. At the same time the types of energy, its diverseness, reliability and availability to people and companies can increase the chances for success in the never ending competition of cities. Examples of this influence abound in history. [16] uses a concept called the "rank clock" to show the impact on the growth of cities in the south of the USA during industrialization. Cities lacking cheap access to energy were, in the long term, dislodged from the largest cities in the US. It is this impact that helps motivate the work carried out in this thesis.

Having abundant energy is not enough. The distribution of it via energy infrastructure is a major city service. It represents the backbone of all other services and it is the role of the public sector to facilitate it. City infrastructure can have very different dimensions, be it physical (e.g. roads, electrical networks) or human (e.g health or education sectors). At the same time, it is expensive to build and will often not be erected by private entities without public financial incentives.

The world's population expected to reach 8 billion by 2025 and 9 billion by 2040. As the United Nations recognizes this worldwide trend towards urbanization, [198], prosperous, healthy and resilient cities are key in this development. Keeping the same quality of city services while providing them to a higher number of recipients requires ingenuity and determination. Of course this trend puts pressure on the public sector to provide adequate infrastructure and resources to cities and their citizens, pushing cities to improve and innovate (there are countless studies comparing cities livability, sustainability, "smartness").

The work of [61] analyzes a speckled (large variation in population size and geographical indicators) dataset of 274 cities in order to demonstrate that economic activity, transport costs, location, and urban form explain 37% of urban direct energy use and 88% of urban transport energy use. With the underlying trend in urban expansion set to continue well past 2050, world urban energy use will increase three times, from 240 EJ in 2005 to 730 EJ in 2050. At the same time, keeping limited resources in mind, increasing the efficiency of energy use is, a low hanging fruit of sorts, as it contributes towards increased stability and reliability of the energy chain inside cities.

The solution presented in this study involves the use of detailed 3D semantical city models that serve as a basis for different domain data coupling. The motivation for cities and companies involved in the city wide energy chain to build such spatial models and integrate them with energy data originating either from production facilities, from transportation entities or consumers, differs. From a cities' perspective, public funding needs to create opportunity both in terms of jobs, but also in terms of sustainability and long term development. The main reasons behind such expensive endeavours are:

- The enablement of smart (energy, water, waste) applications, essential in becoming a smart city
- Public engagement in urban planning and management
- Improved urban planning
- Improved city services
- Utility networks management
- Tourism - boost tourism by displaying cities landmarks online
- Fostering economic development

From an energy utility company perspective, the motivation to pursue smart city initiatives relates to, first and foremost, improved services, while also keeping or reducing operational costs. To that end these companies seek to:

- Accurately and precisely estimate energy demand
- Forecast trends in energy demand and prepare for them
- Partake in the urban planning process, so as to advise and implement the best solutions available for the future clients

- Improve their services, especially in the context of free markets
- Better their management of utility networks
- Access live information already available in smart homes and devices
- Foster economic development
- Reduce their GHG (Green House Gas) footprint

Spatial context

As concluded in [62], the spatial context has big weight in the city wide energy chain and a spatial approach is fundamental. This helps build an exhaustive information system concerning the energy performances of buildings in the urban context. This in turn is used in modeling and simulation by energy providers to estimate energy demand and quantify hourly and spatial variation.

The current proposed solution has been tried out in different geographical contexts. While performing this analysis different organizational models of cities have been observed as well as divergent citizens' expectations. The implementation was tested in distinct legislation and requirements under which companies operate. In this thesis examples related to the following cities were provided: Singapore (Singapore), Berlin (Germany), Geneva (Switzerland), New York (USA), Karlsruhe (Germany), Torino (Italy) and Bucharest (Romania). These cities have different urban sprawls, densities, historical and cultural backgrounds. This is reflected in the spatial and temporal scales at which information is stored and organized.

1.2 Research questions

The following questions were formulated and grouped into four thematic blocks in order to clearly frame the scope of this thesis:

- City wide energy chain:
  What is the city-wide energy chain, and how is information exchanged within this conglomerate?
- Spatial scale in the dissertation context:
  What is the correct scale for modeling and monitoring the urban energy chain?
- Standardization and interoperability for collaborative urban management and urban planning:
  How to achieve interoperability in the management and planning of the urban energy chain?
- Energy domain modeling and simulation:
  How to efficiently model and monitor the urban wide energy chain using spatio-temporal data?

1.3 Outline of the thesis

This section presents a description of the thesis layout.

In Chapter 1, Introduction, the motivation behind the thesis is defined as well as the research questions that are answered in this thesis. In Chapter 2 the City Wide Energy Chain is defined, by presenting the actors that form it. Additionally the very different perspectives that they take on the energy chain and the ways in which they approach geospatial information are presented. Further, the commodity flows are presented, from production to consumption and explored within the current thesis limitations. Chapter 3 presents the spatial context and spatial actors that intertwine at this specific extent. Additionally it discusses issues stemming from the lack of standardized data units and proposes a scale based persistent identifier to manage this issues. Chapter 4 presents the spatial data domains that intertwine in the given context and compares their different approaches to the modeling and sharing of data in this context. In Chapter 5 work related to the coupling of sensors and 3D models is presented. This includes three experiments for data transmission and linking. Furthermore, best practices for spatial data actors are proposed and described in order to achieve collaborative urban management and urban planning. These principles should facilitate work in the city wide energy chain irrespective of the scale at which the actors operate. Data fusion in this context is also explained, with two opposite approaches illustrating spatial actor handling methodologies. In Chapter 6, energy domain modeling and simulation in the city wide energy chain is explained, as well as results from this thesis modeling and simulation work using the previously defined principles, infrastructure and methodologies are presented. Chapter 7 displays the conclusions of this work. Lastly, an outlook for the future of semantic city models and the urban energy chain is provided.

1.4 Why We

In certain sentences and phrases of this thesis the pronoun We is used to describe the authors work. The author of this work believes that such a manuscript is never a single individuals work but it is always team work that helped build, model, direct the methodologies, results, conclusions and findings. A similar example of this philosophy is Diana Nyad, who swam, alone, from Cuba to Florida - the longest recorded unassisted swim. She testifies that her work, a single individual in the water, in what is an incredible feat of human resilience, is actually nothing else than team work.

One can argue that even the main structure of support for this PhD student at the University of Karlsruhe, (KIT - the Karlsruhe Institute of Technology), the GRACE (GRACE - GRAduate school for Climate and Environment) program has a similar perspective and recognizes the need for team work in achieving results like the thesis presented here. This student received support from three parties, trying to balance out the conflicting interests affecting the thesis work, be those scientific, work or personal life related. These were: the supervisor in charge, the direct/personal advisor and the group manager, and also institutional from three institutes, EIFER (the European Institute For Energy Research), IPF (Institute for Photogrammetry and Remote Sensing) and IAI (Institute for Automation and applied Informatics).

Chapter 2

Defining the city wide energy chain

With this dissertation study the focus is on a complex mechanism which relates to public services delivery, namely production, delivery, storage and use of energy that cities deliver to their inhabitants in an urban spatial environment. The current chapter describes the city wide energy chain from the perspective that it is referred to in this thesis. This envelope helps prepare the chapters that follow, supports the solutions proposed and confines their scope. In order to define the city wide energy chain the concepts in it are separated as follows: city, energy and chain. Different definitions are explored for each of those 3 concepts and explained within the scope of the methodological approach by defining its viable spatial and thematic context.

2.1 Spatial and thematic boundaries

2.1.1 Cities

Let us first define the spatial context, which is the city, or the urban spatial environment. Historically, cities appeared first as locations of commerce, government and transportation hubs and also as places where large densities of people live and work. According to the United Nations, [222], urban areas are defined differently from country to country, a single definition has not been agreed upon. Additionally, the geographical extent of a city also lacks a uniform international criterion for boundaries. Habitually a city can have multiple distinct perimeter concepts:

- "city proper", using the administrative boundary definition,
- "urban agglomeration", where the extent of the contiguous urban area, or built-up area construct is used,
- "metropolitan area", in which the number and depth of economic and social interconnections of nearby areas (assessed with economical or transportation indicators) is employed.

Another study, [198], uses absolute values to set thresholds for the definition of urban settlements: the population of the area is at least 2000 people living in a dense settlement, while for other countries this value is set at 5000. Regarding population density, this value tends to vary greatly. The European Commission, [66], defines urban centres as containing a minimum of 1500 persons per square kilometer, a definition to which the authors of [193] concur.

Both of the above definitions are used to determine the urban areas (or urban agglomerations) studied in the present work as: **Places where people live together in a high density urban sprawl that allows city services to be delivered in networks**.

The urban sprawl has a minimum density of 1000 persons per square kilometer and a minimum size set at 5000 individuals. Diversified energy networks are unattainable in rural areas due to larger distances and lower population density and, most importantly, the associated costs. **The thesis scope is further narrowed to the administrative boundary of the spatial entities defined in the chain subsection below**.

2.1.2 Energy and networks in city public services

To further define the spatial and thematic context, a rigorous review of the public services that deliver energy inside cities is performed in this sub-section. In the physics realm, energy is, [41], the capacity for doing work. It may exist in potential, kinetic, thermal, electrical, chemical, nuclear, or other various forms. Energy, after it is transferred, is designated according to its nature (e.g. heat transferred becomes thermal energy).

According to [222], cities are responsible for 75% of global energy consumption. This energy is provided from a mix of all possible national/world and local sources (coal, crude oil, natural gas, nuclear or renewable), [221]. Cities have traditionally had their own production facilities. This is done to complement the national grids and also help stabilize local grids. At the same time it reduces the dependency on external factors. However, in the last years discussions have started on creating national-grid independent cities. This conversation has stayed in general in the realm of planning, with few cities showing interest in achieving near or complete independence. This is mostly due to the high costs associated with this action and, surprisingly, the green house gas footprint that such an action would entail.

City energy services achieve one important goal, and that is to maintain and improve real and perceived quality of life of the people living and/or working in their perimeter. To support quality of life these city services deliver energy in the form of electricity, heat, cold or another energy carrier, with different end applications connected to each of those networks. The delivery of these very precious commodities takes place trough very different network types. Through a user survey among thematic experts and city project managers carried out in the EIFER institute in 2016 city networks that serve as use cases for the development of this thesis were identified. The survey results were presented to a panel of experts at the Utility Network Consortium meeting in December 2017 and together with the expert members an expert reviewed systematic survey of network types in cities was completed, [126]. The following types of energy are, as documented in the working minutes, transported in the cities' utility networks:

- Electricity
- District heating
- Gas
- Waste water*
- Storm water*
- Fresh water*
- Communication**
- Process steam

- Oil
- Waste
- Air pressure

* when used for energy production ** utility networks usually integrate communication tools, or use communication networks for their own data transmission

Out of all energy goods delivered in cities, electricity is the most readily available commodity to all population types. It is also the only one for which world wide database numbers exist (United Nations, International Energy Agency and World Bank), see 2.1. Specific to urban environments, according to the IEA 2017 data, [105], as of 1991, 97% of the world urban population (or 2.28 billion urban dwellers) had access to electricity. Since then, that percentage decreased for a decade (even though in absolute numbers the value increase remained), a process most likely caused by the main communist socialist block disintegration (with large numbers of people leaving rural communities) and the lack of state investment that followed. However, considering the accelerated urbanization trend mentioned in 1 and visible in the numbers presented in table 2.1, the new arrivals to urban environments have moved to cities that already had this resource readily available.

Moreover, regional and national governments strive to provide cities first with access to reliable electricity. The reason is that the main factor in electrical network development is population density, [24]. These trends and policies are further confirmed by numbers from 2017 where the same percentage of 97% of the world urban population is shown to have access to electricity (while the relative number was constant, the absolute value increased in the 25 years span that the table depicts). In any case, this numbers make electricity the most readily available energy resource out of all the types identified in the previous list.

TABLE 2.1: Electrification trends in urban population centers and population, source [105], [220]

Year	Total world population [billions]	Urban population [%]	Urban population [billions]	With access to electricity [%]	With access to electricity [billions]
1991	5.41	43.29	2.34	97.19	2.28
2001	6.22	47.1	2.93	94.72	2.77
2011	7.04	52.11	3.67	95.95	3.52
2017	7.55	54.85	4.14	97.36	4.03

As it is such a coveted commodity, the spatial coverage of electrical networks is not necessarily explained by supply and demand economical mechanisms, meaning that different public stakeholders heavily subsidize the process of network development, thus providing a larger geographical extent than would be available if only spatial parameters would be used (as defined above). This creates geographically unbalanced networks which cover areas that would normally not justify the connection to a national grid. Gas and oil networks in certain countries follow similar patterns as they are prized commodities both for the general population and industrial activities. With most of the other commodities, the price of network expansion plus maintenance and rapid energy losses associated with distance generally act as balance regulators. It keeps such networks in the urban domain or at least to naturally relevant geographical extents.

In cities, energy networks are sometimes physically visible, sometimes inconspicuous and noticeable only through substations (or *"middle men"* such as gas stations), and represent the lifelines of cities as one of the main facilitator to quality of life (QOL). These can be viewed like veins and arteries of a human body that take precious blood (energy) to the farthest corners of the system and (sometimes) bring it back to the power plant (the human heart, e.g. district heating networks). In a similar fashion energy networks are described as the lifelines that cities use to take and bring energy to/from its inhabitants.

2.1.3 Energy chains

In a conservative way, the energy chains in a city can be defined as composed of three separate levels: production, utility (or ubiquitous) networks for energy carrier transportation and consumption. Most often energy comes from outside of the city, with the alternative being that energy is produced inside the city either from centralized facilities or even by prosumers (someone who both produces and consumes energy). The visual description of the energy flow is described in Figure 2.1. In it, the flow is presented using spatially separated entities. In cities the energy source could be fossil, renewable, waste energy from industrial processes and in rare cases nuclear.

The energy produced is either consumed next to the production facility or transported. In an overwhelming majority of cases energy is transported, either via utility networks or vehicles. City utility networks used for energy transportation can be of multiple types: electrical, gas, district heating, district cooling, waste water, steam and oil.

The last level in energy modeling in the city wide energy chain are the consumers. One way of classifying these consumers is as residential, tertiary and industrial. These consumers can also be producers, as mentioned before.

To explain energy chains, a methodology described by Batty in [16] was adopted, which has cities depicted as chains of entities that can be represented with a hierarchical structure. Stocks and flows are used to describe in general how people connect different points in the cities by commuting from one place to another. A similar understanding is adopted in this manuscript with regard to energy chains, which connect energy production centers, delivery centers and final consumers inside the urban space. Batty goes further and defines cities as complex systems and uses system dynamics to explain the complex processes that go on in cities.

With regards to the entities that compose the energy chain the data modeling in this dissertation is confined to all static entities that are encased in the built environment. One can argue that not all energy distribution or consumer objects are spatially static, and therefore not encased in a building. This can be seen for example with trucks that distribute oil products from a gas station to houses. However, the energy which is invested in the final consumer is always modeled in a physical building, where the energy is drained from the mobile entity. This means, the energy consumed will always be modeled in a static physical entity, namely a building.

To continue in explaining the energy chain, the entities that compose it are described. For the purpose of city modeling all the following entities are housed in buildings, or better said, in the built environment:

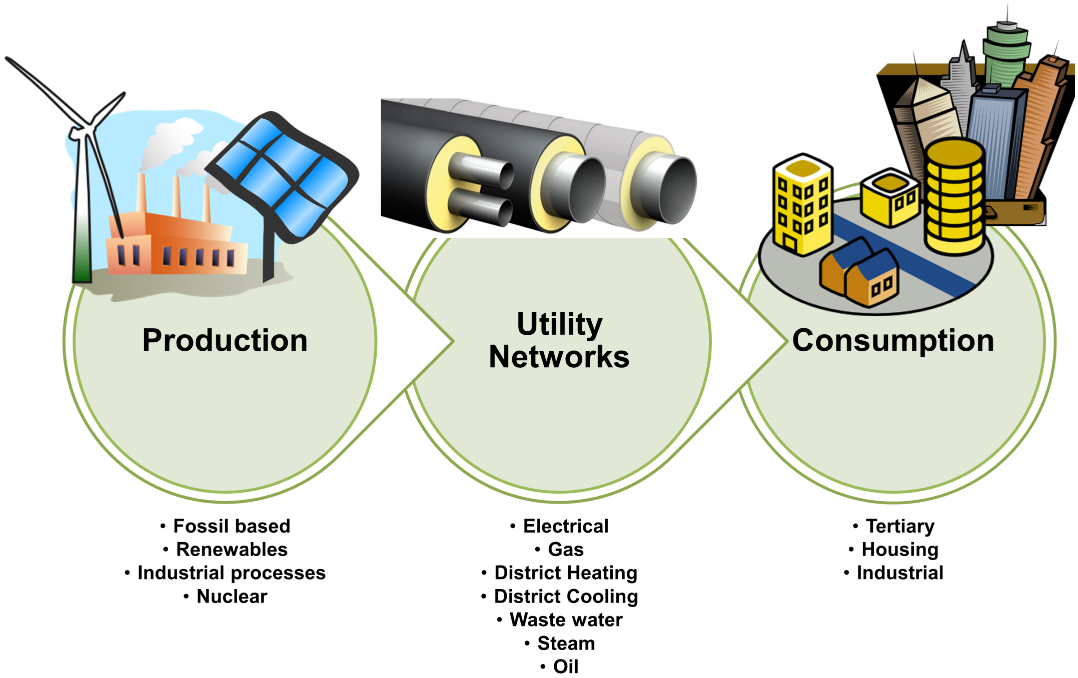


FIGURE 2.1: A conservative perspective of the city wide energy chain, energy flows from left to right, producer to consumer

- energy production centers: built entities that produce or accommodate energy production capabilities (inside or outside)
- the primary network: the energy transmission lines, performing the energy transfer in between production centers and substations
- substations or delivery centers: components of the energy chain that transform the properties of the energy commodity (e.g. lowering voltage for electrical pipes, lowering pressure for district heating networks) to a level that can be used by the final consumer.
- the secondary network: the energy transmission lines that perform energy transfers in between the substations and the final consumer
- storage entities: facilities that store energy, these refer to stand alone entities, as storage can occur most often on the premises of the final consumers
- final consumers: the end energy consumer is defined as a consumer building, with either residential, tertiary or industrial focus.

The classification presented above is rather traditional in the sense that it is uni-directional, energy flows one way to an end consumer. However existing situations that also motivate this thesis (some of them facilitated or brought up by the energy transition) allow end entities to produce energy themselves in order to use, store or sell that commodity back in the chain (e.g electricity from PV panels or heat from boilers). This also reduces the need for energy transportation in the primary networks and increases it in the secondary networks. This change is sometimes triggering the need for additional secondary networks to be built. Additionally, in cities, energy experts have an eye out for sector coupling - the use of residual energy from energy-intensive heating, cooling from (sometimes) industry operations. The entities that compose the energy chain are depicted in figure 2.2 with light blue depicting

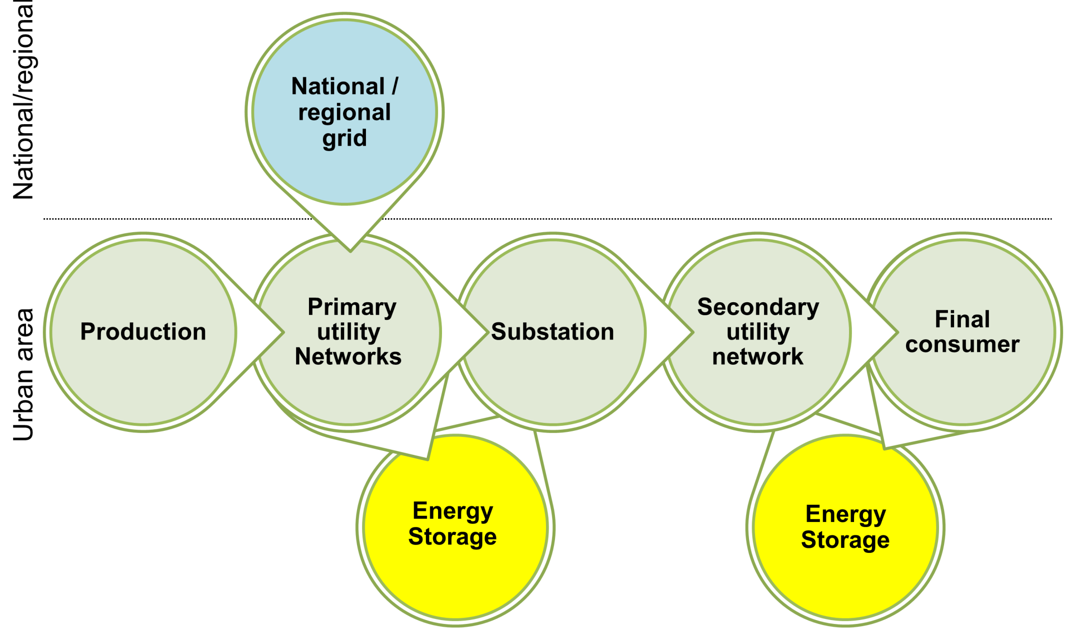


FIGURE 2.2: The energy chain inside urban spaces with its components and flows of energy, special notice to newly added flow directions

national/regional grid input, green the traditional chain components and yellow the most recent additions to the city.

Energy chains help further define the spatial extent of the study areas. This thesis is purposely confined to exclusively study directly (physically) connected entities in the urban domain. For example, an electrical network connected from local source to end user, represents the maximum extent of urban space as described in this manuscript. This extent respects in most cases the administration boundaries of cities but does not constitute a rule. As such, the approach has to consider both administrative boundaries and those of energy utilities.

2.2 The ever changing city - trends in cities and the energy chain

One cannot discuss the city wide energy chain and present solutions to today's problems without understanding the current trends and disruption technologies affecting the different fields and actors intertwining in this context. Furthermore, data from different, intertwining fields is modelled: GIS, energy and IT. Very often domain experts propose solutions from their own field of expertise and perspective. In this thesis a holistic approach that handles the connections in the most efficient and non-partisan way possible is used.

2.2.1 Smart, intelligent and next generation cities

A most significant trend with cities and their administrators is the smart city movement. The term smart city, generally refers to a city that has undergone a process of digital transformation in the public services it provides to its inhabitants and companies. This transformation impacts urban management and urban planning, changing the fabric of cities in real life not just digitally. In [63] a survey of existing definitions was made. The study also proposed a comprehensive definition. A smart city is:

- a well defined geographical area,
- governed by a well defined pool of subjects, able to state the rules and policy for the city government and development.
- using high technologies such as ICT, logistic, energy production, and so on,
- creating benefits for citizens in terms of well being, inclusion and participation, environmental quality, intelligent development;

Based on interviews and reviews presented in [104], the term Smart City is not different in terms of substance that clearly differentiate it from other recognized naming conventions such as Green Cities, Sustainable Cities or Low-Carbon Cities. What the concept does achieve is a definition of the conversion process. Purposefully used by decision makers in the public domain, it has built acceptance within the mainstream and has thus allowed for funds to be dedicated to smart city investments (e.g. infrastructure investments).

In Europe, as far back as 2007, the cities associated together in the Covenant of Mayors program, [44]. Among their declared aims are the improvement of energy efficiency and increased renewable energy use inside cities. The authors of the previous study, provide the alternate term intelligent cities and mention citizens QoL as the most important indicator with which all initiatives should be rated. They look at cultural, economic and social growth in a healthy, safe, stimulating and dynamic environment as the components of positive effects in peoples lives from smart city initiatives. From the authors perspective the central challenge for cities in this initiatives is the integration of new enabling infrastructures and IoT with pre-existing structures while exploiting synergies and interoperability between systems. This facilitates better value services for citizens, and contributes to an improved QOL.

Taking a stakeholder perspective in [16], Batty explains that interviews with decision makers and interested parties show that they are very much aware of and understand the smart city concept. They also see their own cities current situation as one of competition with other cities. That means that in order to stay relevant, cities need to adopt technologies, integrate them in their decision making process and use them in order to provide a higher QoL that convinces citizens to reside in their city, and companies to operate there. This is only possible when manpower and infrastructure combine to provide the right recipe for success.

More generally, digital transformation can be understood as the use of new and frequently changing digital technology to solve problems often using cloud computing while also enhancing capabilities of traditional software products. It is achieved when the digital enables innovation and creativity and stimulates significant change within the professional or knowledge domain, [120]. Specifically with regards to the digital change that cities undergo, these can be structured in 4 pillars, [55], digital:

- infrastructure, the backbone that allows for public digital services to exist in the first place, including WI-FI, 5G and broadband
- services for citizens, direct online citizen relationship to the public administration.

- education, the restructuring of peoples skills so that all citizens can have basic access to digital privileges
- skills valorification, or the building and development of projects and initiatives, to help spread the value creation from digital footprints both for companies and people

As with many profound societal changes that impact citizens lives, it takes time for the term to be universally adopted and agreed to, so alternatives have emerged to the term "smart city". The same term was used to describe very different and sometimes seemingly not converging activities. Parts of the scientific community choose to distance themselves from it, for example Ursula Eicker, one of the most cited authors in the UBEM (Urban Building Energy Modeling) field, prefers the term 'next-generation city', and defines it by linking it to sustainability - a bridge that unifies our common purpose with these transformation actions, [155].

For this present thesis a smart city and a next-generation city are one and the same. They are defined as: *a city that has embraced digital amenities and provides its inhabitants and companies with digitally optimized and innovative public services. This is done in a way that stimulates and enhances open data initiatives, direct citizen participation in city management in order to guide resources allocation and change.*

Smart districts

Smart districts are the embodiment of a smart city on a district scale. In the European Union context, a knowledge platform was built to exchange data on the topic of smart cities. Called the Smart Cities Information System (SCIS), it allows cities to exchange data, experience and know-how required in order to achieve smart cities, [68]. On the same platform, a definition of smart districts as spatial units that rely on three pillars can be found. These are:

- to make efficient use of energy through smart grids,
- the inclusion of district heating and cooling distribution systems,
- to provide new mobility solutions and the use of ICT.

Using the three components mentioned above energy loads can be locally handled, reduced and shifted from peak to off-peak hours. This maximizes energy load balancing and allows for the integration of renewable energies. Even though the definition is very much "alive" in the sense that it changes as things progress in the field, the current one gives us an idea as to the goals of the projects aiming to deploy smart districts. These are most often projects that do so from an energy perspective or at least include it in the proposed work. In addition the use of ICT can enable energy providers and policy makers to use the monitoring data to guarantee a level of transparency in between all stakeholders. This allows citizens to have an impact and even take responsibility in the smart management of energy in their district. At the same time it often leads to changing habits. In addition, largely automatized systems can decrease administrative loads and diminish the costs associated with network operation.

Direct energy generation and storage at building and district level is also identified in [43] as one of the most important trends in local energy management. Benefiting from lower costs and technological break troughs, direct generation from local renewable sources and energy storage continues to gain in presence. The emergence of f

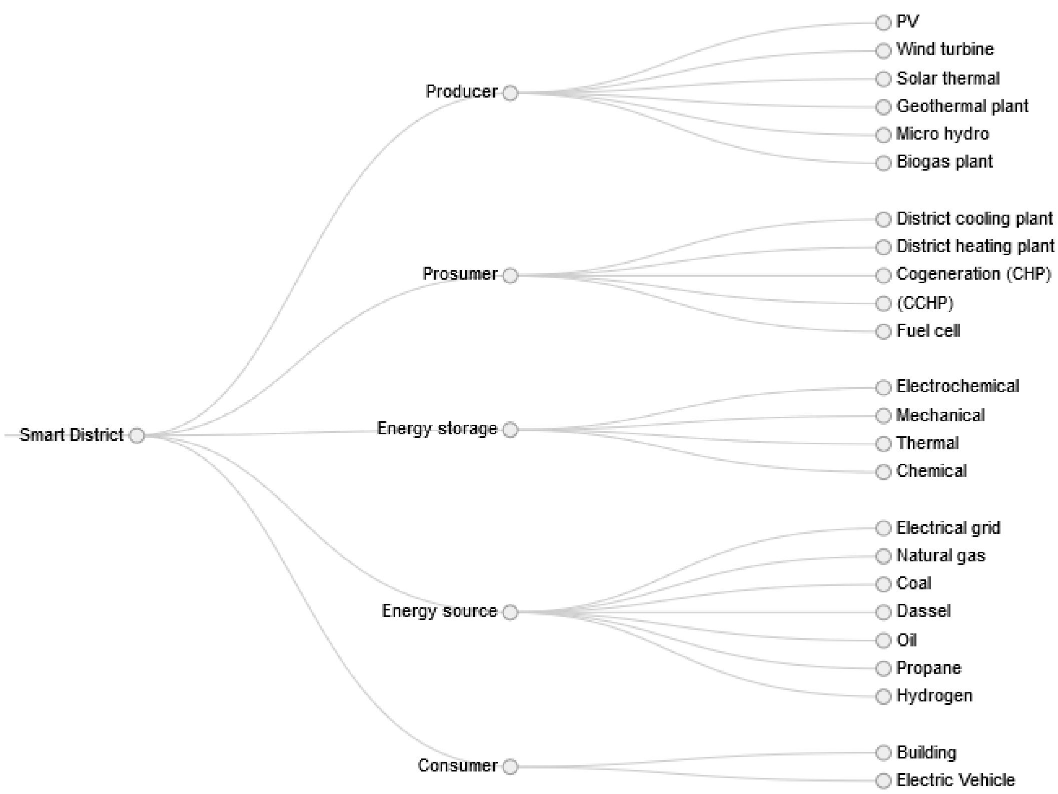


FIGURE 2.3: A renewed perspective of the smart district wide energy chain, adapted with information from [8]

Abbreviations:
CHP - Combined Heat and Power
CCHP - Combined cooling, heating and power

these energy-efficient facilities and their integration into daily lives is what has been identified as one of the most important tasks in achieving next generation cities. Regarding energy transportation, tendencies show that the micro grid and smart-grid paradigms will become the standard in the electricity transportation domain.

The energy chain components at the level of a smart district can be categorized in 5 categories: producer, prosumer (an entity that both consumes and produces a commodity), energy storage, energy source and consumer. These categories and the distinct entity types that are included in them are presented in 2.3.

Open digital governance and open data

Open governance refers to freedom and availability of information. [44] offers the perspective that for citizens to successfully access information and data, inclusive processes for widespread participation are needed. This in turn ensures collaborative decision-making processes and effective value creation from data that the city owns. At the same time ICT offers active and open communication in between citizens and their elected leaders, leading to efficient tax payers money use and higher QOL. One key benefit of open governance is value creation generated from open data. Many city administrators are aware of their data value but not necessarily have the resources or the skills in their budget deprived teams to make use of it. As such, making data public creates opportunities both for the public and private domains to use open data. Of course, not all public data can become open data. The intricate topic of privacy, with different restrictions set in various local, regional and national contexts, requires constant attention and adjustments to data policy. This process of data opening can be quite elaborate and time consuming.

However, open city data has its downsides too, and these come usually in terms of quality and consistency, which affects data processes further down the processing pipeline and most importantly the duration of the treatment required to make data useful. In an attempt to tackle this issue, [169] defined and developed intricate ETL (Extract Transform Load) processes with the help of semantics data and open API called the CitySDK. However, this approach has still to gain notoriety and critical implementation mass. This is why classical standard regression methods in combination with principal component analysis (PCA) used to improve the quality and amount of predicted values are still a common tool within the data scientists community.

Keeping the advantages and disadvantages of open data in mind, one must recognize the increased legitimacy that it provides to city administrators and the fact that it creates a tangible connection of citizens with their own city' institutional representatives, [27]. That is also concurred to by [177] who go further in analyzing the impact of open data in Economy, Education, Energy, Environment, Governance, Tourism and Transportation. Their results show the impact of data initiatives in the 5 smart cities sample they analyzed to be strongest in Governance and Economy. The authors conclude from their analysis that such initiatives provide an inherent innovation pattern which is most important in an "open innovation economy" that brings together citizens, NGOs (Non Governmental Organizations) and companies.

Citizen observatories

The way in which cities are governed is changing in order to adapt to societal changes in terms of life styles, to competition from other cities and to the integration of the digital domain in daily lives. Citizens use a very diverse set of digital applications in their daily lives and expect that their city administrator reacts and provides them with similar ways of information exchange. This data flow needs to work both ways: information is provided and feedback returned. For the scientists involved in studying this phenomena, it is recognized as an established field called citizen science which creates information hubs, called citizen observatories. These observatories, provide further assistance in intricate endeavours such as energy urban planning and management via participatory decision making tools that aim to enable citizens to act as "eyes" of the authorities and policy makers and to monitor land-cover/use changes through everyday activities, [219].

The success of such observatories is already visible in Horizon 2020 projects like SCENT Citizen Observatory (CO), where infrastructure, monitoring systems and legal and administrative frameworks that allow for monitoring and management have been developed, [168]. The methodology used in this project is based on expert judgement (for data quality indicators) and ontological analysis (interpretation), using citizen collected information with the SCENT toolbox. The data quality tests performed by local experts follow the thinking that citizens that are familiar with the specific urban context situations have thematic knowledge and possess an understanding of possible solutions and their implications. This is in turn complemented by a set of Key Performance Indicators that the stakeholders have evaluated. To sum up, COs can be used to enhance communication opportunities and facilitate participatory decision making whilst using local knowledge and human resources rather than seeking top down measures.

2.2.2 Big data and urban digital infrastructure with the Internet of Things - IoT

IoT is identified as early as 1999 as a potential game changer in the context of supply chain management, [7]. IoT relates to one of the aims of this thesis, to show the feasibility of using standardized integrated spatial-temporal data and standards in the energy supply chain. As [213] shows, the Internet of Things can be realized in three paradigms:

- Internet-oriented (middle-ware), Internet is used for data transmission,
- Things oriented (sensors), oriented to measuring devices,
- Semantic oriented (knowledge), linking the measurement and the devices with the help of semantics to things in citizens' surrounding.

Internet availability, the requirement of the first paradigm, is another commodity, usually taken as granted by city citizens and even engraved in citizen rights in some countries. Finland, [17], has even provided a minimum legal speed for nationwide internet access at 1Mbps. 4G internet (60Mbps) has already made IoT possible, however the availability of 5G (22Gbps) is showing to be the tool that allows IoT to achieve critical implementation mass and create what is called the 5G IoT or the Internet of everyone and everything, [239]. It is precisely this availability of data flows in near-real time that creates the so called big data perspective.

So how large does a data set need to be in order to be classified as big data? If the concept of big data is treated from an absolute value perspective it must be mentioned that Bill Gates, the largest stakeholder of Microsoft, considered, at the birth of his Windows operating system, that a data set was considered very large already at 640KB (the big data term did not exist at that date). As of the start of 2019 there is no standard number that the IT community accepts as the threshold for discussions on big data, this varies greatly depending on the skills and tools available to the users in a group. It must also be said that these numbers matter because they form the basis of discussions in between the departments that ensure continuous and uninterrupted data flows (e.g. IT and research groups). As to what exactly qualifies as "big data", this also varies depending on the capabilities of the users and their tools, and their available infrastructure. Ever expanding capabilities and better software make defining big data in terms of information quantity difficult.

At the same time what also matters is the quality of data. Well structured, reliable and well defined data sets are the general standard of data in relational databases. Big data is generally defined as the opposite approach to relational databases, large volumes of poorly structured data that is sometimes provided with gaps and perhaps affected by errors. What these data sets lack in structure and perhaps quality is covered however by simple sample size. These provide scientists and city stakeholders with perspectives that would have otherwise remained unseen and unrepresented.

Within the urban energy chain, data flows remain an issue. It is not as if this data was not produced before, but rather that its flow was and still is largely interrupted.

It is part of this thesis scope to study infrastructure solutions and software that could take advantage of this freshly *freed* energy (or energy relevant) data.

Chapter 3

One scale too many - Divisions of UBEM space

Space is a resource.

This chapter focuses on the components and scales at which information is found in the city wide energy chain. The spatial context is also properly explored, including its dimensions of extent and resolution. As with regards to its components, the use of the word refers to the spatial actors that interact with the energy production, transportation and use. The chapter also includes a classification of this actors. Further, the issue of the right scale for urban building energy modeling is examined and approached from a perspective of spatial relevance for the data used. The issue of spatial data consistency for UBEM modeling is explored together with an assessment of urban fabric contiguity impact. Finally, a scale dependent urban data ontology with persistent identifiers to help with origin-destination relationships is proposed for modeling of the urban energy chain.

3.1 Defining the spatial context

Scale is a fundamental issue in spatial science. In the context of the city wide energy chain different data owners aggregate and maintain data which has an influence on the acquisition, storage, processing, and visualization of data sets from this domain. Scale brings with it the same difficulty in energy modelling. How do spatial actors interact and at what spatial levels? What is the right level of aggregation? Where should we try to gather all information so that the effort is rewarded by enough accuracy? This chapter identifies spatial actors, scales and standards in urban energy planning and management in smart open cities and presents two alternatives used in the manuscript work for spatial scales used in aggregation, one approach that relates all information to buildings and another one to neighborhoods.

3.1.1 Spatial scale as a general problem in spatial sciences

Whether in GIS, in CAD (Computer-Aided Design), or BIM (Building Information Model) or any other spatial discipline, spatial context is intrinsic to the definition of the problem algorithms are seeking to solve. In the manuscript [82], the author stresses the significance of spatial context within many other different fields and applications, such as geography, geology, agriculture, forestry, surveying, water sciences and computer sciences. This chapter makes a case that the impact is as important within the energy domain and defines the spatial scales and the actors that intertwine in order to handle spatial objects in the city wide energy chain.

Similar conclusions with regards to the significance of spatial context in energy planning, modeling and simulation are drawn from the publication of [119]. According to the authors, no specific integrated software for architects, civil engineers and energy designers that combine multi scale indicators such as a neighborhood Key Performance Indicator (KPI) framework, city wide energy simulation and building KPI-optimisation was yet available. The emphasis of the entire effort presented in the publication is neighborhood scale with the authors making a plea for it being the "right" scale at which to perform building energy simulation in the context of a urban area. It is a similar plea that this chapter also supports, seeking a meaningful effort of data gathering, processing, modeling and simulation.

3.1.2 Extent

Scale has many meanings, but in GIS two are of greatest significance: resolution and extent, [84]. With regards to the maximum perimeter or spatial extent of the thesis, this is city wide as defined using the energy chain expanse in chapter 2. That means an urban area where people live together in a high density urban sprawl, where a minimum density of 1000 persons per square kilometer and a minimum size of 5000 individuals is found. Those criteria are paramount in order to allow city services to be delivered in networks (unattainable in rural areas due to larger distances, lower population density and most importantly associated economic costs). Also, the thesis confines itself in the administrative boundary of the spatial entities defined in the chain subsection below.

This scale dimension is reasoned by [136] as the optimal one at which inclusive urban-planning projects that consider full energy cycles cutting across all the presented intervention areas can be performed. Additionally, solitary efforts (e.g. designing and managing independent passive houses and smart buildings) might not be optimal overall, as these cannot facilitate the planning and management of interactions between energy significant actors and their spatial locations. The scale issue is also reflected in the work of [44] who evaluates different smart cities initiatives (from city, regional and nation levels). The work concludes that the scale at which current projects in the topics of energy efficiency in buildings, flexible public transport services, digital infrastructures, smart grids are focused on is still narrow. This is due to the fact that most projects are proof of concepts rather than fundamental change implementation, at city level. As such, building level initiatives are discarded as actual change, and rather as prof of concept implementations.

3.1.3 Resolution

With regards to resolution, this thesis makes a case for using buildings as the basic unit for the specific needs in this thesis. The motivations behind this choice stem rather from a mix of reasons both in the spatial and in the energy domain such as specific data availability, geographical barriers, statistical sample scale significance, and are presented and discussed within this chapter for spatial related issues, and in 6 for UBEM (urban building energy modeling). This choice is significant because the accuracy of the results of spatial analysis methods is linked to both the scale of the data sets involved and the algorithms used. [84] discusses the scale problematic in GIS and its semantic definition in detail and concludes that raster and vector data have to be treated differently.

In raster data, resolution has a direct link to the scale in which the data can be used because of the explicit size of raster cells (pixel or voxel). Spatial urban energy data is sometimes, albeit seldom, provided in raster format. Vector data has no direct resolution compared to raster data. There is no simple definition like the resolution in the raster case.

The vector data's fitness for a distinct scale range has to be defined during its acquisition process. Often, the metadata provides suggestions, if such data is available at all. From within the vector data set itself it is difficult if not impossible to guess scale. Vector data attributes have a scale dependency, too. The detail or number of distinct features classes has a scale dependency which is directly linked to the geometric objects involved.

To tackle different use cases requirements that involve discrete changes, data used in spatial sciences is often provided at different levels of detail (LOD). 3.1 presents the most typical level of detail with regards to data used in the urban domain. Engaging different LODs simultaneously may require different data generation or acquisition strategies. In the urban domain scale, the lowest level of detail may be represented by a digital surface model, meteorological information, air pollution or geothermal applications, with pixels in the order of magnitude of dozens to hundreds of meters. A more detailed LOD with coarse 3D building geometries may be derived by extruding the 2D geometries of building footprints with corresponding building heights stored in a different dataset. The features may be taken from a cadastral dataset, a LIDAR or photogrammetric flight. The higher the level of detail should be, the harder it is to maintain the consistency between the layers [89]. For new structures and buildings, parametric approaches help to enable consistency between multiple scales because they can be generated out of an all digital modeling workflow, [40].

TABLE 3.1: Typical level of detail for data used in the urban domain

Sample description	Data type	Level of detail [m]	Accuracy [m]
Digital surface model	Raster	$10 - 10^3$	0.1 - 1
Air pollution maps	Raster	$10^2 - 10^3$	10 - 30
2D Building footprint	Vector	0.2 - 1	0.1 - 0.5
3D Building geometry	Vector	0.2 - 3 (height)	0.05 - 1
Street infrastructure	Vector	0.1 - 2	0.02 - 0.5
Energy infrastructure	Vector	0.2 - 1	0.02 - 0.1
Underground infrastructure	Vector	0.2 - 0.5 (height)	0.02 - 1

3.2 Spatial actors

What kind of actors do we meet in the city wide energy chain? How can we enhance interoperability among them? To energy data modelers in the city wide context reliable data in the urban planning and management process is extremely valuable. However, at the same time this data is difficult to obtain and rarely standardized, with each stakeholder generally having a narrow perspective of the energy chain, usually an individually centered approach. This slows down the entire data gathering process due to the data treatment and legal procedures required to bring all information and actors together. Spatial data for the city wide energy chain is stored and maintained by legal entities who are first of all entitled (or legally obliged) to do so and whose activity revolves around administration, civil engineering, urban planning or management.

The following list comes from open administrative data provided within a couple of geographical areas that the author has extensive experience. These are depicted as examples in the thesis. These areas are: Singapore (Singapore), Torino (Italy), Bucharest (Romania), New York (USA) and Karlsruhe (Germany). This list is not (and due to the different and unique nature of each city cannot be) exhaustive. It presents the actor and their perspective, first from profession separation:

- Architects and architecture companies. By their profession description, they design buildings and are sometimes involved in the construction process in order to ensure correct building standards. Architects store and work with data in a way that ensures them a functional perspective for the building. Ultimately, an architect aims for beauty and functionality. This often involves the creation of visualization prior to actual construction. This process often uses standards that do not pursue the interoperability perspective. That means semantical city models are rarely used, but rather CAD and BIM related standards with few, if any architects venturing in, and using, GIS domain standards. Due to the often unique nature of the products used, conversions from files created by architects are difficult to complete due to the different perspective taken. For energy use analysis, architects use building scale products such as EnergyPlus with extensions in their own architect oriented software products which focus on the detailed energy estimates of single buildings.

- Urban planner. An urban planner is a specialist that keeps an architectural, economic and political perspective on neighborhood / city development with an ultimate goal of public welfare. As urban planners often work on larger projects, GIS is a necessity but often they are confronted with BIM and CAD input which require costly standard conversions. For the energy chain work this professionals regularly lack proper tools at neighborhood and city scale. This is why a host of research projects have sprung up in the last years to assist in mid and long term energy planning and management. In addition companies like EDF, IBM, Google, Microsoft, Siemens and others have created unique software tools that allow for urban energy planning, see 6 for a presentation of some of those tools.

- Civil Engineers. These are professionals which design, build, destroy or maintain structures or facilities for living, industry and transportation and also do environmental/land improvements works. The civil engineer is a general user of CAD and BIM. A rather large panel of software products exist which support the construction, development, and monitoring of construction works. By the nature of their work engineering products tend to respect standards which facilitates exchanges with the other professionals involved in the city wide energy chain.

- Public Administration / civil servants. These professionals are very often building, maintaining or even designing public works. Most public administrations keep and maintain their own data repositories. These have mostly been converted to a digital format. However, that is not always the case and there are still situations where historical print-ed blueprints are in use, something especially true with utility networks, which often have a history dating back to longer than a century. When dealing with the city wide energy chain these category of professionals are in general the most important data owners / actors in the data chain. They most often control or own the most significant data. In the rare cases that they don't they maintain a network of contacts which facilitates the exchange of information. Standardization in storing this data is rare due to the fact that data is presented to them in many different formats and standards. However, there often exist local or institution based data standards. Standardization of all data in the spatial domain is something that has attracted the attention of strong institutions in the EU and these can be seen by standardization drives like the INSPIRE framework. However, the implementation is an ongoing work with a mid and long term completion goals that this specialists have to accomplish. Regarding software choices both CAD and GIS software are used by these professionals.

- Facility Managers maintain and support the use of built environments. Regarding spatial context they focus on building (rarely also at district) scale. They go in to great detail to store all information related to the smallest uniquely identifiable element that belongs to a facility. Due to the laborious nature of their tasks they tend to be very precise and favor extremely detailed data models like BIM IFC models. Very often, facility managers are involved in optimization tasks for the facilities they manage. In the energy domain that means they maintain a good understanding of the material usage, the occupancy and schedules in usage of their facilities. As shown in [80], building physics and occupancy play a key role in energy demand, which makes these professionals key stakeholders in energy modeling. Additionally, some experts of Facility Management (FM) use sensors and gather data in order to perform detailed analysis in order to support a sustainable use of their buildings. Facility managers are among the first (together with security companies) to ask for and develop smart applications that convert building to *smart* buildings - being trend setters from this perspective.
- Transportation entities professionals (engineers) are transportation operators. They transport people or goods in their networks and, after energy demand in housing, they represent the second highest demand of energy in cities. Due to the larger spatial context that they focus on it is quite common that they use GIS both in data modeling and operation of their networks. There are plenty of software solutions that support transportation planning, maintenance and operation. CAD and BIM is also used when great detail is required, particularly in transportation hubs, but then we go back to civil engineering and facility managers.
- Utility network operators. Professionals working in energy utility network operations are usually engineers focused on the delivery of the energy commodity from the production center to the consumers. These professionals use network specific standards and data models. The operation of such tools require that standardization is respected and maintained, although in most situations internationally recognized standards get a local flavor with specific extensions or entities. It is very often the case that they use both CAD and GIS with some very specific associated data models and software products. The spatial data they use is rarely found in the open domain due to security concerns as most

utility networks are considered critical infra-structure. However, in the last 5 years (the duration of this thesis) there have been some positive developments, both from public funded research projects, [187], and also from cities like Rotterdam (Netherlands) and Nanaimo (Canada). These developments usually relate to electrical networks, with rare samples concerning other network types.

- Energy producer professionals. These are usually engineers tasked with designing, operating and maintaining energy production sites. The spatial standards they use in general are CAD and BIM. The spatial data they own is very rarely shared outside of the companies that own the plants and networks. Exceptions appear when different entities own the utility networks and the energy production site and/or local public entities enforce the sharing of such data. Other relevant data (to the city energy modeling field) that is released by these professionals concerns the amounts and properties of the energy commodities delivered. These impact network sizing and their reliability.

Entities not included directly in the energy chain can also have a significant impact on its management and modeling. These are usually statistical agencies or companies that store data and provide it as a service - DaaS (data as a service). Usually, it is their data that is used in providing UBEM models with input and support for representation of output data.

It goes without saying that the spatial actor list provided in this thesis is a general list, This means that as every city has its own cultural, legislative and historical background the catalog presented here is not exhaustive in terms of numbers but rather tries to be in the sense of categories. Spatial data relevant entities in the city wide energy chain can be unique to a city or country, e.g. the organization that attends the safety of chimneys and heating boilers in Germany, which appeared after accidents with carbon monoxide poisoning from heating installations based on fossil fuel burning. Today the same organization maintains a database with all heating sources for every household in Germany and performs yearly checks for safety in their operation. Recognizing the differences in between different geographical contexts is of the upmost significance. The specifics of each area carry significant weight and this is why scales, actors and standards have evolved naturally and differently in different areas.

3.2.1 Classification of spatial actors

Spatial actors, or spatial data owners can be classified with the scale at which they are active. The spatial scale is illustrated in figure 3.1. The main takeaway from the table is the fact that all spatial actors meet at the district scale. This underlines the significance of this scale in the city wide energy chain as it facilitates information exchange required for UBEM modeling while also carrying sufficient weight in the city system. This means that it is the first scale at which decision making can have meaningful impact and it also the place where data is already in the hands of participating stakeholders.

As the table also shows, multiple stakeholders can store information on the same spatial entity. This brings the issue of data consistency among spatial actors, as the same attribute can have different meanings, values and accuracy in different repositories. Although it may appear to be a common issue, in practice it is not for two straightforward reasons. First, real life data sets are incomplete and second, UBEM

FIGURE 3.1: Spatial actors classification upon the scale at which they operate and maintain data

modelers rarely have access to all the above mentioned spatial actors repositories. As such, most often data sets are completed / repaired with the use of statistical data and ETL (Extract Transform Load) operations.

3.3 A proposal of standardized spatial scales in UBEM

In [42] consistency of both data quality and access were recognized as major challenges in urban energy system planning. Data is indeed dispersed across different physical and administrative scales, and spans the entire urban planning value chain, all while being hindered by privacy issues and lack of standardization. Every spatial actor presented in the previous section has it's own specific spatial scale at which it works and aggregates data. Because of the nature of administrative work and the fact that cities are often not alone in regulating and controlling their energy chain, data tends to be segregated among different spatial actors and scales. These lines of division are connected to topographical, cultural, and historical differences that influence the evolution of cities.

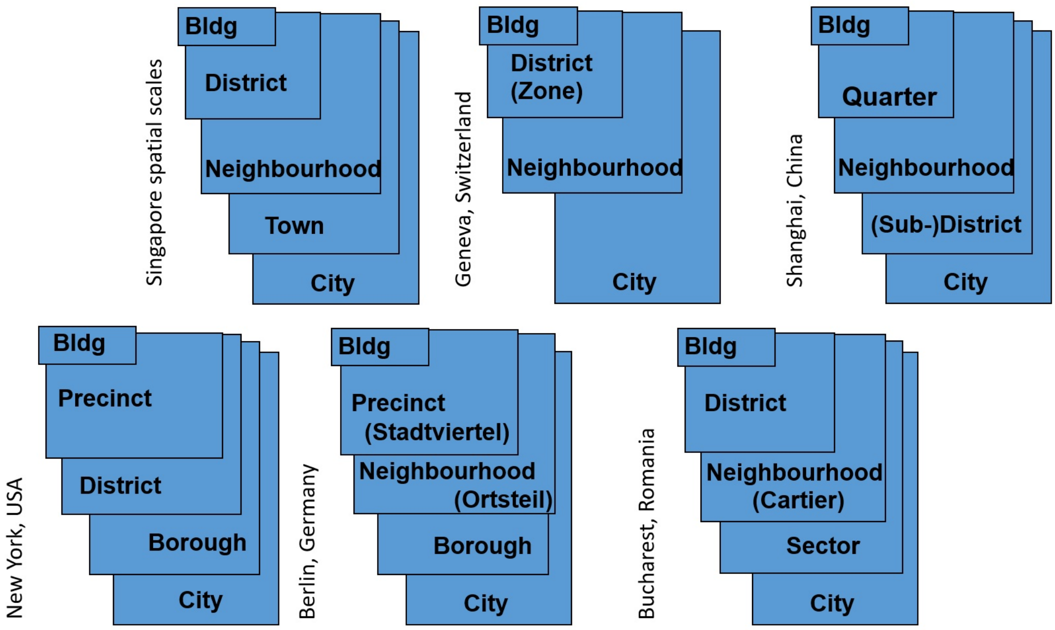


FIGURE 3.2: Spatial scales in the city wide energy chain in different geographical, cultural and administrative contexts

Six cities that the author was able to obtain open data for were selected for a graphical representation of the spatial divisions of data, see 3.2. These cities were chosen to depict different geographical and cultural contexts and share to different extent the open data initiatives that facilitate the work carried out in this thesis. These cities are Singapore (Singapore), Geneva (Switzerland), Shanghai (China), Berlin (Germany), Bucharest (Romania), New York (USA). The cities chosen for the comparison lie in terms of size and population in the same category of multi-million metropolis. What can be observed from these examples is that public administration shares the tendency of organizing data on a similarly called scale, neighborhood. This spatial level contains anywhere in between a few hundred to a few thousand buildings, 3.2. Neighborhood, sometimes referred to as community district, section, borough, precinct, block, area, zone, or in other languages as quartier (French) and Stadtteil (German), is in fact a common spatial scale (in terms of approximate size) used for division across different cities and is something that people of all backgrounds generally refer to when understanding city scale. The same concept can be met with cities that contain smaller population numbers and have a reduced surface size, with the difference that smaller cities tend to lack the the 4th administrative level (referred to as town).

To best illustrate the spatial scales an example in the city of Torino, Italy was created using the six spatial scales present in the city: city (città), town (circoscrizione), neighborhood(foglio), district (isolati), building (edificio) and dwelling / client unit (abitazione). This can be observed in 3.3.

Largely, the spatial scales identified in urban planning and management involve a significant number of spatial data stakeholders and can be classified in six levels, presented in 3.2. Across the world, the keywords used to name the areas vary greatly in meaning and size, spanning municipalities, subdivisions of municipalities, school districts, or electoral district. In this table, all names have been translated to English, using Google provided machine translations. These are deemed correct for English use in Academic Purposes by [87] for most purposes, except grammar. The numbers provided in the table serve only as indication of scale, rather than accurate numbering. In fact few numerical description for these spatial scales have been found, with the exception of China, where neighborhoods are estimated to be at around 2000-10000 families, and quarters at 100 to 600 families, [49], however, numbers like this are seldom provided and do not constitute a fixed rule, rather an indicative value, even in the Chinese context. For this manuscript the following definitions were used:

- city, administrative entity where people live together in a high density urban sprawl with a minimum density of 1000 persons per square kilometer and a minimum size of 5000 individuals
- town, spatial unit, generally smaller (or equal) than a city, that encompasses multiple distinct zoning areas (different area uses such as commercial, industrial or residential)
- neighborhood, geographical entity that usually contains a single use zone, very often residential in nature, with limited tertiary activities present;
- city block, the smallest group of buildings that is surrounded by streets; usually statistically significant for the city administration, and does not have thermal bridges to other surrounding entities (a distance larger than 0.3m - narrowest street in the world);
- building, structure with a roof and walls standing permanently in a location, that has it's own entryway (or multiple);

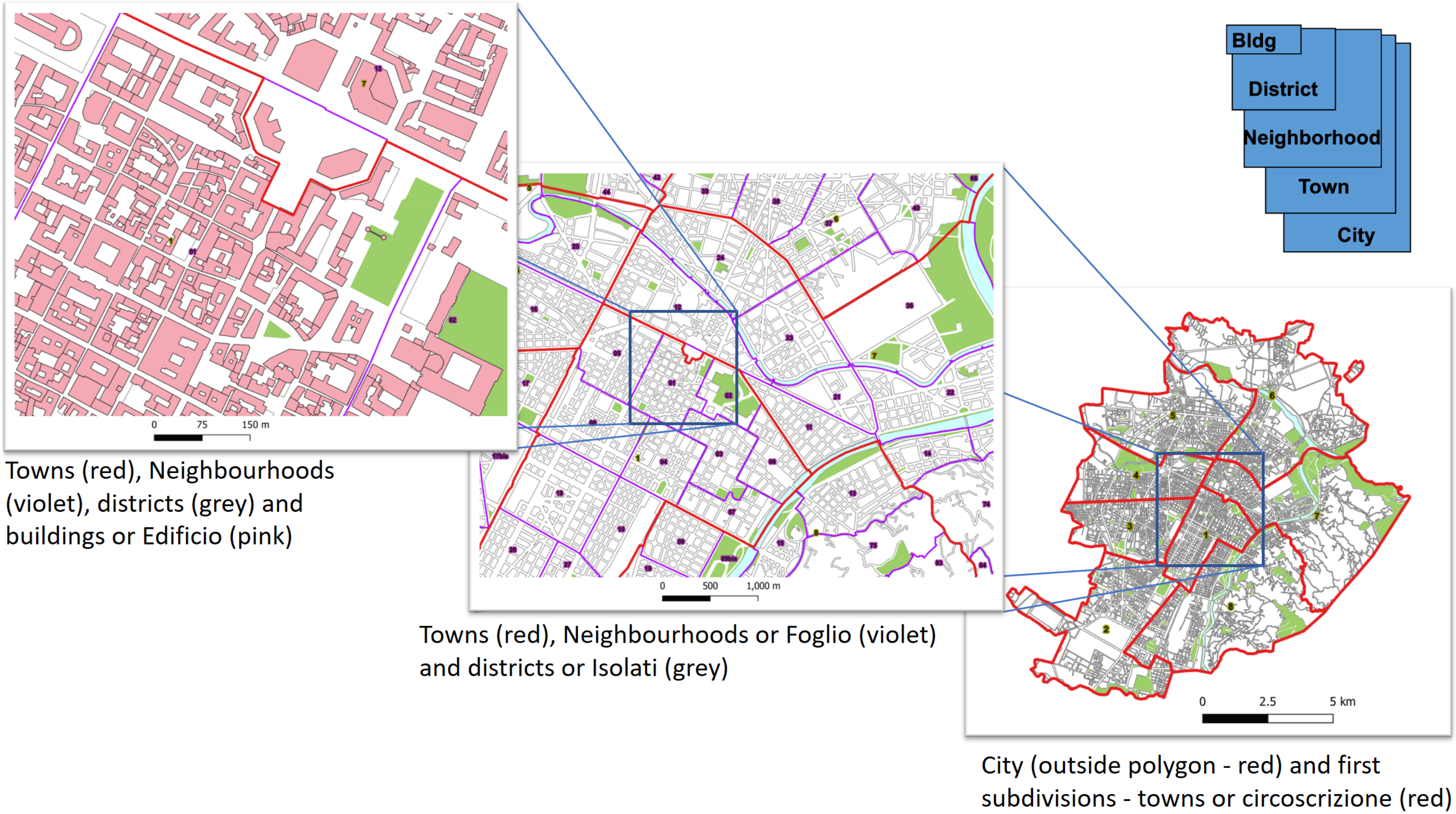


FIGURE 3.3: Spatial scales in the city of Torino, Italy

- energy meter, or the client unit / final consumer, usually having a one to one relationship with an administrative unit (apartments, building or a couple of buildings), is for energy companies a unique entity located on the premises of a residential/tertiary/industrial unit usually within a larger structure/building; a single building can contain multiple energy meters.

The logical way to describe the definitions and the relationship of scales from above is as follows: every city has at least one town, and must have at least one neighborhood with a couple of city blocks and a minimum of a thousand buildings. Within the work presented in chapter 6, the energy data is linked or aggregated at the level of buildings or city blocks. A graphical depiction of the definitions of city blocks, buildings, and energy meters presented before is offered in 3.4.

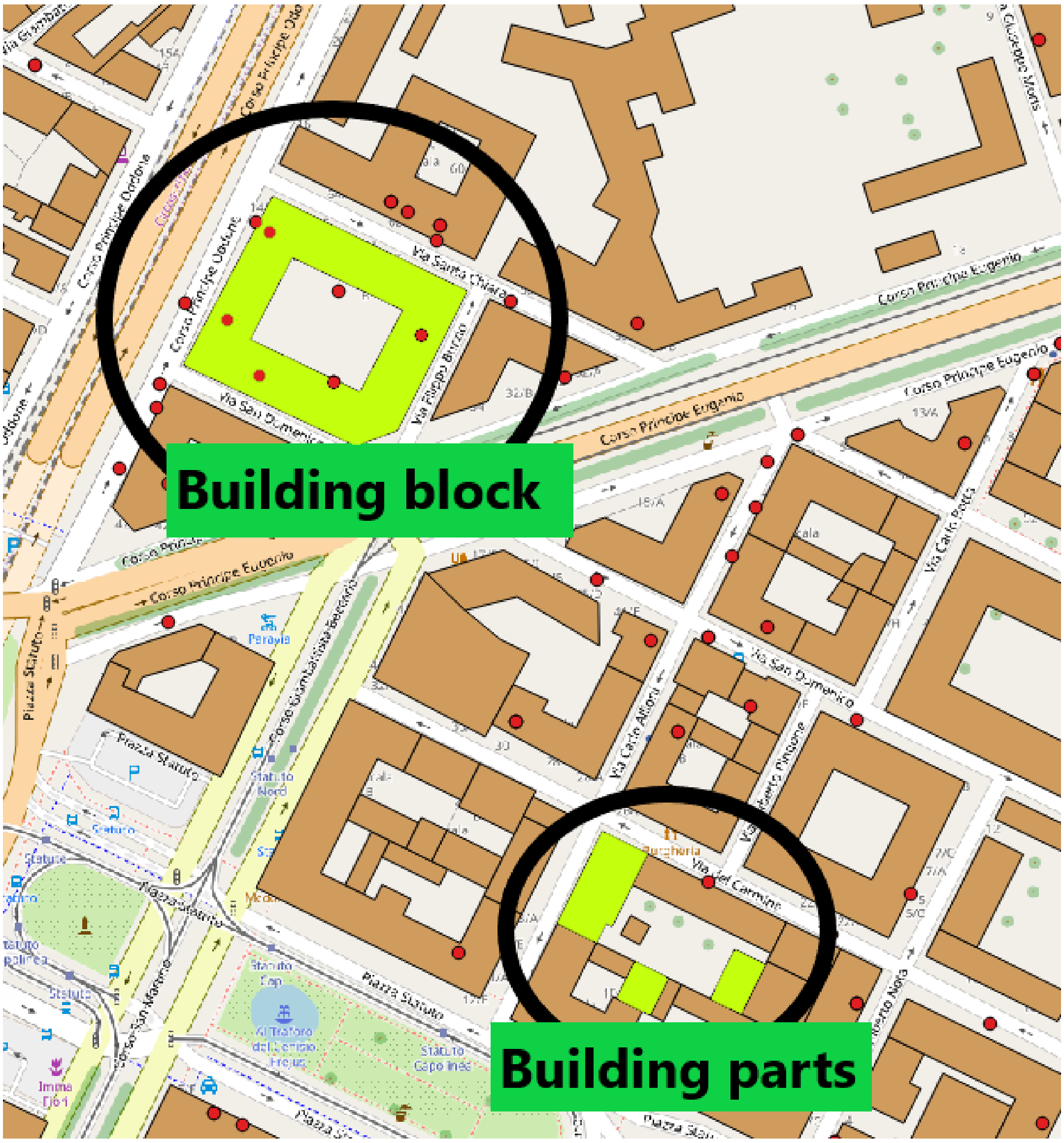


FIGURE 3.4: Building blocks, building parts and energy meters in Torino, Italy

The natural evolution for administrative scales means that often historical data sets exist for these spatial units, which in turn allows for the performance of data-driven approaches that support trend and change detection. For example in the Canton of

TABLE 3.2: Description of spatial scales in the energy chain

Scale name	**Substitute names**	**Rough number of buildings**	**Approximate number of inhabitants**
City	Urban center, Polis, Metropolis, Municipality, Conurbation	$10^3 - 10^6$	$10^4 - 10^6$
Town	Borough, Sector, District	$10^3 - 10^4$	$10^3 - 10^5$
Neighborhood	Community district, Section, Borough	$10^2 - 10^3$	$10^3 - 10^4$
District	Precinct, Building block, Quarter	10 - 10^2	$10^2 - 10^3$
Building	Building parts, Structure, Ward	1	0 - 1000
Energy meter	Final consumer, Client unit, Dwelling	1	0 - 1000

Geneva in Switzerland, urban and energy data is centralized in an online database. The data is kept open, by law, to researchers, citizens, journalists, and companies, while protecting personal data. In the repository energy data is organized by scale following the various planning scales, from building, to localized plots, to neighborhood, town and canton, [42]. In certain cases, large scale urban or agglomeration projects also warrant dedicated data structures exiting the usual spectrum of administration based data structures.

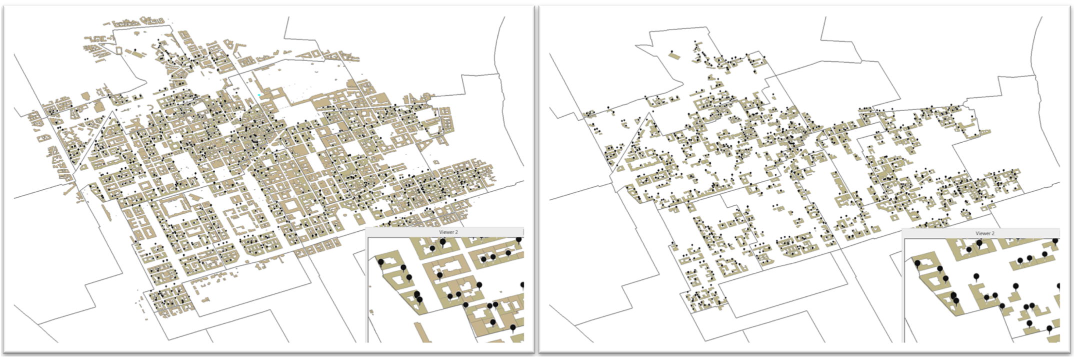


FIGURE 3.5: Spatial energy data inconsistency example: buildings and energy meters data, left presents the entire building stock, right, only the buildings with associated energy meters

3.3.1 Spatial relevance of data unit sample size - scale sensitivity in UBEM

A noteworthy aspect in the city wide energy chain modeling is the relationship between the extent of what is modeled and the size of the data unit at which information is aggregated/modeled. This entire chapter makes a case for the use of buildings as data units and neighborhoods as the "right" spatial scale at which aggregated and disaggregated data facilitates city-wide energy chain modeling and city data modeling in general.

The lack of standardized input data means that UBEM models often make use of "building archetypes" in order to reduce the simulation inputs required, [3], [228]. Archetypes are groups of buildings definitions that share similar properties, such as use, age and belong to the same spatial extent. A review of existing national and local archetypes was presented in [191] and further methods for developing multi-scale (national, city, county and district) archetype development methodology using different data-driven approaches can be found in [3], and [46]. The procedure generally consists of five steps: data collection, segmentation, characterization, quantification, and results modeling. Archetypes and the entire effort of the UBEM scientific community represent another testament to the potency of using buildings as data units in energy modeling.

This chapter states that city wide modeling should be done at neighborhood level with buildings as the data aggregation unit. However, one can argue for the possibility of going beyond building data units, and providing more, or less, data aggregation. For example, [121], approaches this issue from the perspective of sensitivity of the energy model, and converges on the idea that larger data collection processes for higher details in UBEM result in statistical improvement of the results only with small sample size calculations (a small number of data units). In his case the data unit is the household, what is previously identified in 3.2 as the energy meter. The author compares simulation results at different scales against measured energy use in the case of single residential units to a multi dwelling building in the area of Karlsruhe, Germany. The results depicted in A.6 show the increasing coherence of a predictive energy demand model with increasing sample size of one, five, ten and thirty residential units of the same building. The statistical model uses the mean daily outdoor temperature as single regressor to predict daily heating needs. The findings are an indirect indication of the dependency of scale in the aggregation of energy use data. The correlation of both time series at the scale of individual units is relatively weak and derives strongly from the line of equality, (top left figure). With increasing scale (larger number of energy meters / consumers / residential units) a much tighter fit of both samples is reached. This matches well the line of equality (30 samples, bottom right). As the predictive model is not adapted to represent individual behavior, the results indicate that these are clearly visible at a small sample size. However, with increasing sample size the results indicate that different behavior is resulting in an increased as well as diminished energy use so that the results can be predicted using an uncalibrated model. This corresponds well with the purpose of UBEM, city and town scale modeling, where large numbers of individual usage patterns combine to form the total expected demand.

This statistic effect is also underlying the use of statistic load profiles, in which individual behavior is aggregated to a smooth profile. This perspective was studied in electric systems, with the aggregation described by [181]. In the case of thermal demand, [85] used the aggregation methodology for developing synthetic load profiles. The aggregation of load profiles allows for accurate prediction of energy use at larger scales. In the electricity market this results in reliable predictions based on standard electrical profiles for a large number of similar consumers as well as in the gas load prediction for regulatory zones, [18].

The results for the scale of urban neighborhoods presented in [121] show similar statistical effects. This leads to two main conclusions based on the same effect. On the one hand statistical modelling approaches such as the energy signature model proved to be applicable at this intermediate urban scale. This conclusion is in line with [136] and is the main thinking behind UBEM applications that include the prediction of energy needs for district heating systems or common systems for groups of buildings. On the other hand, the fact that differences in energy use by individual behavior neutralize each other clearly indicates that the scale is not suitable to detect excess energy use or under-supply of individual units. Therefore, energy-monitoring strategies are highly scale sensitive and must consider different scales in their design to validate their results.

At neighborhood scale (hundreds to thousands of buildings), regardless if data is provided from statistical records, data-driven approaches, or disaggregation, taking a modeling approach where buildings are data points will result in input error data leveling itself out due to significant data and energy use volumes. At the same time solutions can be provided in a meaningful manner, in the sense that local energy issues can be tackled with a sense of .

3.4 Spatial consistency

Providing spatial data at the same scale and with similar resolution and accuracy is crucial for energy modeling algorithms. Most of the times energy specific data is inconsistently spread in corresponding spatial levels. Data sets consistency issues in the spatial domain can be separated in three categories, first border / boundary consistency, secondly multi scale coherence and thirdly data contiguity. Regarding the first, the boundaries for energy relevant stakeholders in the city wide chain differ widely, as can be seen in the case presented in the figure 3.6. Depicted with red are the public administration spatial units while blue illustrates the ones of the statistical agency. The example provided is located in the city of Torino and illustrates border consistency issues and why data treatment and generalization are often required to the information UBEM models require.

Multi-scale coherence regards the provision of attribute data at different levels of resolution, e.g. buildings, districts, neighborhoods. Per se, GIS data has two components, one regards spatial entities, their coordinates and information to their coordinate systems (European Petroleum Survey Group identification code) and the second, its attribute data. Attribute data are appended to their corresponding spatial feature. The first example of such association is the papyrus of Torino, drawn around 1150 BC, where gold and silver mines, associated with the mined quantities are depicted. The map was used for estimating financial revenue of the state, [118]. In UBEM's case, attributes regarding occupancy (such as number of inhabitants, number of dwellings), or energy sources (independent source or grid district network connection) are indispensable in energy demand calculations. This is why having coherent data sets across multiple spatial levels is among the first targets of ETL operations.

In addition to differences of borders, the issue of data contiguity plagues these data sets. Contiguity is a concept used in GIS to define in topology the adjacency of spatial objects. In the case of polygons they are contiguous to each other if they share a common arc. Extrapolating this concept to the specific case of energy data, data is contiguous when entities belonging to a string inside a data set, or ones that share a spatial border with one another do not have data gaps.

3.4.1 Scale related ETL operations

Harmonizing data across different spatial scales is performed through a procedure entitled generalization of energy data. This operations involve various ETL procedures:

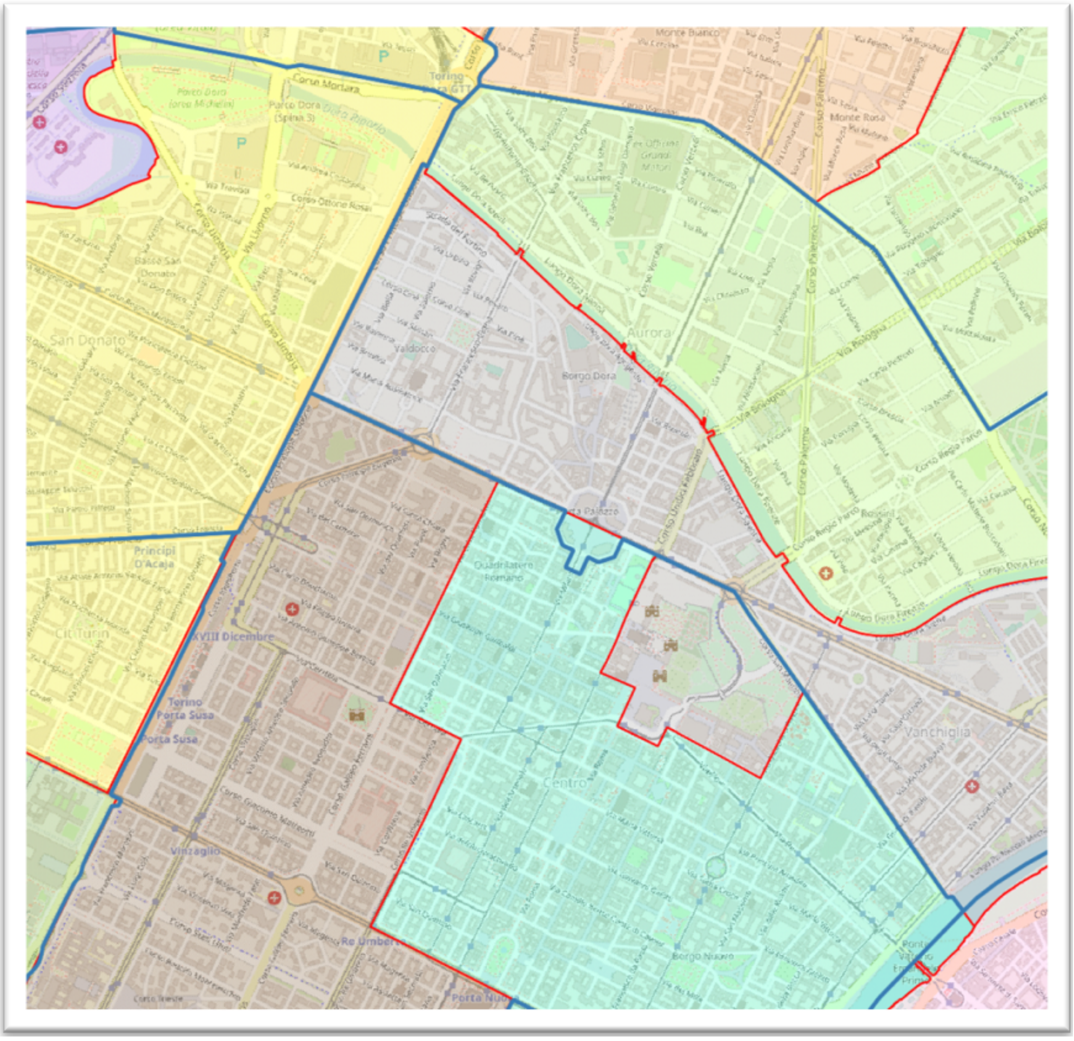


FIGURE 3.6: Spatial borders inconsistencies in the energy chain spatial data units(red – public administration, blue – statistical agency, map created from open repository offered by the municipality of Torino and the Italian National Statistical Agency)

- Creation of hierarchies and spatial relationships. Establishing hierarchies, ownership, and spatial relationships are a major formal requirement of semantical city models and UBEM modeling. These are explored further in sections 4 and 3.6.
- Direct allocation from a data owner / source. It can be performed using spatial methods (see figure3.4) or directly (when identification using unique identifiers exists), by using the previously created hierarchies. Entity to sub-entities relationships.
- Generalization by (dis)aggregation and clustering. Relying again on the previously mentioned hierarchies, (dis)aggregation is required for UBEM modeling, but also for interactions with modeling other kind of urban activities such as waste management or water distribution, that interact with energy systems. These other urban modeling activities are often performed at different spatial and temporal resolutions. It is for this reason that there is a need to (dis)aggregate information at smaller/larger spatial units. This processing

can also be required because of data inconsistency with regards to the spatial levels in between stakeholders / institutions.

- Extrapolated from direct neighbours
- Disaggregation from the larger scale spatial unit to which the building belongs to
- Aggregation from smaller residential units that belong to the building in question
- Spatially randomly distributed from statistical data sets
- Clustering using socio-economic indicators
- Measurement based allocation

This procedures ensure consistency and provide an interrupted flow in UBEM methodologies. They are required either for providing data as input to the UBEM algorithms (where they become the simulation backbone), either for the purpose of sharing results and providing immersive decision-making in the analysis sequence. Providing data to the simulation is usually the first part in a inclusive and iterative process where multiple stakeholders contribute input that help energy and urban planners devise processes that involve multiple steps in the preparation. These processes can even impact spatial segregation of planning units due to energy resources allocation (over production in certain areas, energy poverty in others).

To facilitate the theoretical approach proposed in this thesis with regards to buildings as data units for energy information aggregation an example is offered in the case of a neighborhood in the city of Torino, Italy. The energy and socio-economic relevant data is linked to a 3D CityGML model of the buildings. The buildings are a 3D data set, while the certificates are point data in 2D. Both data sets originate in the open domain and prior to handling within UBEM algorithms underwent several ETL procedures in order to further clean the data (eliminate incomplete, inaccurate or conflicting entries). To facilitate the merger of such data two GIS related methods were used: contained within and nearest neighbour assignment. These very potent operations facilitate the assignment of energy certificates data to the buildings. As can be graphically observed in figure 3.5, only one in five (19%) of all buildings with an energy consumption located in the center area of Torino receive a direct assignment from energy certificates. This incomplete data set is a common occurrence in the UBEM field and is handled on a case by case basis.

In addition to missing or incomplete data, there is also the issue of contradicting data. For example, energy certificates of dwellings of the same building can present conflicting indicators (e.g. different year of construction or energy use class). Issues such as this one can be tackled using automated data validation processes (for example data-driven using rule based algorithms, as presented in [47]). However, in certain cases, manual intervention is required on behalf of the modeler, and this is the reason that the data processing itself is usually about 40 to 50% of the estimated work load in the case of UBEM modeling.

3.4.2 On generalization by (dis)aggregation and clustering

With raster data, projects dealing with multi scale data resample image pyramids. This is done in order to facilitate the use of algorithms for analysis and visualization, so that the optimal amount of data is appropriated for each different scale. In the case of vector data multiple scales can be derived by means of data generalization such as the aggregation of geometric objects and aggregation of features classes, approximation of geometric objects by simplification or replacement with geometries of lower dimension. These energy specific ETL (Extract Transform Load) data processes are time consuming and usually entail discrete changes to features with every change of scale. [207] discuss how continuous changes of individual objects can be applied for visualization processes. Even though the work of Sester and his colleagues focuses only on visualization, the same approach to decompose data generalization into simple geometric and topological operations is arguably similar to what is proposed within this chapter. The objective is to generate a chain of entities that reduces both data gaps and redundancies while also allowing for the transmission of information via hierarchies and references which are needed when energy and spatial algorithms working at other scales occur. The ETL operations can be classified either as discrete or continuous changes, a similar classification to the one used in the aforementioned manuscript that occur to spatial objects and result in changes in topology and/or in geometry. These changes can be:

- Discrete changes of individual objects: the topology and the geometry of the object changes, e.g. collapse operation, where the original geometry is replaced by a completely new geometry;
- Discrete changes of groups of objects: usually occur when larger scale data need to be brought together and thus the abstraction level and often the type of object changes, e.g. typification, where a group of objects is represented by a new group consisting of fewer objects;
- Continuous changes of individual objects: the topology remains the same, however geometry and/or attributes change by moving either the whole object or individual points of the object, e.g. continuous simplification of objects where the original object points are relocated.

Creating UBEM data sets is often performed with statistical methods, namely clustering. This methods are employed when some data sets with regards to morphology, 3D models, socio-economic indicators are missing or difficult to acquire (in terms of time or finances). The three methods identified in the UBEM field are PCA (Principal Component Analysis), k-Means and SOM (Self Organizing Maps). Once a cluster is identified, similar values are allocated to the entire lot. PCA is a hierarchical unsupervised vector space transformation techniques used for exploratory data analysis, [14]. It is used to reduce the number of dimensions of multivariate data sets, particularly when the dimensions are related to each other. PCA shows variation in the data set by constructing a set of new uncorrelated variables derived from a linear combination of the original variables. An example of this work with energy modeling is found in [152]. With PCA clustering analysis the significance of the different parameters was performed in an urban data set. From that five representative typological classes of the urban fabric for a Brazilian city were highlighted (sparsely built, open-set mid-rise, colonial compact centre, modern high-rise, and densely built low-rise). This allowed for statistical values to be allocated to the entire data set. k-Means is a simple structured iterative clustering method based on the centroid method, [141]. Prior to processing taking place the number of cluster (k) is defined and the optimal solution for the given k distributes centroids as far apart as possible with data observations being associated with the nearest centroid. Cluster centroids are recalculated based on mean distance to all observations in newly generated clusters. This process is repeated with cluster centroids shifting locations after each iteration, until the clusters stabilize. SOM, a specific case of ANN (Artificial Neural Networks), [241], are particularly useful with multi-dimensional data sets. It projects input variables into a low-dimensional grid which is then used in visualizing and exploring properties of the data. The SOM grid is user-specified in terms of size and shape with cell values being initialized randomly. Through multiple iterations the values are adjusted. A good example of a UBEM publications using clustering methods is found in [200] where SOM is applied on a dataset of approximately 40k buildings to identify clusters of similar socio-economic parameters that can be used for energy retrofitting business cases.

3.5 On urban fabric contiguity impact on scales in city energy modeling

For UBEM applications, most often, neighborhood, city block and building are the data units scales at which information is aggregated when energy modeling is performed. This, combines with the importance of statistical significance for energy urban planners, an item which cannot be understated in UBEM, and is explained further below in this chapter. Habitually, one city block is provided with energy from the same source, and energy renewable and refurbishment applications are applied at either such a scale or at the level of a building due to one often ignored aspect: urban contiguity - non-interrupted urban fabric of building envelopes. The significance of this aspect in the field is furthered in [152]. The fact that buildings share envelope surfaces allows first of all for energy savings, due to the reduction of weather related losses, a fact which, for example, further justifies the continued use of row houses or apartment blocks. Moreover, energy production, storage and distribution network infrastructure can be shared with minimal additional expenses and distribution losses. Detailed energy related use cases and applications are further explored in 2 and 6.

With regards to the building spatial definition given above, for the author of this manuscript, who has both a geomatics and an energy engineering background, the 25cm distance rule is required for both spatial and energy factors. The first and simplest reason is that the narrowest streets in the world are around the same size, and streets are used generally to separate administrative units. Another factor is that most UBEM simulations are habitually incapable of spatial segregation of buildings and in some cases even of city blocks, [148], either for a lack of standards or useful input data, more on this topic in 4 6. Another argumentation lies behind the speckled or patchy quality of varied spatial data used. In the 2D domain for example, surveyors and photogrammetry experts can have a different understanding of what constitutes the polygon describing the building footprint, as seen in 3.7. The outside polygon can be, for the ETL experts who create the data set, either the physical side of the building, the edge of a balcony or the edge of the roof. With regards to the 3D part, the absence of underground structures (or cellars) in the definition, or the lack of roof top description are both explained by the general lack of general information at administrative units. However, cellars are confirmed by [80] to be playing a significant role in UBEM calculations and such, the lack of such information, deprives UBEM applications of enhanced accuracy in calculations. Together the 2D and 3D issues can have a considerable impact on volume calculations, itself a major regressor in thermal energy demand, [121].

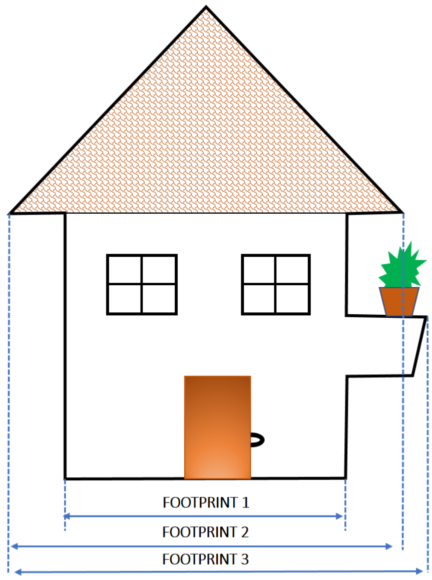


FIGURE 3.7: A building footprint and its conflicting understanding for several professionals

Even though, the spatial scales presented in 3.2 are the results of historical developments, it is noteworthy the fact that scientists also started recognizing neighborhoods as the optimal scales for energy urban-planning projects, [136]. Due to its significance, a definition of neighborhoods was sought in existing literature. In general the concept is described by sociologists as the place where social interaction occurs. As such it is essentially a social construct, made of ever-changing and layered systems of personal and collective links within and across cultural boundaries, [156], which is a rather similar concept to energy networks. What makes people share an energy network and an energy source is arguably very simply explained by their vicinity and accessibility to/of it. This corresponds to Tobler's first law of geography in [217]: **"everything is related to everything else, but near things are more related than distant things."**.

This is especially true with regards to energy trends presented in chapter 2, specifically the fact that buildings and their inhabitants are now concurrently consumers and energy producers (prosumers) and there is a lack of energy transportation networks able to take that extra energy in order to deliver it. For electrical networks the issue is not a lack of networks, but rather a design "flaw" as originally networks were designed with uni-directional flows from large production centers to consumers, see 2.2. For all other energy utility networks, in most cases they simply do not exist, and represents one of reasons for which building data units are recommended for use in city wide energy modeling the by this manuscript. Buildings, per se, constitute contiguous entities. An appropriately scaled framework, in this thesis's case, the building unit, on which a more continuous, more spontaneous urban pattern may be formed.

3.6 A scale dependent urban data ontology with persistent identifiers

One of the main impediments in the adoption of UBEM tools is the lack of a standardized ontology, [77], as well as the lack of standardized city wide energy data, 3D city models and test cases. Within the economically well developed states of the European Union the spatial part of the issue has been slowly finding resolution with the advent of CityGML models constructed from aerial photogrammetry and LiDAR. This data availability has also been facilitating use cases, presented in chapter 2 and 4. In addition, there are open data initiatives providing energy relevant data, such as [187] and [179]. However, the ontology matter remains a blocking issue, as does the lack of unique building identifiers in the data sources, [51]. To facilitate future development of UBEM tools and their connection to other urban modeling activities this chapter proposes a scale dependent urban data ontology that can be easily used by modelers and stakeholders of different backgrounds. The proposal presented addresses separately energy and network thematic data by making use of the Energy ADE and the Utility Network ADE. As such the focus in this sub-chapter is not on the entire ontology, rather on the spatial relevant issues that guarantee ownership tracing and energy chain linking.

The problem of valid identification at multi scale is well known in the Information Technology domain and alternative solutions exist in the GIS field (for example UUID or GUID). However, none of the solutions that the author identified are adaptable and human readable for the multi scale issue presented in this chapter 2. UUIDs (Universally unique identifier for all the rest) or GUIDs (Globally Unique Identifier for Windows developers) are large numbers that are nearly guaranteed to be unique. They are usually 128 bits long and look as follows in hexadecimal:

30dd879c-ee2f-11db-8314-0800200c9a66

The format is a well-defined sequence of 32 hex digits grouped into chunks of 8-4-4-4-12. This gives us 2^{128} or about 10^{38} numbers. UUIDs are standardized by the Open Software Foundation (OSF) and integrated in the CityGML standard to identify spatial entities. This identifiers complement the proposal made in this section, as they serve IT systems and existing standards with predefined integrated capabilities.

For an identification system that allows architects, engineers and IT experts to interact together simple, scale dependent solutions are required. These identifiers should be persistent, regardless of ownership changes and allow users with minimal IT background the ability of controlling and handling ID allocation and configuration. That makes the system multi-domain accessible, and similar in a way to XML (eXtensible Markup Language), where the language was designed to store and transport data and at the same time be self-descriptive.

The proposal follows in the same rationale as presented in [16], where hierarchical clustering is used to explain relationships by using origin-destination flows. As presented before, all the complex relationships between the objects in the city wide energy chain require an ontology complete with a spatial representation, formal naming, definition of the categories, and relations between the entities. The spatial structure of city, town, neighborhood, district, building and energy meter is presented in 3.8. An example of the spatial units was previously illustrated for Torino, and is depicted in figure 3.3.

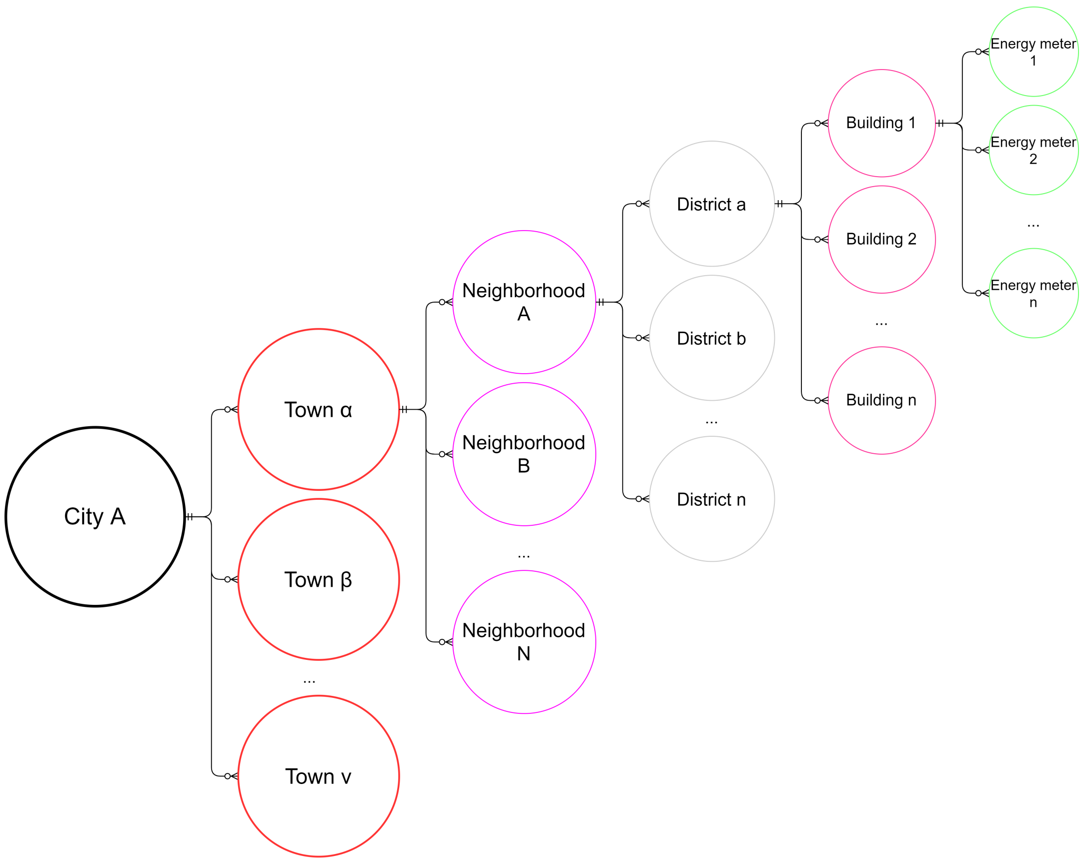


FIGURE 3.8: Hierarchical proposed structures

The unique parent of all entities in a city is the object City. It contains at least one town. Each town contains at least one neighborhood, which in turn has at least one district. Districts can contain multiple buildings. Finally, the building meters can exist in buildings, however, it is often the case that multiple buildings can be allocated to a single meter. The parent - child relationships are all of type one mandatory to many optional. Thematic attributes are allocated to each entity and can be aggregated and disaggregated from one scale to another using specific calculation methods. Certain attributes are present at all scales, for example, number of inhabitants, while others could be scale specific, such as building physics parameters, as can be observed in 3.9.

Within the energy chain energy thematic information are often linked to spatial units building entities. Ergo, buildings host this data as attributes. This further facilitates the links to the Energy ADE and the Utility Network ADE which also use buildings for the storage of specific information. The larger scales, that go beyond buildings, are relevant to energy planners and usually host energy related KPI indicators such as total energy demand, GHG production, CAPEX (Capital expenditure) or OPEX (Operating expense).

In order to facilitate modeling activities, the entities host two identifiers as attributes. This are referred to as PID (Persistent IDentifier) and GUID (Globally Unique IDentifier). No single PID system for urban energy planning was identified during the writing of this thesis. Typically, in information technology, PIDs are used for the foundation referencing of digital assets in scientific publications, books, and digital repositories, [229]. PIDs usually contain metadata, and the same thinking was applied to the realization of the ones included in the thesis. The method proposed here embedded spatial levels in the identification, thus facilitating modelers understanding of the spatial entity that they are working with. The presence of the GUID, sometimes referred to as a UUID (Universal Unique IDentifier,) as a second attribute, is made so that the IT systems to be provided with a GIS typical unique identification mechanism (also used in CityGML and the entire GML - Geographic Markup Language community). Table 3.4 presents the two identifiers used in this proposal.

The number of digits used at each spatial level in PIDs was determined using samples available for the previously mentioned cities, while also giving future users flexibility, in case certain spatial scales are non-existent in their own use case. The length of the string depends on the number of spatial scales that precede it, see tables 3.4 and 3.3.

The flexibility of the proposed method allows for different sized cities to be modeled. For example in the case of the city of Karlsruhe, Germany, with an approximate population size of 313k and 85k buildings the scale Town is not required (nor can it be identified within the existing administrative structure). In fact, most UBEM modeling situations do not include an entire city, but rather sections of it. Also, multiple cities can be included, as often data repositories contain multiple applications of the same UBEM models. The modeled area can be referred to as a study area, in case that the entire city is not being modeled, and replace one of the spatial scales in the proposal.

TABLE 3.3: Description of a PID and GUID identifier in the case of a town

ID type	General	Example
PID	Cxx Txx	C01T05
GUID	xxxxxxxx-xxxx-Mxxx-Nxxx-xxxxxxxxxxxx	1s3e45t7-z22c-15h3-a431-423sdf1567193

TABLE 3.4: Description of PID and GUID identifier in the case of a building

ID type	General	Example
PID	Cxx Txx Nxx Dxxx Bxxxxx	C01T05N11D023B00141
GUID	xxxxxxxx-xxxx-Mxxx-Nxxx-xxxxxxxxxxxx	1s3e4w67-e82b-1xd3-a451-423612114123

Following the identification and modeling of all spatial scales relevant to the study area and use case, the existing spatial structure can use the EAV (Entity Attribute Vale) data model for populating all relevant information. The proposed ontology helps mitigate issues such as data heterogeneity and consistency in urban domains by offering a binding platform to all stakeholders from a non-biased perspective.

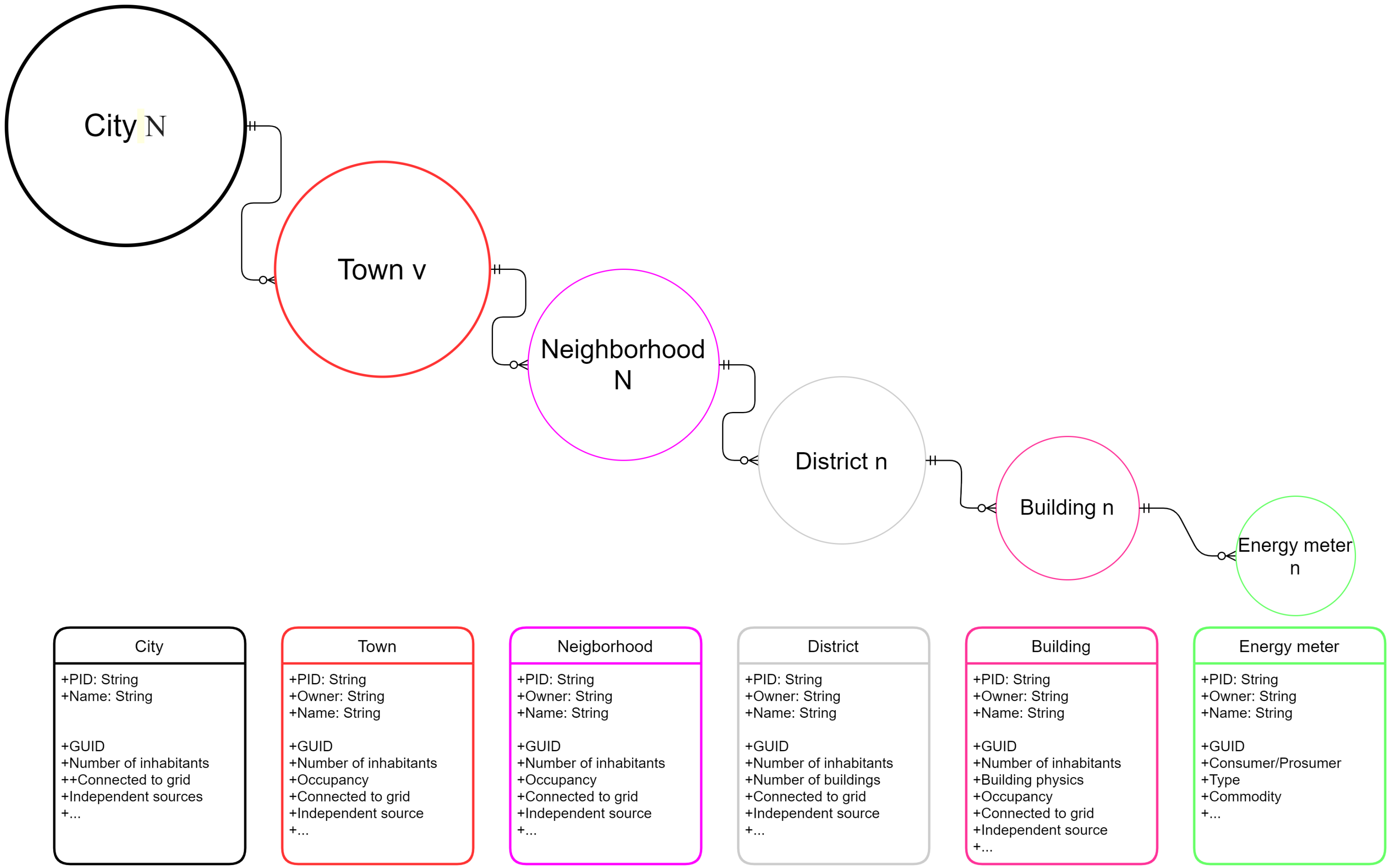
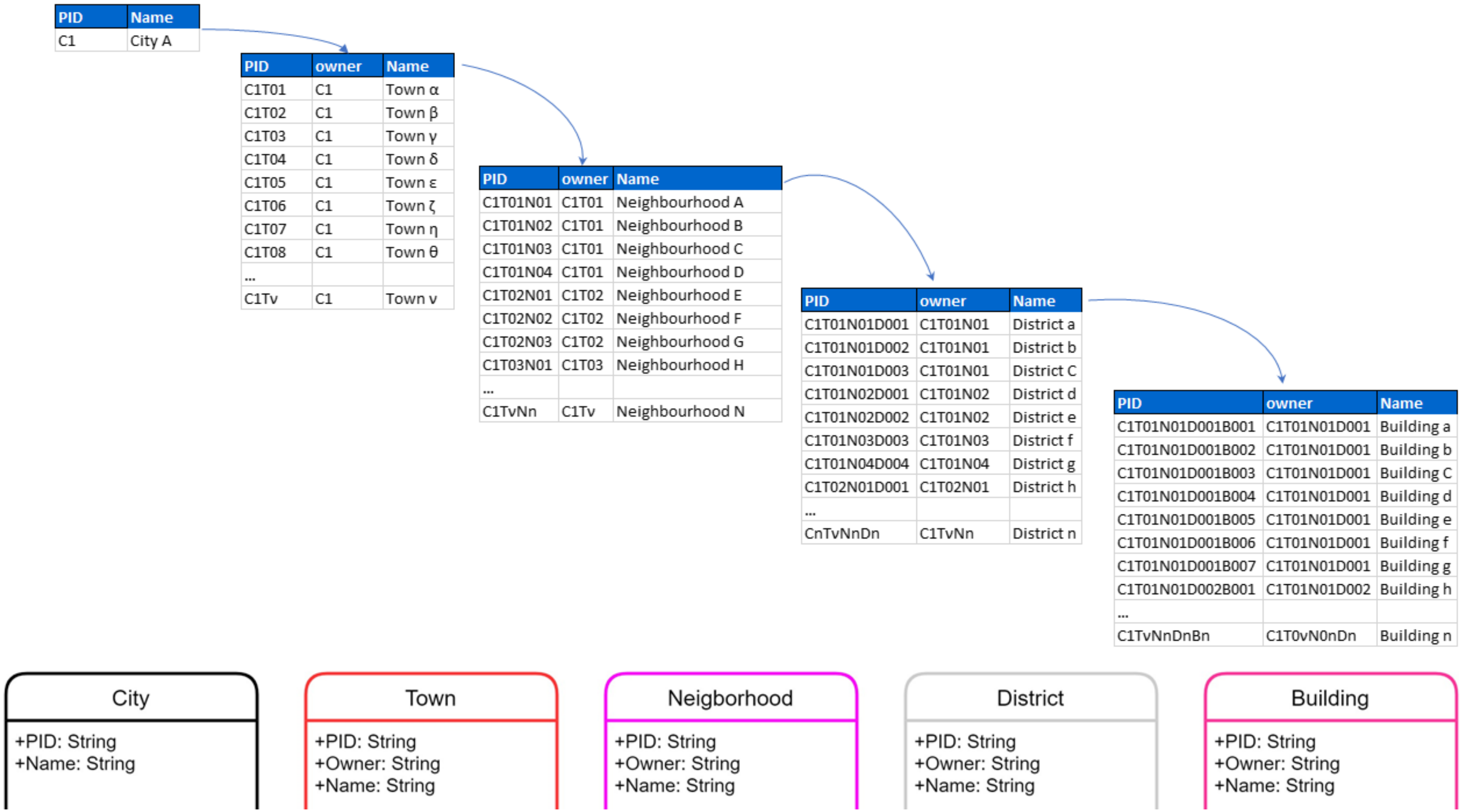


FIGURE 3.9: Hierarchy based on spatial levels



PID	Name
C1	City A

PID	owner	Name
C1T01	C1	Town α
C1T02	C1	Town β
C1T03	C1	Town γ
C1T04	C1	Town δ
C1T05	C1	Town ε
C1T06	C1	Town ζ
C1T07	C1	Town η
C1T08	C1	Town θ
...		
C1Tv	C1	Town v

PID	owner	Name
C1T01N01	C1T01	Neighbourhood A
C1T01N02	C1T01	Neighbourhood B
C1T01N03	C1T01	Neighbourhood C
C1T01N04	C1T01	Neighbourhood D
C1T02N01	C1T02	Neighbourhood E
C1T02N02	C1T02	Neighbourhood F
C1T02N03	C1T02	Neighbourhood G
C1T03N01	C1T03	Neighbourhood H
...		
C1TvNn	C1Tv	Neighbourhood N

PID	owner	Name
C1T01N01D001	C1T01N01	District a
C1T01N01D002	C1T01N01	District b
C1T01N01D003	C1T01N01	District C
C1T01N02D001	C1T01N02	District d
C1T01N02D002	C1T01N02	District e
C1T01N03D003	C1T01N03	District f
C1T01N04D004	C1T01N04	District g
C1T02N01D001	C1T02N01	District h
...		
CnTvNnDn	C1TvNn	District n

PID	owner	Name
C1T01N01D001B001	C1T01N01D001	Building a
C1T01N01D001B002	C1T01N01D001	Building b
C1T01N01D001B003	C1T01N01D001	Building C
C1T01N01D001B004	C1T01N01D001	Building d
C1T01N01D001B005	C1T01N01D001	Building e
C1T01N01D001B006	C1T01N01D001	Building f
C1T01N01D001B007	C1T01N01D001	Building g
C1T01N01D002B001	C1T01N01D002	Building h
...		
C1TvNnDnBn	C1T0vN0nDn	Building n

FIGURE 3.10: Spatial scales based identification system

Chapter 4

Not all data are created equal - elephants in the room

This chapter focuses on the spatial-temporal data in the digital domains that interact in the urban energy chain of spatially aware cities. This includes the standards, Spatial Data Infrastructure (SDI) and methodologies used when modeling and monitoring the city wide energy chain domain, as previously defined in chapter 2. The entities and interactions modeled with the standards included in this chapter are approached from the perspective of urban energy modeling and planning. Modeling activities require the detailed description of the existing situation in order to carefully predict the behaviour of existing systems. Ideally, energy modelers are looking for situations where the modeling input (the entire information required for the description of the *status quo ante*) is made available, in uninterrupted workflows, that allow for contiguous energy chain modeling, from standardized sources. In order to process all this information good infrastructure and consistent data flows are required. As Hodges and his colleagues explain in [101], declarative modeling languages already facilitate the definition of semantic relationships with such quality that it makes them suitable for enabling machine-to-machine (M2M) interoperability (interaction and understanding). They continue to identify gaps with regards to proprietary ontologies and cascading effects from their use, an conclude that these semantic models are unlikely to play a large part in resolving the pressing horizontal and vertical interoperability issues currently encountered. *These proprietary ontologies hold the promise of interoperability in name only.* This judgement further motivates the work presented in this section as it focuses on using open semantically organized information as support for interoperable spatial-temporal data flows and applications in energy modeling.

4.1 Use cases - just what the doctor ordered

The doctor that the title refers to is in this case a software expert. The doctor uses, as a standard today, the AGILE software development method, [54]. One of the cornerstone of this method are use cases, [203]. Producing city wide information models in the city wide energy chain usually involves people with multiple competences surrounding the spatial, energy and environmental domains. To justify the infrastructure and the data collected there is a need for wide ranging application potential. As such this section of the manuscript presents use cases that involve the solutions proposed and used. The work of Biljecki and colleagues, [26], identifies a total of 29 use cases where 3D city digital twins are deployed. These use cases are defined as situations that cannot be resolved without 3D or where the use of it adds significant value in the use case. In the thesis context of the city wide energy chain, 25 use cases were distinguished (presented in table 4.1) that help plan and manage the energy chain in a city.

What can be observed from the table is that the use cases touch all the energy chain components previously identified in Chapter 2, either directly in production centers, substations, consumer buildings (residential and tertiary) or networks. With such an inclusive list in terms of energy chain components the case for modeling the entire energy chain by using CityGML as a data hub is strengthened. The fact that 3D semantical city models are currently used in at least 25 use cases with ramifications in the energy chain in cities also provides a justification for cities and mapping agencies to provide data in this formats.

Specifically in urban energy planning both [183] and [196] highlight the importance of open data and standard formats for decision support tools in urban planning. This is essential for three main reasons: (i) to streamline workflows processes, from input to output, (ii) to give software tools the ability of functioning in different geographical settings and at different spatial scales, and (iii) to facilitate comparison and interpretation of results.

Planning of cities for urban infill (greenfield) or redevelopment (brownfield) investments both requires good data availability for successful design and implementation decisions. However, brownfield is far more data intensive, as [42] indicates. To handle the large amount and variety of related data in a structured way, this thesis has adopted the CityGML standard, extending it in the process to cope with the generation of many urban scenarios through interactive optimization.

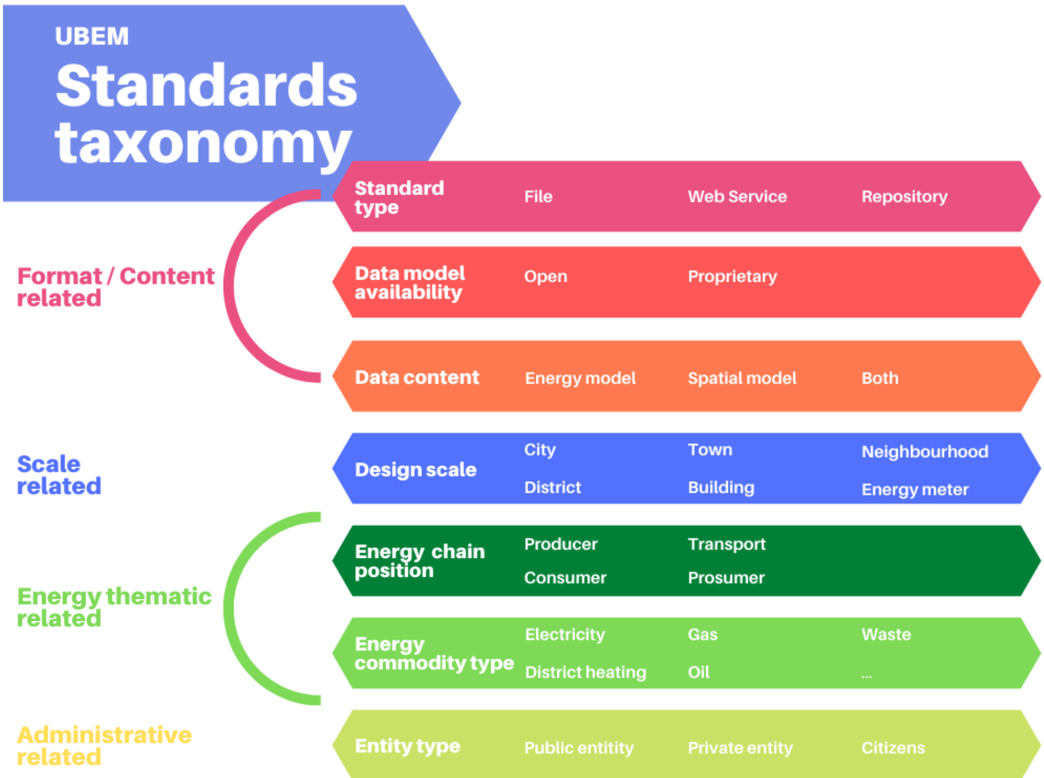


FIGURE 4.1: Taxonomy categories for UBEM standards

TABLE 4.1: 3D City Models uses cases in the energy wide chain, columns Use Case and Application example are developed from [26]

§	Use Case	Application example	Energy chain component
1	Estimation of the solar irradiation	Surface suitability for installing photo-voltaic panels	Producers
2	Energy demand estimation	Assessing the return of a building energy retrofit	Consumer buildings
3	Facility management	Managing utilities	Networks
4	3D cadastre	Correct owner registration	Consumer buildings
5	Classifying building types	Semantic enrichment of data sets	Consumer buildings
6	Geo-visualization and visualization enhancement	Electricity towers positioning	Producer
7	Estimation of shadows cast by urban features	Determination of solar gains/envelopes	Consumer buildings
8	Visibility analysis	Finding the optimal location for a new generation facility	Indirect
9	Estimation of the propagation of noise in an urban environment	Wind mills positioning	Indirect
10	Estimating the population in an area	Demand estimates	Consumer buildings
11	Visualization for network navigation	Navigation	Networks
12	Utility networks	Conflict detection	Networks
13	Urban planning	Designing green facades and roofs areas	Networks
14	Aiding positioning	Map matching	Indirect
15	Visualization for communication of urban information to citizenry	Virtual tours	Indirect
16	Reconstruction of sunlight direction	Object recognition	Consumer buildings
17	SAR images treatment support	Interpretation of radar data	Producers/Consumer buildings
18	Emergency response	Coordinated utilities interventions	Producers/Consumer buildings
19	Lighting simulations	Planning lighting of landmarks	Indirect
20	Computational fluid dynamics	Predicting air quality	Indirect
21	Determination of the floor space	Valuation of buildings	Consumer buildings
22	Routing	Understanding accessibility	Networks
23	Flooding	Mitigating damage to utility management	Networks
24	Change detection	Urban inventory	Consumer buildings
25	Volumetric density studies	Urban studies	Consumer buildings

4.2 Taxonomy of spatial standards and scales

As the simple mention of standards in energy modeling can lead to multiple interpretations there is a need to harmonize the different perspectives that these standards take and attempt a classification of those items. From the mindset that this manuscript is written in, these perspectives are fourfold:

- format/content related (standard type, data model availability, data content),
- design scale,
- energy thematic (energy chain position, energy commodity type),
- entity type (public, private, citizen).

Figure 4.1 presents the different aspects of the taxonomy. In the following subchapter explanations referring to these multiple viewpoints are given, so as to detail the proposed terminology with the mention that the standards presented below are not an exhaustive list of standards used in urban energy management, planning and modeling.

With regards to the data format there are three types of standards in use: file, web services or repositories (data bases) based. In the case of files, information is contained within a single or multiple file based standard (or within extensions of that standard), for example CIM (Common Information Model), IFC (Industry Foundation Classes), SHP (SHapeFile), CityGML (City Geographic Markup Language), DWG (DraWinG). Another method of sharing data is with open standardized web services, such as WFS (Web Feature Service) or BCF (BIM Collaboration Format) while a third is represented by standardized interconnected frameworks or databases. These can be structured - SQL (and even standard bound, as is the example of the 3DCityDB) or non structured - NonSQL.

The content related perspective pertains to the actual data that the standard stores. From the perspective of UBEM this can be distinguished in two separate items: the spatial and the energy model. This feature separation of modeling of the spatial entities and modeling of the energy thematic information associated to each physical entity is present within most standards as energy domain relevant standards rarely provide associated spatial information capabilities. In fact the only standard encountered that has advanced capabilities in both directions was CityGML with its extensions EnergyADE and the UtilityNetworkADE. Energy thematic standards go around the issue of not containing actual geographical information with the use of schematic diagrams, as can be seen with CIM. This standard is not to be confused with the City Information Model, mentioned in Chapter 3. The lack of support for such interoperability with spatial data has been recognized and there is ongoing research into how this can be addressed, as [98] shows for CIM's and GIS's combined use in the modeling of electrical distribution networks.

Standards also have a design scale at which they were designed to be operated. In this manuscript, the subject of scales in urban energy chain modeling and monitoring has been discussed broadly in chapter 3. The spatial scales used to define spatial units in the thesis are City, Town, District, Neighborhood, Building, Energy Meter. With the spatial domains in use within UBEM, their design scale is presented in figure 4.2. It can be observed that the number of entities is highest in the lower part of the pyramid (high total number of buildings compared to a low total number of cities). Similar observations can be made with regards to energy thematic standards.

However, the scale design and dependency is also linked to the energy commodity that is transported. For example electrical network data using the CIM standard can encompass cities, regions and even national and continental networks while district heating networks standards such as DIN 4747-1 and the associated spatial model are limited in scope to districts, neighborhoods and cities.

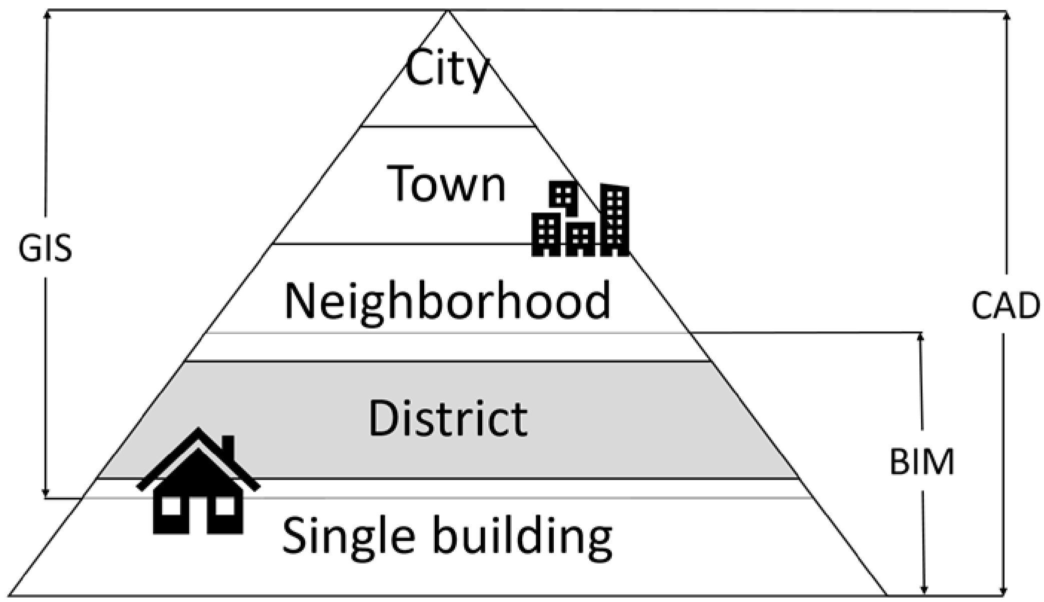


FIGURE 4.2: Spatial model standards and design scales for BIM, CAD and GIS

The third perspective taken in the taxonomy is energy thematic. First with regards to the position of a modeled element in the energy chain, these entities can be energy producers, transportation/utility networks entities, consumers or prosumers, all described in detail in chapter 2. Standards used for producers often relate to the technical requirements of the operator, this can be a heating production plant. For the UBEM modeler the purpose is to be able to model the commodity exchange with the producer, e.g. heating sent in the network has certain technical parameters describing it (temperature, flow) and other parameters on the return flow. Secondly the standards refer to different commodities (electricity, heating, gas, waste water, storm water, fresh water, communication, process steam, oil, waste, air pressure), for example DIN 4747-1:2003 or VDI 2036:2009 for DHN components and annexes.

The last perspective included in the taxonomy is based on the entity type using them. Standards in use by public entities - city administrations, standards in use by companies or standards in use by citizens. These three are separate because of the different use cases that they serve. In general public entities are incentivized to maintain factually correct records of the energy chain components that they are responsible for, and often use standards with open source data models, such as CityGML or CIM. As such their data records tend to cover larger areas and use cases, while the ones in use by companies tend to be narrower, as stipulated to legislation and related to the specific business case of the company. In the context of the digital revolution digital devices such as smart homes installations, smart phones and low cost sensors have also facilitated the emergence of citizen scientists that collect data and often provide it via open data repositories. These treasure data trove often facilitate the work performed by UBEM modelers.

4.3 Digital domains for spatially aware cities

Digital city (wide) information models, also referred to as digital city models and by others as digital twins, are digital representations of cities composed of a computational model and (sometimes) a real-world system, [188]. At its core the fundamental characteristics of a digital twin are:

- data centricity, an architecture where data is the primary and permanent asset,
- 3D capabilities, data contains three dimensional location attributes,
- analytics capacities, the system has the ability of performing analysis on the data supplied,
- computational sensing of an artefact or a process.

Collectively, these aspects ensure informed decision making is taking place, and also allow for data-driven methods to be implemented. Digital twins are used in smart cities applications as the basis for information repositories and form the backbone of data models. Specifically, in energy urban planning projects they allow for building plans and implementation steps to be tested before actual change begins, [80] and also facilitate and improve city management and communication with citizens, [64]. However, building precise digital city models is expensive and *a never ending* task. That means that with cities experiencing constant change, the digital version built is *obsolete* by nature/design. Sensors and complex sensing equipment (such as self driving cars with their sensor multitude collection capability) come in to play to facilitate up to date models.

Existing building and urban models are very often proprietary and non-standardized (non-interoperable), and only intended for high resolution, single building analyses. Government legislation is being used to incentivize and bring forward standardization in this field. This in turn enhances quality of services and the use of UBEM methods, [51]. It is the purpose of this chapter to present standards and standardization in the field, while urban digital twin models and their use in energy modeling are described in greater detail in 6.

The most popular tools and standards used to describe the spatial part of the energy chain can, depending on the scale, be assigned to one of the following categories: CAD (with a focus on design), BIM (centred on buildings and their data models) and GIS (a repository of geographical data). All of these standards can be used to hold, connect and process relevant spatial energy data. The following subsection describes CAD and BIM standards in use in the urban energy chain.

4.3.1 CAD and BIM technologies

CAD software tools are found in a wide variety of applications when it comes to technical design. The most popular file formats for CAD files include DWG (from Drawing) and DXF (Drawing Interchange Format), both developed by Autodesk, and the non-proprietary STEP (Standard for the Exchange of Product Data) file format, recommended mainly for 3D drawings. CAD drawings generally contain the geometric information of the object, without semantic information stored within it (rarely through connections with external data repositories). As this technology is well-developed and mature, there exist energy simulation software tools that allow for the interoperability with CAD files. One of this software is the DesignBuilder, which is a graphical user interface that connects to the EnergyPlus simulation engine, [65]. It allows for the import of DXF files and is used to perform day lighting analysis, thermal simulation of natural ventilation, visualization of site layouts, solar shading and sizing of HVAC systems, [151]. Most energy simulation software that allow for the import of CAD files are slowly being discontinued, these include i.e.: Ecotect, [11]), and Project Vasari (now replaced by REVIT), [10]. Frequently the software take the form of plugins, such as on AutoCAD products: EnergyPlugged, e-QUEST, Energy + or Insight 360.

However, due to initial design decisions pure CAD data formats show their limitations when it comes to data models in the built environment. CAD was designed to cover different and very broad use cases related to computer-aided design. To cover the emerging demand of complex and integrated engineering design and planning standardized Building Information Modeling (BIM) has emerged. A BIM model is a digital representation of physical and functional characteristics of a facility. As such it serves as a shared knowledge resource for information about a facility forming a reliable basis for decisions during its lifecycle from inception onward, [142]. It always contains a 3D graphical representation of the building and additionally focuses on the information data model regarding the object. Given that, not only geometry, but also semantic information, such as ownership, construction material, costs and energy-related data are stored in the BIM model.

Fundamentally, BIM assists with constructing buildings virtually, before their construction in the real world. However, it does not stop there, and further supports the building use throughout its life cycle, all the way into dismantling and demolition, [67]. What it brings in addition to CAD systems, is that it enables the creation of single and unified representations of the building, so complete that it can generate all the necessary documentation, [165]. Typical for BIM projects is a high level of detail, which helps in conflict detection and improves planning and construction. Overall, it delivers an optimisation in the quality of the entire building related process while enhancing interoperability when connected with the use of single all-inclusive software packages or open data formats, [93]. However, producing very detailed representations of the building results in labor-consuming processes. Hence, the usage of BIM is usually limited in terms of spatial scale to that of single buildings, or at most district.

Specifically for storing energy-related information, the BIM community has developed the Green Building XML (gbXML) data format, which enables interoperability between BIM software and Building Simulation Tools. According to Hardin and McCool, it supports model-based analysis for the energy simulations of the whole building, [97]. gbXML is an XML file that includes over 500 categories of elements and attributes used for the description of all aspects of a building related to its energy footprint: building physics, occupancy, meteorological parameters and the geometry. The gbXML schema was designed so as to enable interoperability between energy analysis tools and BIM software.

Concurrently, significant research effort has been made to provide automatic data flows that built gbXML data such as presented in [227]. BIM software tools that support gbXML include: Autodesk Revit, Bentley Architecture and ArchiCAD. gbXML files can be used in multiple simulation software, for example Open Studio ([94]), a full-featured software framework that supports rigorous, multidisciplinary building simulation. It is a collection of software tools that enable whole building energy modeling. With regards to the simulation engine used, it is that of EnergyPlus, [60], as well as advanced daylight analysis using the Radiance software, [230].

The most used and produced data format in the BIM field is the IFC format, [97]. The Industry Foundation Classes (IFC) specification is a neutral data format developed and maintained by buildingSMART International, an international organization which focuses on improving interoperability for software applications used in the building industry. IFC is developed as an open specification data format for describing, exchanging and sharing construction and management data across multiple applications in the building industry. The data model is object-oriented and contains elements used throughout construction or facility management projects.

The IFC format is also the international standard for openBIM, [109], a collaborative effort that is vendor-neutral with processes that supports seamless teamwork throughout project lifecycle. This initiative seeks to enhance interoperability by providing reliable data exchanges and collaboration workflows. The latest version of the standard is provided as an ISO standard - ISO 16739, [110], and can be described using either the STEP (.ifc) or the XML (.ifcXML) data structure.

One other potential data source for UBEM practitioners is the BIM Collaboration Format (BCF), a format used in the BIM community that allows multiple users to work on the same file concurrently, [13]. It grants multiple BIM applications the ability to communicate on model-based issues on already shared IFC models. There are two alternative solutions, one based on file exchanges between software platforms and another one using a web service. BCF is developed by buildingSMART International, and provided as a standard. What makes BCF valuable in the energy chain is that it can be easily exchanged or "roundtripped" (sent back and forward in between users). This stands true as long as all users maintain the integrity and uniqueness of the shared BCF file, [114].

Both gbXML and IFC formats represent a source of heightened interest for UBEM experts and provide interoperable data exchange while BFC creates a framework for data sharing within the same team and among institutions. The shift in the constructions sectors towards standardized BIM formats and the similarity of design spatial scales with UBEM is creating larger volumes of BIM data, therefor becoming a good source of information in the modeling of the energy chain. Table 4.2 presents both file formats and the BCF framework with their main characteristics.

While CAD tools and data may be encountered on each of the energy chain relevant spatial scales mentioned in chapter 3, BIM's use cases are narrower in terms of size. CAD software can be used for the detailed drawings and plans of single buildings as well as to work on larger regions, i.e. while planning and designing electrical grid within the city. Due to the fact that BIM models are usually very detailed and accurate depictions of reality, they are usually used to describe single buildings and rarely districts. However, with the advent of BIM in industrial workflows, it is expected that the amount of data available will steadily increase. Figure 4.2 depicts the spatial scales for which CAD, BIM and GIS technologies are employed.

For the handling of data within the applications related to this manuscript CAD and BIM data has been converted to GIS relevant formats described in the next chapter, SHP and CityGML. This was done mostly for the purpose of standardization of input data to the simulation tool aEneAs (Energy Assessment) and various user interfaces that were used. The conversions, energy model, workflow and architectural solutions used are described in detail in chapter 6.

TABLE 4.2: BIM data formats used in the urban energy chain

Format	Design use case	Stored elements	Data representation
IFC	Supports all building life cycle applications and software interoperability	Comprehensive object-oriented approach for all building parts and related processes (with class definitions)	STEP, XML
gbXML	Facilitates building design and energy efficiency studies	Building physics, occupancy, meteorological parameters	XML
BCF	Grants multiple BIM applications file sharing capabilities on IFC models	Sharing format	XML

4.3.2 GIS's take on cities - semantic city models and CityGML

The Geographic Information System (GIS) is defined as "computer-based tool for mapping and analyzing things that exist and events that happen on Earth", [206]. This conceptualized framework provides expertise for acquiring and analyzing spatial and geographic information. It provides urban planners and designers with very potent tools, [150], and has become a standard utensil in the field of urban building energy modeling, [191]. In the urban context, GIS had the advantage that it allowed for multiple layering and analysis of data, closely modeling real-world complexities. The spatial scales for GIS modeling activities in this setting are districts, neighborhoods, towns and cities as shown in 4.2. GIS is rarely used to describe or model one single building, see the definition of a building in 3, even as the standards presented in this section are able to do this, in order to facilitate conversion from BIM, CAD and other sources into it.

Multiple standards with the ability to store city data have been developed. They cover very large data types and specific GIS needs. Largely, these can be classified in mostly vector and seldom raster data types, as previously discussed in chapter 2. A very popular file format for GIS is shapefile (SHP). A shapefile stores non-topological geometry and attribute information for the spatial features in a data set. Shapefiles are composed of a set of vector coordinates that store the geometry of a feature. However, as their original design was based on a use case that was purposely wide in purpose, SHP files lack standardized data models useful for narrow, or specific, applications such as in the energy chain. Shapefiles only ensure uniformity in terms of unique identification of each geometrical item via the Object IDentification (OID)/Feature IDentification (FID) tool. Both OID and FID refer to a unique identifier of an entity within a data table associated to a SHP file. FID is primarily used in shapefiles while OID appears in geodatabases, a large repository solution that replaces files in the case of geographical data.

GIS data sets constitute a crucial component of the input of energy modelling and simulation tools and they have a significant impact on the quality of results, as shown in [80, 171]. Such input and its significance are discussed in detail in chapter

6, where the energy models library aEneAs - that uses Python/SQL is presented. The library seeks to avoid issues of standardization by using semantical city models. In an urban context, for creating and standardizing digital urban twins semantic city models have been the main solution employed in the field of UBEM, [191]. This combines both vector and raster data sources and represents a good basis for urban management and urban planning as [90] testifies. These models include:

- 3D geometry
- Thematic coverage (e.g. building, vegetation, city furniture)
- Topology (spatial relations between connecting or adjacent geometry)
- Semantics (e.g. the meaning of a surface door, window)
- Appearance (surface texture)
- LOD (level of details concept)
- Open standardization (to provide interoperability)

CityGML is the international standard of the Open Geospatial Consortium (OGC) for the representation and exchange of 3D city models. CityGML is based on XML and is used for the storage and exchange of such models. It has gained traction over the years to become a significant part of the open 3D data available in the developed world. CityGML distinguishes between geometrical and thematic models. A geometry model grants consistent and homogeneous definition of geometrical and topological properties of spatial objects within 3D city models. The thematic model of CityGML employs the geometry model for different thematic fields like DTM (Digital Terrain Models), sites (e.g. buildings; bridges and tunnels), vegetation, water bodies and transportation facilities.

CityGML 2.0 includes a concept of Levels Of Detail (LOD), presented in table 4.3. It is a hierarchical detailed approach where objects evolve to include more details with higher LODs with regards to their geometry and thematic differentiation, [91]. LODs are defined using data/point accuracy, minimal dimensions of objects and related model scale. Simply put, the LOD0 includes a 2.5 dimensional Digital Terrain Model, a ground mesh. Starting with LOD1, building blocks and classes are included while LOD2 brings in differentiated structures, textures and thematically differentiated surfaces. With LOD3 external building architectural models, as well as high-resolution textures are included (facilitated by the intricate surface description). The highest level, LOD4 finally includes interior structures, and has been seldom used outside of academia, as such models are usually built and managed within the CAD or BIM domains. CityGML can concurrently include buildings with different levels of information, which facilities the merger of multiple source data sets as well as efficient data visualization and data analysis.

What makes the standard stand out is the fact that it includes not only graphical aspects of city models and the representation of the semantic and thematic properties, taxonomies and aggregations, as seen in [91]. CityGML's classes and ownership relations for the most relevant topographic objects in cities and regional models with respect to their geometrical, topological, semantical and appearance properties support the aforementioned design choices.

National and city level public authorities have started providing CityGML data, as shown in [148] and [149]. This statement is valid for more than half the territory of

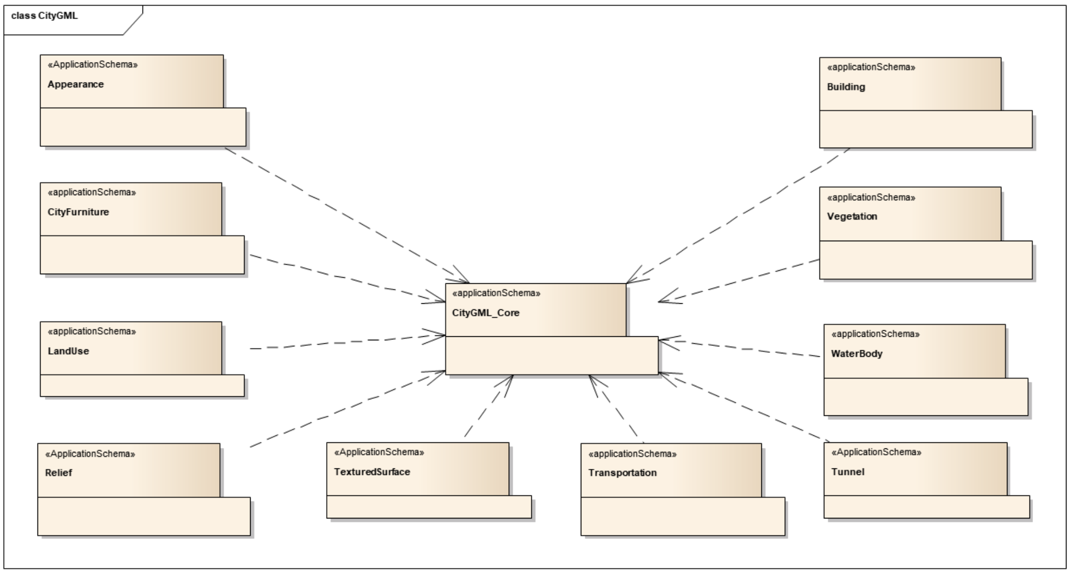


FIGURE 4.3: Classes of the CityGML standard, screen capture in Enterprise Architect

the EU. The EU has also adopted CityGML building feature class as a standard with the INSPIRE directive. In fact the main part of CityGML used in this thesis for description of the entities in the city revolves around the building class, significant for modeling information used in the energy chain. This simplifies any ETL operation that originates in the INSPIRE format. Figure 4.3 presents a diagram of the CityGML core and the included classes.

When the concepts surrounding the thesis were developed, CityGML had been updated to its second version 2.0. Summarized from [90], CityGML is:

- XML data model for exchanging 3D city models
- Open and international standard from the OGC since 2008
- Containing geometry, topology, semantics, visual appearance
- Based on Geographic Markup Language (GML)
- Describing entities and their spatial and non-spatial attributes, relations, and their complex hierarchical structures in five levels of detail.

Beside the building class, there are other CityGML classes that did not come into use in this thesis' experiments, because the standard is used to describe all entities in a city. On the other hand, CityGML has its own limitations, which a reference benchmark study can attest to, [170]. As this manuscript finds its way to the printer, CityGML is having a new upgrade in terms of content and syntax top, to 3.0; [124]. The scientific community is divided as to whether a new version was required at this stage or rather increased use and implementation methodologies that facilitate wider use of the standard. Certain constraints and drawbacks are corrected with the upcoming version 3.0 of the standard, however, the solutions presented can also have an enhancement effect with regards to difficulties in implementation.

As a data model's advantages and disadvantages can only be properly understood and tested when integrated in data flows, different implementations of the standard were analyzed. The open database schema version of CityGML, the so called

TABLE 4.3: CityGML 1.0 and 2.0 level of detail concept scale and accuracy requirements, adapted from [91] and [242]

Aspect	LOD0	LOD1	LOD2	LOD3	LOD4
Model scale	region	city/region	city districts	architectural model external	architectural models external and internal
Accuracy	lowest	low	middle	high	very high
Position / Height absolute accuracy	lower than LOD1	5/5m	2/2m	0.5/0.5m	0.2/0.2m
Generalization	maximal	as object block	as object	real as object	constructive

3DCityDB, offered a great tool for connecting existing information and standards to spatial entities. The open software tool, described in full detail in [238], offers solutions for storage, analysis, management, interaction, and visualization of the 3D city models stored in the standard. It works both with PostgreSQL and Oracle, the second and third most used SQL (Structured Query Language) DB engines on the planet in 2019, numbers presented in [205].

In the particular case of this manuscript the applications developed to test the research questions use the aforementioned database schema within a spatial relational database, PostgreSQL with the PostGIS spatial extension. The package has an associated set of procedures that allow one to import, export and manage virtual 3D city models according to the CityGML standard. These procedures greatly facilitate the use of CityGML. 3DCityDB was used in association with other non standardized data sources. For that purpose the persistent identifier PID system presented in chapter 3 was used and links to the entities in the schema by calling on the UUIDs of each object were made.

TABLE 4.4: CAD, BIM and GIS standards/methodologies summary

Method	CAD	BIM	GIS
Spatial scale, see figure 4.2	Single building, District, Neighborhood, Town, City	Single building, District, Neighborhood	District, Neighborhood, Town, City
Formats	.dwg, .dxf, .stp	.ifc, .gbXML	.shp, .gml
Software	AutoCAD, Inventor, FreeCAD, QCAD	Revit, ArchiCAD, BIMserver, BIMVision	QGIS, ArcGIS, CityEditor, PostgreSQL/PostGIS
Energy simulation tools	DesignBuilder, Ecotect, EnergyPlugged, e-QUEST	Open Studio, Green Building Studio, EnergyPro	CitySIM, aEneAs, EDF Curtis, SimStadt, CitySim+, City 4.0

Table 4.4 summarizes the information that regards CAD, BIM and GIS spatial scale, formats, software, and energy simulation tools presented in this section of the paper. Further information on simulation tools that use GIS is presented in chapter 6. For all intents and purposes CityGML remains the only GIS related standard with

capabilities that support performing energy simulations at city scale using buildings as the smallest measurement unit. Complete data sets (spatial and energy thematic data combined) from single sources describing the characteristics of the objects to be simulated are difficult to come by, a topic discussed in chapter 3. As such, these models are generated with the help of different data sources from different energy chain actors.

4.3.3 Bridging the divide - ADEs to the rescue

One of the great advantages of using CityGML is the possibility of employing the Application Domain Extensions (ADE) mechanism. This allows users to store additional information, e.g. energy-relevant data, directly in the spatial model. The ADE mechanism has been proposed as a solution to assist in the use of standardized city wide models in energy modeling. It extends the schema with new classes and attributes which are not explicitly modeled in CityGML. An ADE is defined in an extra XSD (XML Schema Definition) file (a special XML document that defines the properties of, and allowed relationships between XML features) with its own namespace, [57]. This file has to be explicitly imported with the XML Schema Definition of the extended CityGML modules as it is used to validate the integrity of an XML file. By these means it is guaranteed that all the elements found in the file satisfy the conditions laid out in the XSD file. ADEs can be defined by information communities which are interested in specific application fields. Further, the UML (Unified Modeling Language) modeling approach for ADE's is to convert UML to XSD relational DB schema which in turn allows for database working methodologies to be developed surrounding these data models. As [25] indicates there are multiple advantages to applying the ADE approach:

- formal specification of extensions
- semantic and syntactic interoperability for application specific information
- validation via CityGML and the respective ADE schema
- multiple concurrent ADE usage possible

As ADEs use their own unique namespace when defining features, naming conflicts are avoided between different ADEs when using multiple. This makes it possible to employ multiple ADEs simultaneously within the same CityGML document in a *jigsaw* like manner. ADEs can therefore add new features to a semantic city model, without adding reliability issues. Figure 4.4 depicts such a situation where three extensions-the Energy ADE, the UtilityNetwork ADE and the Dynamizer ADE, are proposed as an all encompassing data model for the energy chain.

Modeling buildings' energy data. The Energy ADE

The Energy ADE is an extension of the standard CityGML that uses the standardized extension mechanism in order to create a standard where energy information required for urban modelling can be found. It aims to provide interoperability to existing standards and to be useful at all scales in between single building and city wide, [2]. The Energy ADE has been published in a first and second stable versions, v1 and v2, which has allowed for its exploiting in research and development projects. For example [194] demonstrates a workflow designed to work in the UK spatial context for the creation of residential scenes encoded using CityGML and

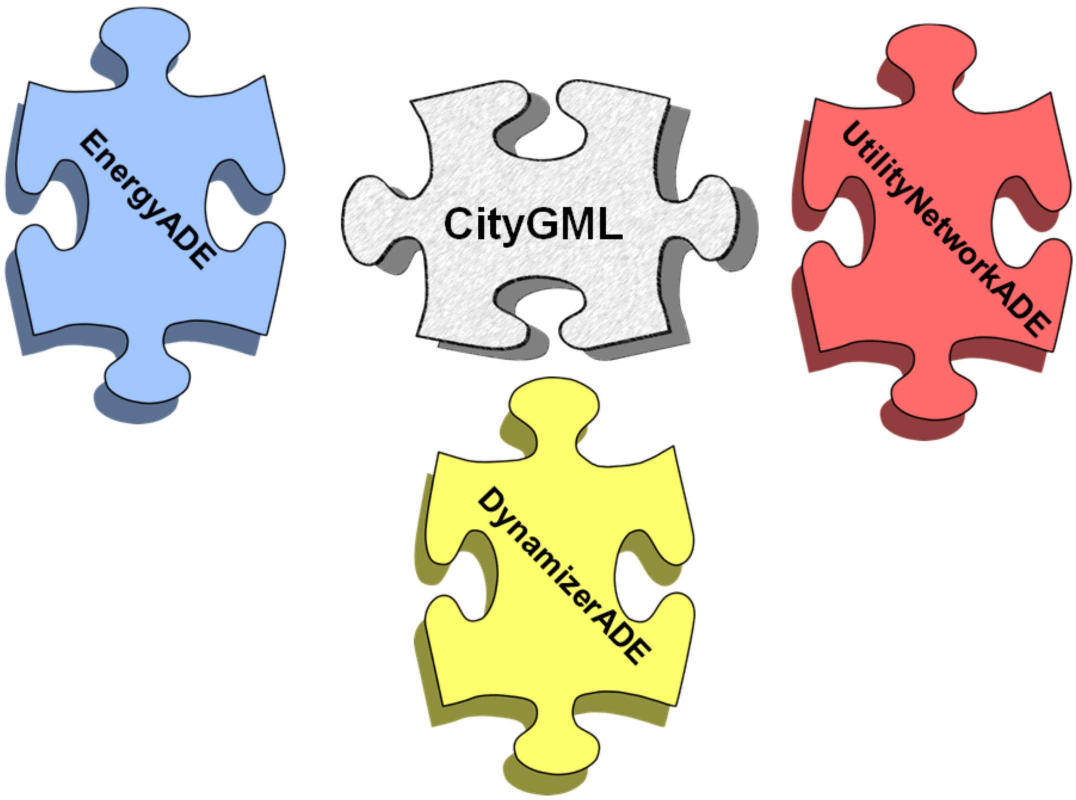


FIGURE 4.4: CityGML and extensions jigsaw, presenting the CityGML extensions Energy ADE, UtilityNetwork ADE and the Dynamizer ADE connected to the core

Energy ADE which facilitates energy simulation at urban level with building scale results.

Figure 4.5 depicts the way in which the Energy ADE extension is built on top of the Building class of CityGML and on entities and features defined in the international standard ISO 19136 GML. Also included is the Timeseries ADE, another extension, used internally at the KIT. In this specific case it is employed in order to accommodate time series data (a common occurrence in the energy domain, e.g. occupancy profiles or measurement data from meters).

The CityGML Energy ADE enables the calculation of energy flows in buildings, based on stored and managed data, [79]. For performing detailed energy calculations such as energy gains and losses, the envelope of the building (including external installations, e.g. outdoor window shades or solar panels) is considered the physical boundary of the data model and is stored in the new attributes, ThermalOpening, ThermalBoundary and ThermalZone. At the same time, the envelope is also associating it with the geometry of features in the *old* Building class.

The data model also allows for energy systems mounted inside buildings to be modeled. A significant design choice was not to include large energy producing installations that supply multiple buildings and are connected to grid infrastructure. Another ADE, the UtilityNetwork ADE, can be used to handle such installations, see the following section for more details. The Energy ADE allows for buildings energy input to be modeled within each unit inside the data model. This is also true for energy output, in case the building houses a prosumer.

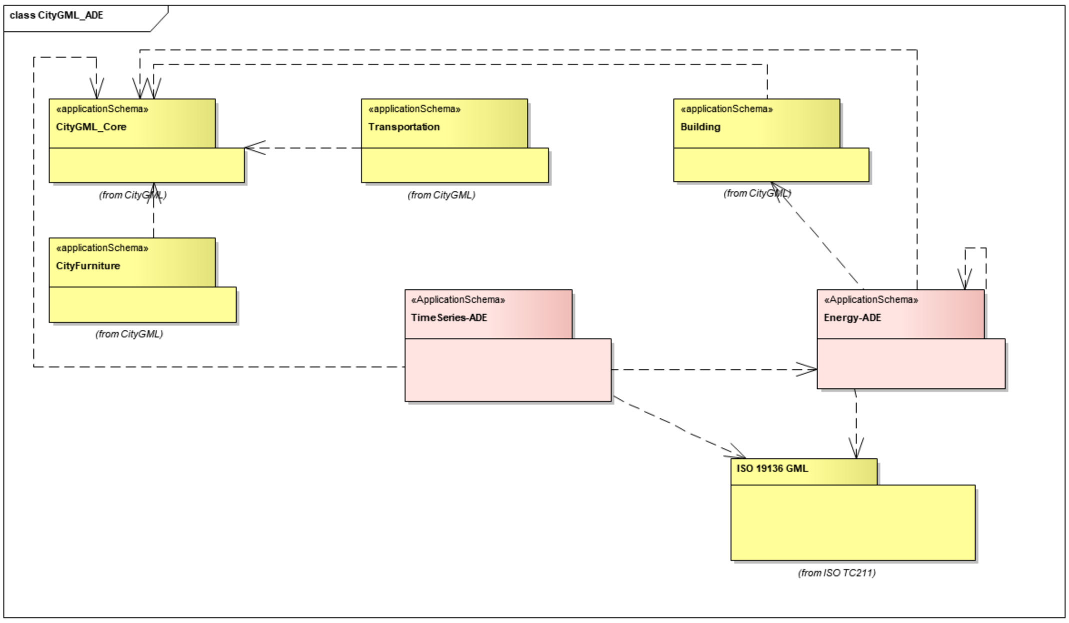


FIGURE 4.5: CityGML and the Energy ADE linkage, screen capture in Enterprise Architect

Figure 4.6 presents the internal structure of the EnergyADE v1. It consists of four modules: Occupant Behaviour, Material and Construction, Energy Systems and Building Physics, supporting classes and a core model. This structure can be extended by adding new modules or adding new classes to the modules (a standard ADE feature). It includes the possibility for further data/concept model extension and includes interoperability between the modules. The modules in the data model can be filled on an optional basis, depending on the purpose of the utilization.

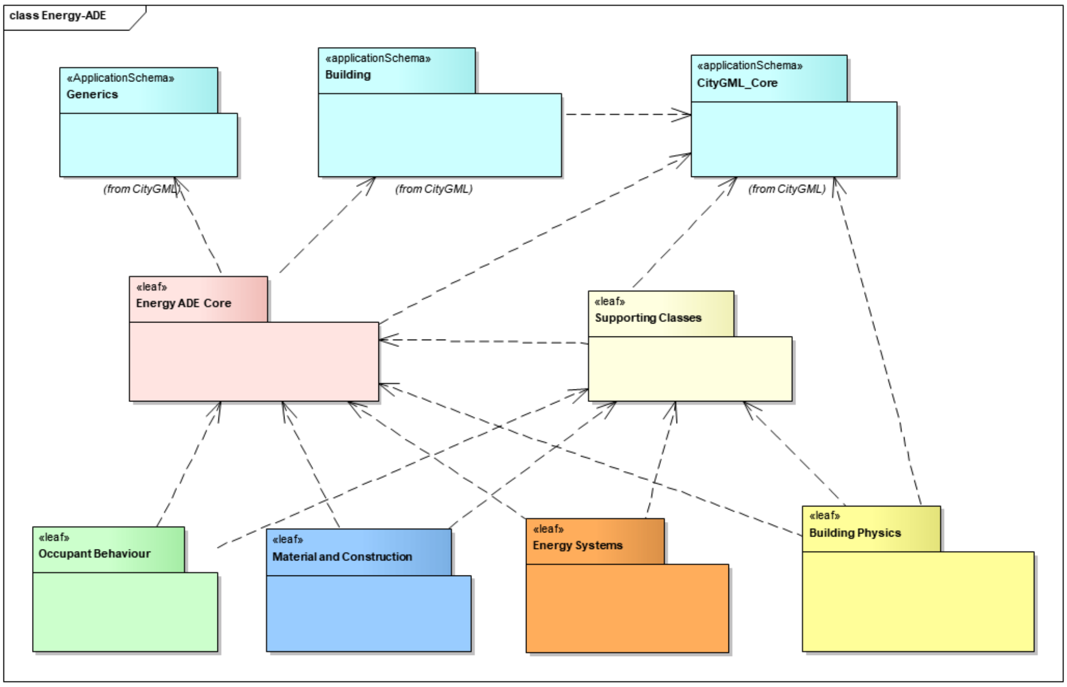


FIGURE 4.6: Energy ADE classes, screen capture in Enterprise Architect

Urban lifelines. The UtilityNetwork ADE

A utility network allows the flow of commodities. These are defined and classified in chapter 2. Modeling such complex environments is very often done using proprietary and unique solutions developed to satisfy local criteria, [21]. In [34], the requirements to build a multi-network (that includes multiple types) model are identified, that is a standardized set of defined feature types capable of modelling all the constituent pieces of the network. The model needs to include standardized feature definition, the ability to model different network types, have both network and feature hierarchy, support topological connectivity and include both explicit topographical and topological feature representation. A number of alternative standardized data models exist for representing utility networks on an abstract level. This means that they are able to represent different kinds of networks (such as water, gas, electric, telecommunications, etc.) within the same data model.

Assessing three of the most well used data models, Becker and colleagues, [21], looked at the IFC network data model, the "Network" package of the Infrastructure for Spatial Information in Europe (INSPIRE), as well as the ESRI ArcGIS Network Model, used in many GIS applications. INSPIRE is an EU drive for a standardized general spatial data model in Europe, while ESRI is the largest software manufacturer in the GIS domaine. Becker and colleagues found that two of the models had limitations in the context of spatial scale (among other various ways). The INSPIRE Network package did not have the capability of decomposing individual network component elements into constituent elements, meaning that a single model is really only suitable when working in single spatial scale. The IFC network data model features are specified only in a local reference system, meaning that they are suitable only at the single building scale. The ArcGIS Network Model, however, has the capability of aggregating and decomposing network features to represent networks at multiple spatial scales, and it can be used in any spatial coordinate system, meaning that it would be quite suitable for modelling at different spatial scales.

The Utility Network ADE is a more recent and in-development extension for CityGML that uses the ADE mechanism. Figure 4.7 presents the extension and the links to the GML standard, as well as CityGML and its Core, Building and CityObjectGroup classes. It is designed to help store, by use of semantic city models, features related to the generation, distribution and consumption of media such as water, electricity, and gas which are considered essential in the modern urban environment, [33]. The extension is meant to cover all conceivable spatial and semantic scales, and to deal with existing gaps in current network data model alternatives, [20].

The consortium working on the development of the standard is also looking at the definition of the level of detail concept for utility network features. The current version of the data model is defined in such a way that it allows for networks and network features to be defined as sub- or super-ordinate to others. Functional, protective or otherwise special elements along networks such as valves, meters, storage devices, protective coatings are also supported. One of the key aspects that distinguishes this standard to other existing ones is the inclusion of a topological network model, [125]. It comprises the core of the Utility Network ADE data model and ensures that all kinds of networks defined in a single CityGML model are represented and can be operated upon in a unified way. This means that multiple types of networks, transporting different types of medium and at any scale, as well as the relationships and interdependencies between them can be effectively modelled using the UtilityNetwork ADE. In addition, due to the nature of the ADE mechanism

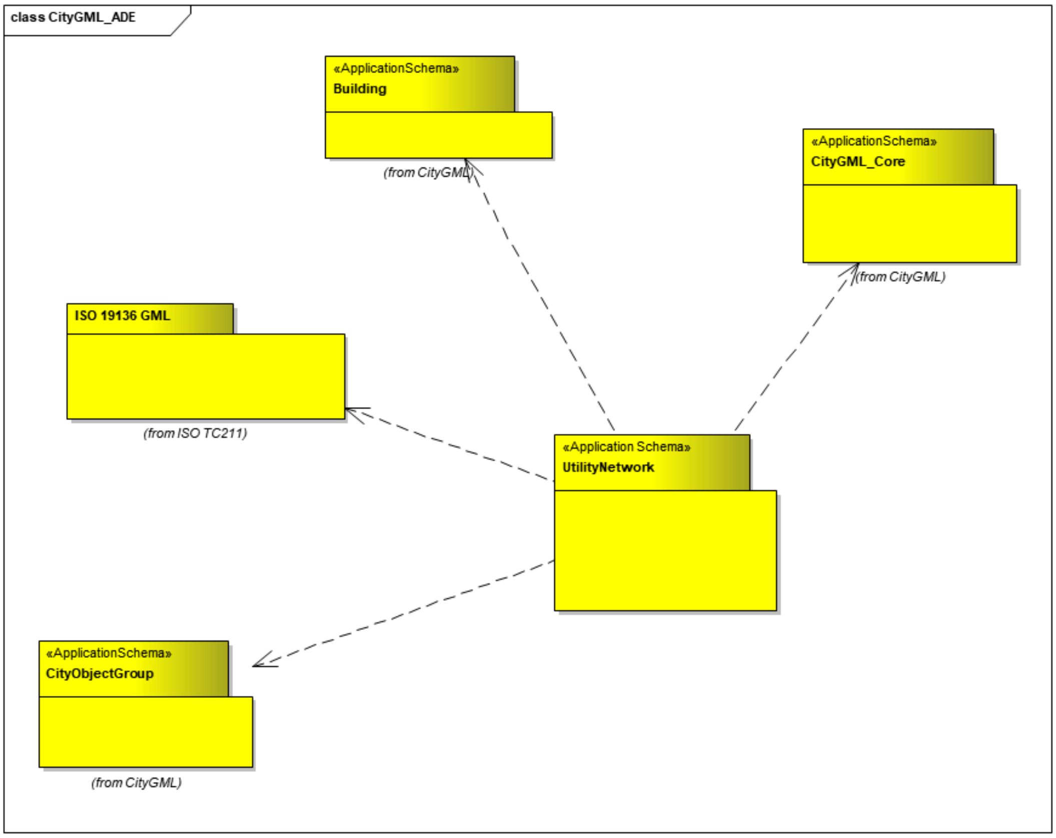


FIGURE 4.7: CityGML and the UtilityNetwork ADE link, screen capture in Enterprise Architect

such modelled systems are already compatible with existing CityGML-based city models. Figure 4.8 depicts the core module of the UtilityNetwork ADE.

A more recent attempt at comparing standards in use and development for utility networks, is found in [34]. The conclusion of the study is that the UtilityNetwork ADE can handle multiple types of networks and features in a common infrastructure. At the same time the functionalities that are required with modeling complex utility networks interactions are present. Additionally, a use case for interconnected water and electrical networks is presented in the aforementioned manuscript. It leverages the extension for modeling the two intertwined networks with water flow being used to generate electricity.

Summarized in table 4.5 are the capabilities of different utility network modeling solutions. These are compared from the different perspectives of the integrated features: standardized features definition, different network types capabilities, network hierarchies, feature hierarchies, topological connectivity, and explicit topography versus topology. The table also includes work from [23] which focuses its analysis on the LandXML standard. This standard is recommended by the OGC for use with utility network modeling and includes a very simple topology model. The standard is specifically designed for open canals and waterways that transport water, wastewater, storm water, and other similar products, and the associated properties are similar in scope. This design decision is the reason for which the relevant pipe element model structure is only basic (parametric geometrical representation, and a spatial reference that includes point coordinates).

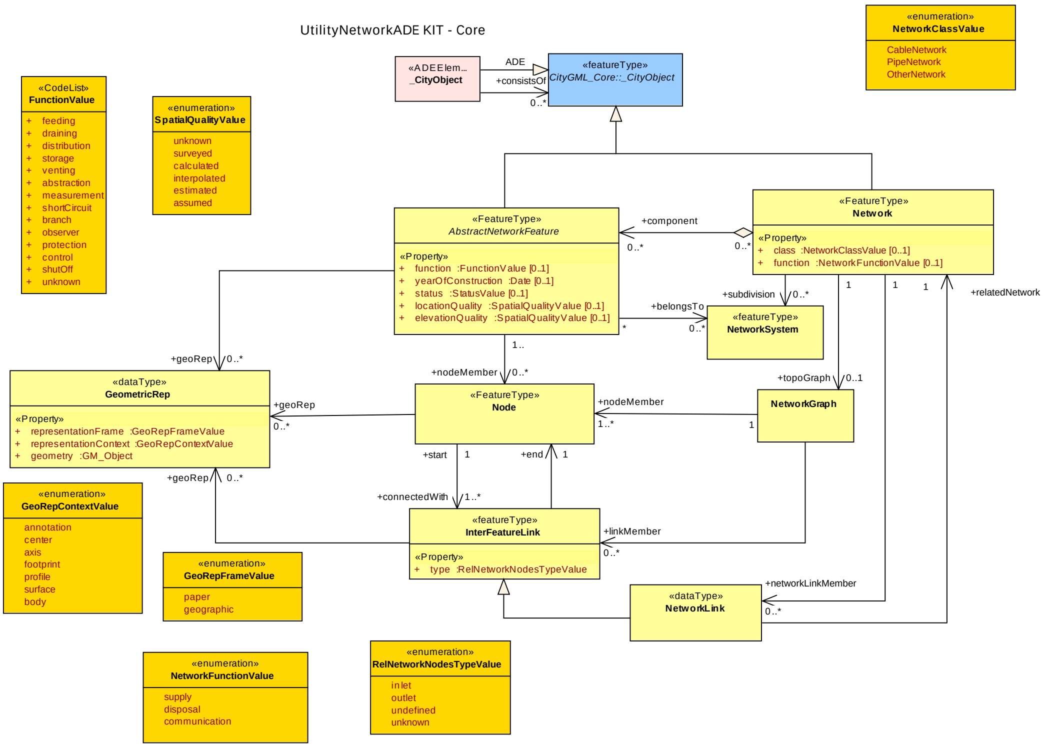


FIGURE 4.8: The core class of the UtilityNetwork ADE link, screen capture in Enterprise Architect

In order to classify the solutions depending on the level of support provided for different features, each support level was categorized from 1 to 3 in the table (1 - poor, 2 - basic, 3 - good). This indicates that the most complete solution is provided by the UtilityNetworkADE, with runners-up being a tie in between LandXML and IFC. These three solutions have different spatial design scales and scopes, the first (CityGML coupled) is city oriented, the second (LandXML) directed to use at national level, while the third (IFC) is aimed at building to district level. As such, this further indicated that the UtilityNetworkADE matches the use cases presented in the dedicated section, 4.1. The other two runner-up solutions can facilitate similar use cases, however, at different spatial scales.

Making a case for comprehensive data models of utility networks, [178] explains that in the context of utility network renovation processes there is a lack of knowledge of other existing networks. Most often, utility network operators operate their own networks in close vicinity of other networks. When cities or companies concurrently refurbish their networks, each organization is responsible only for their own network. However, due to to the nature of underground work and possible interactions they should be forced to work together and share responsibility. Cities often oblige these utilities to make public works at the same time, so as to avoid unnecessary repeated long duration construction sites in the same locations. This justifies case for the further development and deployment of the UtilityNetwork ADE.

4.3.4 CityJSON, a convenient alternative to CityGML - data interchange formats

In the GIS world two data interchange formats are currently most used, XML and JSON. Selecting an acceptable data interchange format is usually dependent on the

TABLE 4.5: Utility network modeling solutions, after [34], [23] and own work, with three support levels, 1 - poor, 2 - basic, 3 - good

Aspect //////// Solution	Design scale	Standard feature definition	Different network types	Network hierarchy	Feature hierarchy	Topologic connectivity	Explicit topography vs. topology	Result
Anylogic	Dwelling to National	1	2	2	2	2	2	11
ArcGIS	Building to National	1	2	1	2	3	1	10
CityGML Utility Network ADE	Building to City	3	3	3	3	3	3	18
IFC	Dwelling to District	3	2	1	3	3	3	15
INSPIRE	City to National	3	2	1	1	3	1	11
LandXML	Regional to National	3	2	3	3	2	2	15

data transmission rates and performance available to the proposed applications. XML or the eXtensible Markup Language, is a language designed for universal data representation that is human readable and has a fundamentally simple design, [57]. It also includes a user-defined hierarchical structure, which allows for and represents another advantage in using the scale based unique identifier solution proposed in this manuscript in chapter 3. On the other side, engineered to be a data exchange language, JSON is equally human readable and directly supported inside JavaScript, Python and Ruby. This is also the reason for which it is the most popular data format for sending API requests and responses. A standardized schema or meta-data definition was proposed for it in [184]. That allows the developers to specify the structure of JSON documents.

In the case of this manuscript, the applications developed used a central data repository based on a database whose import-export mechanism revolves around XML based files. In literature, comparisons of the use of JSON and XML have yielded different results. Where [172] find JSON to be significantly faster (parsing is done up to one hundred times faster than XML), [96] find that XML comes slightly on top considering multiple user perspectives. This is why JSON usage is a significant alternative to XML and is mentioned in this manuscript as a worthy alternative.

As stated in its dedicated sub-chapter, the standard CityGML was developed based on XML and GML. However, GML encoding has been shown to inhibit development, [133], by way of difficult CityGML files parsing, different geometrical data representations, XLinks, and a lack of implementation support for ADEs. Providing an easy-to-use alternative to CityGML, CityJSON uses a simpler JSON encoding that streamlines file parsing operations and allows for 3D models/digital twins inside these files to be easily visualised, manipulated, and edited.

4.3.5 Accommodating the leviathans - an ETL funneling system

The title of this section hints at the *elephants in the room*. That is, the BIM, CAD and GIS domains data served/provided either as files or as web services to energy modelers responsible of running city wide simulations. It is one of this thesis' hypotheses that CityGML supports performing energy simulations at city scale using buildings as the smallest unit. The hypothesis furthers that using CityGML as a backbone for data modeling and linking of entities allows for energy data to be properly allocated not only by identification means but also spatially. This section of the paper looks at the infrastructure backbone that facilitates data input and conversion into the CityGML format.

In the previous sections the types of data (format and transmission mode) from these domains are discussed and their significance to the UBEM field is emphasized. Table 4.6 summarizes the information that regards CAD, BIM and GIS spatial scale, formats, software, and energy simulation tools presented in this section of the paper. Further information on UBEM and software developed during this thesis for energy simulation is presented in chapter 6.

TABLE 4.6: Standards/Methodologies information sheet, aggregated from own work and [79]

Method	CAD	BIM	GIS
Spatial scale, see figure 4.2	Single building, District, Neighborhood, Town, City	Single building, District, Neighborhood	District, Neighborhood, Town, City
Formats	.dwg, .dxf, .stp	.ifc, .gbXML	.shp, .gml
Software	AutoCAD, Inventor, FreeCAD, QCAD	Revit, Autodesk BIM 360 Glue, ArchiCAD, BIMserver, BIMVision	QGIS, ArcGIS, CityEditor, PostgreSQL/PostGIS
Energy simulation tools	DesignBuilder, Ecotect, EnergyPlugged, e-QUEST	Open Studio, Green Building Studio, EnergyPro	CitySIM, aEneAs, EDF Curtis, SimStadt, CitySim+, City 4.0

Complete (this attribute should be taken with a grain of salt) spatial and energy thematic data sets originating in single data sources that describe the characteristics of the objects to be modeled and simulated are difficult to come by, a topic discussed in chapter 3. These have rather a speckled and inconsistent character. As such, the digital real world replicas (models) are generated with the help of different data sources from disparate and very diverse energy chain actors. Therefor, the creation of accurate digital city models and their following enrichment are time consuming and expensive ETL data processes. To expedite this process and ensure the usefulness of data, it must be usable for a wide range of applications. One way to ensure this

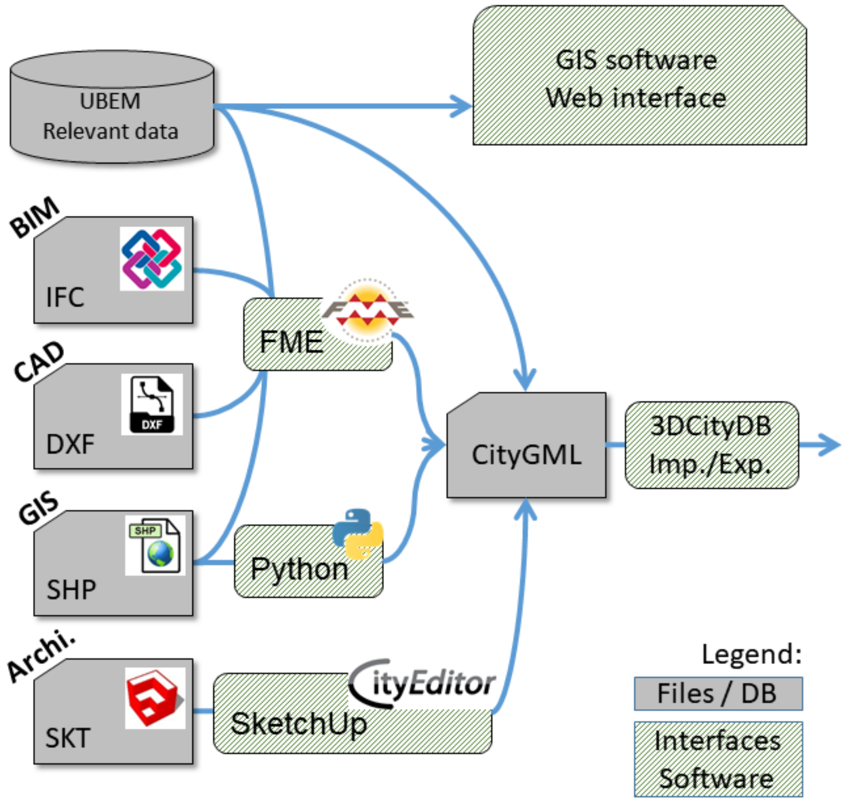


FIGURE 4.9: Accommodating the leviathans - an ETL funneling system

is the use of open standard formats such as CityGML. That was also the solution chosen within the applications developed to help test the research questions posed in this thesis. This aspect is discussed in [231] where different methodologies for the rapid production of energy modeling prone city models are presented. For the work performed to test the hypothesis in this manuscript the common feature of all those formats is their open data structure.

Often, differences arise in between coordinate systems, which are at times local Cartesian coordinate systems in BIM and CAD formats and geographical or geodesic for GIS. The second category is most often included in the standardized 4-5 digit EPSG system. This lack of *foresight* in data creation with local coordinates creates further issues in handling combined GIS-BIM-CAD workflows. For this thesis the conversion issues were handled in software tools such as PostgreSQL (with the extension PostGIS), FME, FZKViewer, QGIS or ArcGIS, City Editor, gModeller (both of the last two are software extensions built on top of SketchUp).

Figure 4.9 depicts the funneling system and the tools used in the data conversion. Most often, data provided is non-standardized. As such it requires extensive quality testing before any standardized workflow can be applied to it. These processes, referred to as ETL, Extract Transform Load, are previously discussed in chapter 2, and can sometimes use more than half of the actual UBEM modeling time.

CityGML spatial models can be generated in multiple ways. In the case of the work behind this manuscript it was mainly produced with three software solutions. In the first case BIM or CAD (in IFC, gbXML or DXF format) data are extracted and converted using FME. The same software can also add other UBEM relevant data to CityGML, providing quite a potent solution to ETL operations as it can generate complex buildings in multiple LODs. In addition, the Energy ADE and the UtilityNetwork ADE extensions can be used to populate multiple types of data used in UBEM simulations.

An extensive literature search showed that first endeavours in merging data from IFC into CityGML were published in [129] with the GeoBIM extension being developed for CityGML. The extension is integrated in the BIMserver serving as another potential entryway for GIS-BIM links. Other attempts at linking the two worlds have started to bear fruit. These efforts provide a similar tool to the database version for CityGML DB, the 3DCityDB, [100]. However, as can be seen in the cited manuscript, these efforts achieve limited success, for example on visualization, less so on accurate IFC data positioning, as conflicting design choices and use cases behind GIS and BIM make the merge difficult.

For GIS data that is stored in SHP files, Python code was developed that can create a CityGML building based on the building footprint and the elevation. This generates a flat roof building in LOD1. With additional surface classification provided by Python code a hybrid LOD2 model is created. A significant downside is that it contains only flat roofs, however, the volume calculations of the buildings interior can be adjusted statistically in order to take that into consideration (for example via building cadastre code). As such the data produced can provide significant advantages in energy modeling.

The last method that the figure presents uses data that stems mostly from architects and town planners (urbanists). These professionals often use graphical design tools that provide less in the direction of a standardized data structure. This was recognized by software companies, and extensions have been developed that offer tools to export standardized data. This is the case for SketchUp and the CityEditor extension. This tool includes a workflow that allows the extraction of CityGML data from a SketchUp 3D model, in this thesis' case LOD2 building models.

Out of all the tools presented, FME is the most complete solution. It can help ETL processes by providing semi-standardized funneling conversion systems. In addition it can also integrate code snippets (in Python in this manuscripts' case) to provide for the use of already implemented tools. Finally, the funnel system presented above works for file based standards. With regards to web services integration the next section is dedicated to them.

4.4 Web services lend a hand

The need to retrieve data from distributed sources brings with it connectivity and functionality problems. Here, the adoption of a Web Service could provide remedy, especially in the context of near real live data streams that produce large volumes of data. An Application Programming Interface (API) considers both the provider and the receiver in a IT system communicating over the internet, [123]. The web service is the provider (server-side) portion of the API, an interface programmed to include a pre-defined request-response message system. In very simple terms, web services provide solutions that allow the sharing of data in between parties that host data and others that require it without the sharing of the entire data set, or as [122] explains, web feature services allow clients to only retrieve or modify the data they are seeking, rather than retrieving a file that contains the data they are seeking and possibly much more. Web Services generally involve communication over the internet using protocols like Hypertext Transfer Protocol (HTTP), Simple Object Access Protocol (SOAP), REpresentational State Transfer (REST), and eXtensible Markup Language-Remote Procedure Call (XML-RPC) as a means of communication.

As [162] demonstrate, being able to easily discover relevant data is a prerequisite to unlocking the potential of data. Online sharing of spatial data faces similar challenges to other distribution methods, such as file based. These include provision of descriptive and structural metadata to enhance data's discoverability and reusability. The W3C (World Wide Web Consortium) recommends the following best practices for data sharing on the web, [226]:

- valid licensing information,
- transparency regarding data provenance and quality,
- data versioning to deal with data update on a scheduled basis
- provision of unique identifiers,
- interoperability enforcement by way of standardized, locale-neutral and rich data formats,
- improved data accessibility and usability from flexible and stable API.

Several web service standards, often used with sharing spatial data in the energy chain, are covered in this section. Easy and stable access to data on the web allows users to take advantage of the advanced web infrastructure which often means that by default the HTTP method is used. This provides access to data using atomic transactions (atomicity is the defining property of database transactions, where each SQL query gets wrapped in its own transaction). In practice most data providers or delegated third parties support data access either through bulk download or through an API, or both of them.

When specifically discussing the issue of sharing spatial data, most often interoperability issues arise due to the data heterogeneity of this data category. More specific, most data providers might use different data formats with their own data models, which requires a special and tailored software solution to share their data. These data-driven implemented web services cause issues for data consumers, as they are locked to specific vendors and increase the difficulties of sharing of the spatial data. To overcome the interoperability issues, the OGC has been established. This consortium aims at providing standards both in data model and web services. The purpose is to make spatial data sharing across any network, application or platform routine.

4.4.1 GIS web services

Responsible for the standardization in this sub-domain of the GIS field is the OGC. The organization has a long list of web services to show for its activity and these have been crucial for the development of WebGIS, or GIS on the Web (Internet), the use of Geographical Information Systems data online. For the better part of the last two decades OGC has been a great *locomotive* for change within the GIS community. It has helped develop and encouraged the implementation of different web services. This section presents web services used for sharing spatial data produced in the GIS *world* that are relevant for the city wide energy chain.

- Web Map Service (WMS)

- Web Coverage Service (WCS)
- Web Feature Service (WFS)
- WFS in the 3DCityDB

In order for web services to function together, the OGC Web Services Common Standard (OWS) was developed. This is a standard which specifies those aspects that are common to all or multiple OWS Interface Standards. In its scope fall contents, parameters and data structures, and XML, KVP (Key Value Pair) or SOAP+XML encoding of Operation Requests and Responses, [39]. The OGC Web Services are implemented on Distributed Computing Platform (DCP). DCP supports HTTP, where the online resource of each service operation is the Uniform Resource Locator (URL). The HTTP Post, Get and Response methods are the mainstay of these web services.

The Web Map Service

The WMS provides a simple HTTP interface for requesting geo-registered map images from one or more distributed geospatial databases, [19]. It can be used to request images from multiple WMS servers concurrently so that images can be overlaid into a single view for users or included into a data handling *pipeline* (workflow). The output format is of multiple types, among those: PNG, JPEG, GIF, TIFF, SVG, PDF, KML. It also provides three request operation types, invocable in the form of URLs. These are:

- GetCapabilities, a mandatory request that obtains service metadata (description of server information and request of parameter value)
- GetMap, also a mandatory request that provides a map image whose geospatial and dimensional parameters are well defined
- GetFeatureInfo, an optional request available to communicate information about particular features shown on a map

The Web Feature Service

WFS is a web service used for querying and editing vector geographic features. It has radically changed the way in which geographic information is created, modified and exchanged over the internet, [158]. The ability of sharing geographic information at a feature level rather than file level has given users previously unmatched flexibility. WFS is primarily used as Feature Accessing Service, but it can also be a Feature Type Service, a Coordinate Conversion Service as well as a Geographic Format Conversion Service. For encoding both XML and KVP are available. WFS supports ad hoc and stored queries. According to the specification document, [225], it provides multiple operations:

- Discovery Operations, allows clients to determine its capabilities and retrieve the feature type,
- Query Operations, enables the client to retrieve features or values of feature properties,
- Locking Operations, grants data lock for single and exclusive access to (a) feature(s) in order to modify or delete it (them),
- Transaction Operations, provides the ability to create, change, replace or delete features,

- Stored Query Operations, sanctions custom made parameterized query expression stored by server.

A version of the WFS, designed to work with semantical city models stored in the 3DCityDB based either on PostgreSQL or Oracle is presented in the experiments dedicated section of this chapter, 5.2.3.

Web Coverage Service

The WCS is a web service that supports retrieval of geospatial data known as "coverage" (geospatial information for space-varying phenomena), [176]. In this context, with coverage, the OGC implies a feature that acts as a function to return values from its range for any direct position within its spatial-temporal domain. This web service can be employed for any property, e.g. surface reflectance, concentration of pollutants, number of lanes, and can originate from different domains such as a utility network/grid, or a transportation network. The web service can handle requests of geospatial raster images, process results for complex modeling and analysis and perform complex querying of coverage data. It is composed of a Core (that specifies a set of requirements the WCS must fulfill) and an Extension (that brings further functionality and defines means of protocol binding). Similarly to WMS, the WCS returns raster formats, however with the WMS the spatial data to be portrayed is static while the WCS adds dynamicity and provides available data and descriptions. The binding is more flexible as it can be any of the protocols defined in the WCS Extension. The web service includes three operations:

- GetCapabilities, provides metadata regarding the server's capabilities and coverages offered
- DescribeCoverage, seeks details on one specific coverage
- GetCoverage, asks for a selected range in a coverage according to properties of an elected spatial-temporal location

4.4.2 CAD and BIM web services

The focus of this manuscript lies with open source standards and data. Within this domain there are few web services developed that can serve CAD and BIM data, even though as far back as 2010, members of the academia, [116] and [216], were underlining the need for such services, that can provide similar functionalities to the ones shown in the previous section. Open source architecture and prototype solution have surfaced, as seen in [243], however, none have gained similar traction and notoriety when compared to the ones in the GIS domain. This is due to a number of factors, the most cited ones being issues with handling identification and multiple editor issues. [135] proposes solving the identification issues using blockchain with the drawback that the proposed system could be architecturally slow. The potential of web services development in the CAD domain can also be seen in the attempt of one of the largest software manufacturers to complete user surveys that help define use cases, such as [9].

In fact the only non-proprietary web service provided in the BIM community of buildingSMART Internatonal is the extended web service-based (RESTful) API mode for BCF, [114]. In its web service version, the BCF-API is a RESTful service that allows separate BIM file users to connect directly or to be connected via a server. This requires operating a BCF server in parallel (that can double as the BIM server). All

BCF data is stored in the cloud and enables concurrent users to sync BCF data in one centralized location. The existence/presence of dedicated third-party BCF servers are the main purpose for the development of the OpenBIM initiatives and help act as the hub for such communications enhancing shareability among BIM users. BCF is a powerful tool, similar to the WFS of the OGC, for BIM related data in the energy chain and will impact the management of energy related facilities once BIM and BIM data become part of the *modus operandi*.

The alternative to using the BCF-API for data sharing is the use of the open source software BIM server. The Building Information Model server (BIMserver) enables the handling and stockpiling of the information of a construction project by using the open data standard IFC, [128]. The BIMserver uses a model-driven architecture approach where IFC data is stored as objects, similar to a database where operations such as query, merge and filter are possible. It also contains extra features (e.g. model checking, versioning, project structures, merging). Other similar software in terms of capabilities are all commercial solutions, [144].

4.4.3 Sensor frameworks

The previous sections take a rather spatially oriented perspective to the data shared in the energy chain, be that file or web service based. The two following subsections briefly introduce solutions that bring a sensor/*things* mindset. These are the two sensor enabling frameworks that were used within this thesis' experiments, the OGC Sensor Web Enablement (with two alternatives, SOS-SES-WNS or SensorThingsAPI) and FIWARE. The three solutions are different from the conception perspective, as they use a SOAP (SOS-SES-WNS) and a REST architecture respectively (SensorThingsAPI and FIWARE). [123, 211] analyse the performance of each solution (SOAP vs. REST). Standardized SOAP is language, platform, and transport independent, where REST requires use of HTTP. As SmartBear, the author of [211], puts it, SOAP is *preferable* in distributed enterprise environments, as REST assumes direct point-to-point communication. However, REST is easier to operate with a smaller learning curve. In addition REST is closer to other web technologies in design philosophy and is more efficient with regards to message size (SOAP uses XML for all messages, REST mostly uses smaller message formats like JSON).

OGC Sensor Web Enablement

As discussed before, the OGC is a recognized authority in the standardization of the geospatial information field. This has included an active role in standardizing sensor data inputs with both spatial and temporal characteristics. Historically such data originates from sensor like devices, or *things* that sense. This is why these standards' names also reflect that. The Sensor Web Enablement (SWE) allows for different kinds of sensors and their data to be communicated, compared and combined. It is a suite of standards developed and maintained by the OGC, [36]. The trigger event for the development of the SWE was the advent of the Sensor Model Language (SensorML), an XML based standard that allows for the storage of geometric, dynamic, and radiometric properties of dynamic remote sensors, [37].

Within the SWE, a conceptual schema encoding for observations (and for features involved in sampling when making observations) was developed, [58]. Called the Observations&Measurements (O&M) data model, it represents the corner stone behind the web services used for this thesis' experiments as it helps define the markup of sensor measurements results. An observation consists of information about the observed geographic feature, the time of observation, the sensor, the observed phenomenon, and the observation's actual result.

To further facilitate modeling sensor data flow as well as proper storage and use of the generated data, the SensorML was developed. It is currently in version 2.1, [38]. It enables interoperability at the syntactic level and at the semantic level by building upon ontologies and semantic mediation. The modeling language allows the description of sensor data flow (during and post measurement) in a robust and semantically-tied way. It aims at completely automated machine processing in complex workflows.

The three musketeers: SOS-SES-WNS

The Sensor Observation Service (SOS) is one of the standards developed by the OGC within the SWE. In simple terms it offers a standardized web service for transmission of sensor description and its associated data. It provides an interface to sensors, their measurements and associated metadata and supports on site, mobile, remote or networks of static sensors, [242]. Observations are available with an associated computational representation of observed features. The standard version 2.0 was used during the experiments for this thesis, which employs the SensorML standard for meta description sensor encoding and the Observation and Measurement standard (O&M Model) to encrypt data gathered by sensors. The implementation used in the experiments led by this thesis' author is open source and provided by 52°North, an Open Source Foundation enterprise based in Münster, Germany. The implementation used in this thesis' experiments, 52°N Sensor Observation Service 3.5, is developed in Java and it is of Server and Client type, [242], [140].

The Sensor Event Service (SES) is a web service for standardized monitoring of sensors and the triggering of notifications. The software implementation was also developed by the same entity as with SOS, the 52°North, [115]. SES is a conceptual enhancement of the Sensor Alert Service (SAS) - a previous SWE web service specification used by clients (information/data consumers) for subscribing to self defined alert conditions. The basic SES produces notifications and provides methods to subscribe for notifications and retrieve the latest notification. It offers the functionality of registering new sensors dynamically and enables the creation of new notifications. For the experiments developed during this thesis it permitted the monitoring of near real-time sensor data and the measurement of continuous phenomenona. It allows a publish/subscribe-based access to sensor data and measurements, being a standard event notification system, [140].

The Web Notification Service (WNS) is a web service specifically designed for standardized sending of notifications via different protocols (e.g. SMS, e-mail). It acts as a transport *transducer* by changing the protocol between incoming and outgoing messages. It was designed for use with SES and allows the users to receive an alert on their device based on this pre-determined alarm indicators. WNS handles the message and its content as a black box. The WNS does not have any knowledge about the message content. Combined together, the SOS, SES and WNS provide complete functions to manage sensor data, its flow, and event management, [36].

The SensorThingsAPI

The SensorThingsAPI was designed for *things* that sense and equips their users with an open, geospatial-enabled and unified way to interconnect, thus constituting a standardized interface to the Internet of Things (IoT), [139]. It is an alternative solution to the previously presented SOS-SES-WNS, and can be regarded as a lightweight SWE profile. The OGC SensorThings API architecture can be largely separated in two parts, the *Sensing* part and the *Tasking* part. For this thesis, the first was used as it provides a standard way to manage and retrieve observations and metadata from heterogeneous IoT sensor systems. The implementation used in the experiments was developed in Java EE by Fraunhofer IOSB, and is called the SensorThingsServer and is of type Server.

FIWARE

FIWARE is a systematized framework, similar in scope to the SWE, provided however from a different perspective, that of the larger IT domain, and not specifically from the geospatial community. It does however offer integrated support for geospatial data attributes. It is being developed by the FIWARE Foundation, also a non-profit organization that promotes open standards across domains such as Smart Cities, Smart Energy, Smart AgriFood and Smart Industry. FIWARE provides an open alternative to existing proprietary internet platforms and it's software ecosystem is successfully constituting a reference option for developing digital twins, [56].

It's development has received substantial funding and support from the European Comission (EC) and has been supported by a large number of academic and private entities, [201]. The project was started in 2011 and today provides cloud hosting services based on OpenStack technology. It has a set of open standards APIs that offer a number of added-value functions "as a service", the Generic Enablers (GE), [75]. These mainly facilitate the connection to the Internet of Things, the processing of data and media in *near* real-time, and at large scale, the performing of Big Data analysis. These APIs also incorporate advanced features for user interaction making FIWARE an open alternative to existing proprietary internet platforms.

Important assets provided by using FIWARE are the abilities to manage spatial data and multiple sensor networks (a specific use case that usually gives developers and providers issues with handling different protocols and data structures), [74]. Specifically for spatial data it includes functions for handling and querying (of interest for this thesis' use cases). Context information (data that characterizes the state of that real-world object at a given moment in time) is allocated in the form of attributes associated with an instance of an object. One of the GE REST APIs, the Context Broker, is used for that end. It can handle high data volumes, a concept known as Big Data (or massive data), via a Hadoop framework. Similarly to the previously presented SWE WNS alert manager, FIWARE contains a Complex Event Processing (CEP) GE.

Technical summary

A short summary of the capabilities of the two sensor frameworks with the three available implementations included in this thesis's experiments is presented in table 4.7. The characteristics presented in the table purposefully do not include any performance metrics, such as request response time or CPU use as the applications in which the software solutions were used are different as per the intended use/characteristics. Ultimately, a choice of the *best* solution can only be done when

TABLE 4.7: Sensors/*things* web service summary

Aspect	SOS-SES-WNS	SensorThings API	FIWARE
Binding	SOAP, KVP	REST	REST
Encoding	XML	JSON	JSON
Sensor operations support	Full	Minimal	Full
Data handling complexity	High	Low	High
Data handling efficiency	Low	High	High
Critical rules for event handling	Supported	-	Supported

considering specifications of the intended use case. To propose a *blanket* recommendation for all applications would not be target-oriented. This is why a table summary brings more value for future readers of this manuscript.

Chapter 5

Energy relevant coupling of sensors and 3D models

Within the energy chain one aspect is most important with regards to energy spatial data, that of phenomena with both spatial and thematic varying properties. Long before the term Internet of Things (IoT) or its associated data revolution, the energy domain was producing large data sets of *spaghetti* like data sets, or better described as continuous data bundles. Energy domain experts handling this data have focused their attention to the thematic content and its functionality (energy relevant) and have had less of an interest in the spatial character of these data sets. This can be seen with the specifications of the most used electrical networks domain standard, the Common Information Model (CIM), also a semantical standard, [59], developed using UML, [236]. In general, network standards have been thematic oriented and sensor data repositories had no specific association to location information.

5.1 The impetus

Most literature in this domain refers to measurement data sources as *things*. As discussed in chapter 2, one of the ongoing trends in cities is the IoT revolution. [92] find that *things* have a constantly changing definition as technology has evolved. Regardless of the changing nature, the main goal, that of making a computer sense information without the aid of human intervention, has been ever present. What is pushing for a evolution of current systems is the fact that the multitude of interconnected objects that harvest information increases steadily every year. At the same time these objects interact with the real world, and use existing internet standards to provide services for information transfer, analytics, applications, and communications.

In a general fashion, spatial attributes can be looked at as only but another facet of data, another property of an entity/object. However, the combined use of spatial-temporal data in analysis in the energy chain has historically been less often done, due to limitations in sensor connectivity and spatial/thematic data segregation in data repositories. Scientists and practitioners alike are looking for ways to integrate the two components in efficient ways that allow for value creation. However, only a fraction of potential value creation has been accomplished, according to the report [4], within EU wide state institutions (10-20%) and even retail-entities (30-40%) there is much potential.

Nevertheless, the spatial characteristics of energy data gather more attention from data owners in the energy chain as use cases develop, where location know-how brings added value, see table 4.1. These requests for connectivity are now coming from a variety of actors including architects, urban planners, facility managers, spatial scientists and policy makers. Often, different actors interact within the data collection, sharing and consumption process, as they act both as a collector/intermediate data publisher but also as a data consumer (for analysis and decision support) which in turn produces further data. This emerging trend gives access to spatial data and byproducts to an increasing number of parties and thus enhances reproducibility, collaboration and decision making. As such, considering the definitions presented in 2, energy data with spatial attributes plays a vital role in smart city energy applications.

As with any nascent field no specific domain can claim ownership till expertise and workflows are developed. That meant that both the W3C and the OGC have looked for solutions for the problem of sharing large bundles of data that also holds spatial characteristics. One of the main sources for this heightened interest is the current Internet of Things (IoT) revolution, already mentioned in the ongoing trends with next generation cities in 2. IoT requires functional interoperability with other products, [226], and is interesting for this thesis as large parts of the data produced by this domain are energy relevant (e.g. accurate occupancy for buildings, detailed weather information, advanced forecasting for the aforementioned themes, energy use values, air pollution measurements). To achieve this capability, platform independent APIs for application developers are needed, and a means for different platforms to discover how to inter-operate with one another, [226]. The specific needs of this thesis requires the present work to focus on standardized frameworks that allow for measured data (from sensors) and its spatial information to be stored together. The suitable identified candidates are FIWARE (The Future Internet platform) developed as a by-product of research paid for and organized by the European Commission, and the SWE (Sensor Web Enablement) of the OGC with the SensorThings and SOS implementations.

5.2 Experiments

Within the scope of this thesis was managing the flow of energy related measurements and incorporating them within the digital representation of a city. This is why solutions were explored that allow *things* to be spatially located and for their observations to receive location attributes. Such capabilities help enable distributed energy production and a quantification of the environmental influences on that production. Understanding where a measurement is made can make a world of difference. To provide a simple example, temperatures measured in the shade or in the sun provide very different results even if the location is only but a few meters apart.

In order to abide by the good practices (presented in section 5.4) that this thesis seeks to respect experiments were carried out as to the suitability of the previously presented web services solutions. The focus lies with the distribution and availability of open standardized spatial and thematic data with public and private authorities. A fully functional workflow, that facilitates the flow of information with regards to UBEM models, in a solution that is purely based on web services was tested. The implemented workflows aimed at paving the way for a seamless integration of spatial CityGML and sensor data. On behalf of sensor data, SOS, SensorThingsAPI and FIWARE successfully linked to spatial data and allowed for the linking of these different data sources, while for CityGML four solutions are presented.

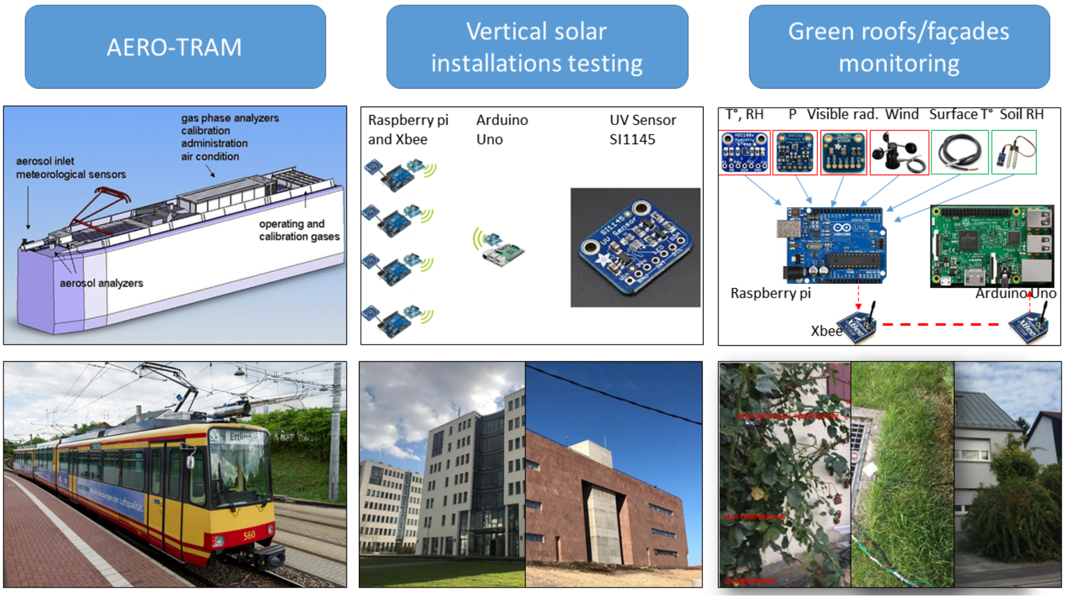


FIGURE 5.1: Sensor installations used, left hand side presents images from [95], down-middle EIFER and ULPGC building test site, down-right from [224] (left and right pictures: green façades on vineyard house in Karlsruhe, Germany, an multi-family house in Ettlingen, Germany, middle depicts a green roof installation in Illingen, Germany)

5.2.1 Sensor installations

Three use case experiments were performed to test the proposed solutions with the objective being to integrate near-real live sensor data and 3D standardized data in the format of CityGML. Figure 5.1 presents the general sensor configuration and images from the installation sites/tram. All three configurations follow the principles of open infrastructure presented in this chapter, with the software solutions belonging to different open source licence models. The upper pictures depict the sensor installations while the ones below the location of the installed sensors.

AERO-TRAM

In the first sensor use case, data is presented in the form of air quality data measurements obtained from a KIT-IMK (Karlsruhe Institut für Technologie - Institut für Meteorologie und Klimaforschung) measurement campaign (belonging to the AERO-TRAM project). The project was sponsored by the environmental ministry of the German state of Baden-Württemberg and aimed to be an extensively automated, long-term study of concentration, spatial distribution of parameters, particle size distribution and total particle number in the area of Karlsruhe. The sensor installation is a high quality mobile installation that was situated on top of trains (S-Bahn trains in German, a train that can travel both on high voltage and low voltage powered lines allowing it to connect rural and urban settlements on similar gauge tram and train tracks: "*from the hinterland through the city center back into the hinterland*") belonging to the local transportation authority (Verkehrsbetriebe Karlsruhe, VBK).

This mobile laboratory "AERO-TRAM" for measurements of urban ground level pollutants could perform fully automated measurements of the gaseous components:

$H2O(g)$, O3, NO, NOx, CO and CO2, in addition to weather indicators like air temperature, RH (relative humidity), P (air pressure) and horizontal wind. Its location was measured using an integrated antenna for Global Positioning System (GPS) measurements, [95, 242]. Data for both of the monitored train lines (S1/S11 and S2) was offered in support of this research work for the purpose of sensor infrastructure experiments. Part of the train track runs in the Rhine valley, while a second part continues up into the Black Forest (German: *Schwarzwald*). Measurements were made at short time intervals as the train moved along its track, making the sensed data a perfect candidate for our study with changing spatial, temporal and thematic characteristics. The measurement campaign occurred prior to this thesis, and there was no possibility to test live data feed from the tram onto the software platforms proposed into this thesis. Data was provided in R format and then converted and stored into a local database from where it was further used as sensor data for this thesis's infrastructure experiments.

Solar installations testing for vertical façades

The second use case includes data from a UV measurement campaign on a static location in two buildings. One was located on the university grounds of Las Palmas de Gran Canaria (ULPGC), , the most populous city in the autonomous community of the Canary Islands, Spain. The second monitored facility was located in Karlsruhe, Germany, the second-largest city in the state of Baden-Württemberg, in southwest Germany, near the French-German border. Both buildings host tertiary activities, namely research and development. One sensor installation was mounted on each building facade (four in total for each location). The objective of the sensors installation was to provide data support for modeling activities. The model's purpose was to determine suitability of commercial solar energy installations on vertical façades in two very different locations, central European, with continental climate and southwest European/northwest African, with sub-tropical climate. Both implementations have been realised under similar conditions and used the same equipment.

The sensor installation was designed and built in-house from separate purchased parts. The sensors were directly situated on a window of each facade of each building, eight for the entire experiment. For radiation measurements the model SI1145 Digital UV Index/IR/Visible Light Sensor, from SiLabs was used. This equipment can measure infrared as well as visible light and using a calibrated light sensing algorithm it can provide the UV Index, which provided the input to further data modeling in this experiment. The sensor was attached to a Arduino Uno Board REV3 with a UartSBee V5 - Xbee Adapter attached to it. It picks up the measurements of the sensor and sends it to a centralising repository by using the Bluetooth capabilities of the Xbee adapter. The central repository unit receiving these data was a Globmall ABOX Raspberry Pi 3 Model. As such, all data produced was able to flow in *live* fashion to the data repositories and make use of a standardized web services protocol.

Green roofs/façades monitoring

The third use case was a sensor installation developed to quantify the impact of green roofs and green façades on a building's energy demand. The sensors were installed in three buildings located in the region of Baden-Württemberg, in southwest Germany. One building included a green roof installation and the other two had partial green façades on the side of the buildings. The objective of the installations was to measure the impact of green roofs and façades on the building heating and cooling demand. Parallel to the sensor installation a model was developed that can provide estimates of the impact of such external building extensions on the heat transfer based on the heat balance principle of foliage, soil, and structural layers, [224]. The measurement campaign was performed during the summer and validated the model with regards to the estimates of impact on cooling.

The sensor installation was developed starting from senseBox, an educational sensor toolkit, built on top of an Arduino, [173]. It provides (in this experiment) well documented, portable interactive sensors for reproducible research based on open source software. The data produced was released as open data on the platform's interactive website. It included multiple off the shelf additions and performed the following measurements: air temperature and humidity (sensor HDC100x), air pressure (BMP200), visible radiation (TSL45315), wind speed (Aenometer) and concrete and substrate temperature (DS18B20). In two of the locations internet was available and the data was able to flow live to the central data repository, located at the KIT, while electricity was provided from large batteries. In one location (Karlsruhe vineyard house) no electricity or live data feed was possible, as such data was stored on a memory card and then transferred to the repository post measurement campaign.

TABLE 5.1: Sensor installations summary

Aspect	AERO-TRAM	Vertical solar installations testing	Green roofs/façades monitoring
Location	Region of Karlsruhe, Germany	Las Palmas de Gran Canaria, Spain and Karlsruhe, Germany	Karlsruhe, Illingen and Ettlingen, Germany
Sensor installation	High accuracy scientific installation	Built from scrap	Built on-top of educational version of senseBox
Location	Mobile	Static	Static
Objective	Measurements of urban ground level pollutants	Validate model for solar panel installations on vertical facades	Monitor impact of green roofs and façades on building heating and cooling demand
Measurements	Location, H2O(g), O3, NO, NOx, CO, CO2, T°, RH, P and Horizontal wind speed	Infrared, Visible light and UV Index	T°, RH, P and Visible radiation, Wind speed, Concrete and substrate T°
Live data feed	No (measurements delivered *as is*)	Yes (two sites)	Yes (two sites) No (Karlsruhe site)

A summary of the use cases and associated measurements and technologies is presented in table 5.1. This table indicates the location, the objectives of each experiment and the measurements that were taken in each location and whether or not live data feed was achieved within the trial.

5.2.2 Measurement data stream

With regards to sensor data flow, both the OGC (with SOS, SES, WNS, SensorThings API and the WNS) and the FIWARE perspectives were implemented and tested. The place to start was determined by analyzing the standards delivered by the OGC, as this organization delivers products aimed at both spatial data and their thematic components.

AERO-TRAM In the case of the AERO-TRAM data the implementation occurred long after the measurement campaign, as such no live data feed could be organized. However, the data was stored in a 52°North SOS database and later was made accessible via a RESTful extension. The specifications of the sensors were described using SensorML and stored within the SOS Service. The measurements were first converted from R format into a CSV file and then further stored using the O&M model inside an XML which was then finally imported in the SOS database using the SOS Service. The 52° North SOS adheres to the OGC SOS standard and has online support available, [242].

With regards to the complete SOS-SES-WNS profile, after the SOS database is populated an SOS-SES feeder picks up the data and feeds it into the SES. From here it is fed into an application layer by use of a SOAP Subscribe/Notification mechanism. An SES-WNS translator, SOAP to XML, allows for the WNS service to receive notifications.

In the present experiment context, the extraction of sensor data to the web interface was performed using a WFS showing static data present in the database. For both OGC SWE options, a useful solution with regards to visualization of results is presented in [50] with the 52° North Timeseries API.

Solar installations testing for vertical façades In this use case there is a need to separate in between the two constellations used in the Karlsruhe experiment and the one in Las Palmas. Even though both locations present the same use case and both achieved live feeds to the repositories, the frameworks deployed were different. In Karlsruhe, the experiment used the SensorThings API while the one in Las Palmas used FIWARE.

For Karlsruhe, the sensor data flowed from the Arduino Uno via the Internet to a server at the IPF, where an interface would allow connections to the SensorThingsServer. This had the sensor specifications previously stored on it using SensorML. The server would receive data via the Subscribe/Publish mechanism and store it accordingly on the SensorThingsAPI database. Access to the data would be allowed via the Request/Response mechanism.

In Las Palmas, the FIWARE framework worked exclusively to store the data coming from the sensors. The workflow includes two software components that were not discussed before, ORION and Cygnus. ORION connects to the sensor and broadcasts data in a standardized format, while Cygnus manages subscriptions to ORION sensor data broadcasts and finally writes the measurements to a database.

Green roofs/façades monitoring

For the third use case, certain distinction must be made in between the three use cases. In both Illingen (green façade) and Ettlingen (green roof), the SensorThings API was used, while in the Karlsruhe vineyard house use case a lack of access to internet hampered any possibility of a live feed. For the locations with feasible data

TABLE 5.2: Comparison of the spatial workflow technologies

Method	Developing entity	Advantages	Disadvantages
3DCityDB WFS*	3DCityDB consortium	Readily available with associated DB solution	Only selection function, No data management options
Standard WFS and GeoServer*	Boundless Spatial, GeoSolutions, Refractions Research	Most used server for geospatial data	Does not cover the Generic Object class, No support for complex data types, complex geometry, No grouping of features within feature collections (hierarchies)
Cesium JS interface*	Analytical Graphics and Cesium Inc.	Simple ETL processes make large 3D non-standardized data sets available	
REST based interface**	[186]	Acquisition of CityGML data based on semantics, Queries can be made, using topological and semantic rules, Easily include standard extensions (ADEs)	Focus only on the building class

* in-house implementation and testing, use case AERO-TRAM data and 3D city model of Karlsruhe, Germany
**alternative found in literature

stream the chosen protocol, although the simplest of the three proposed solutions, offers sufficient capabilities for this use case: data gathering from sensors, well established software solutions, with associated documentation and standardized data storage.

The sensor data flowed from the sensors attached to a Arduino Uno via two Bluetooth adapters to a Raspbery Pi which then sent the data further to a central data repository located on an EIFER PostgreSQL server using the SensorThings API database structure. With regards to the location without internet access (Karlsruhe vineyard house use case), the data had to be transferred manually, every second week during the measurement campaign as the power source (battery) was depleted.

Figure 5.2 encapsulates the concepts presented above. An important mention is that a real live implementation of the SOS-SES-WNS would not use the R format file as an in-between the sensor and SOS. The R format was chosen only due to availability on the side of the data hosting entity (KIT-IMK) and is valid in the presented experiment as the measurement campaign occurred prior to this thesis.

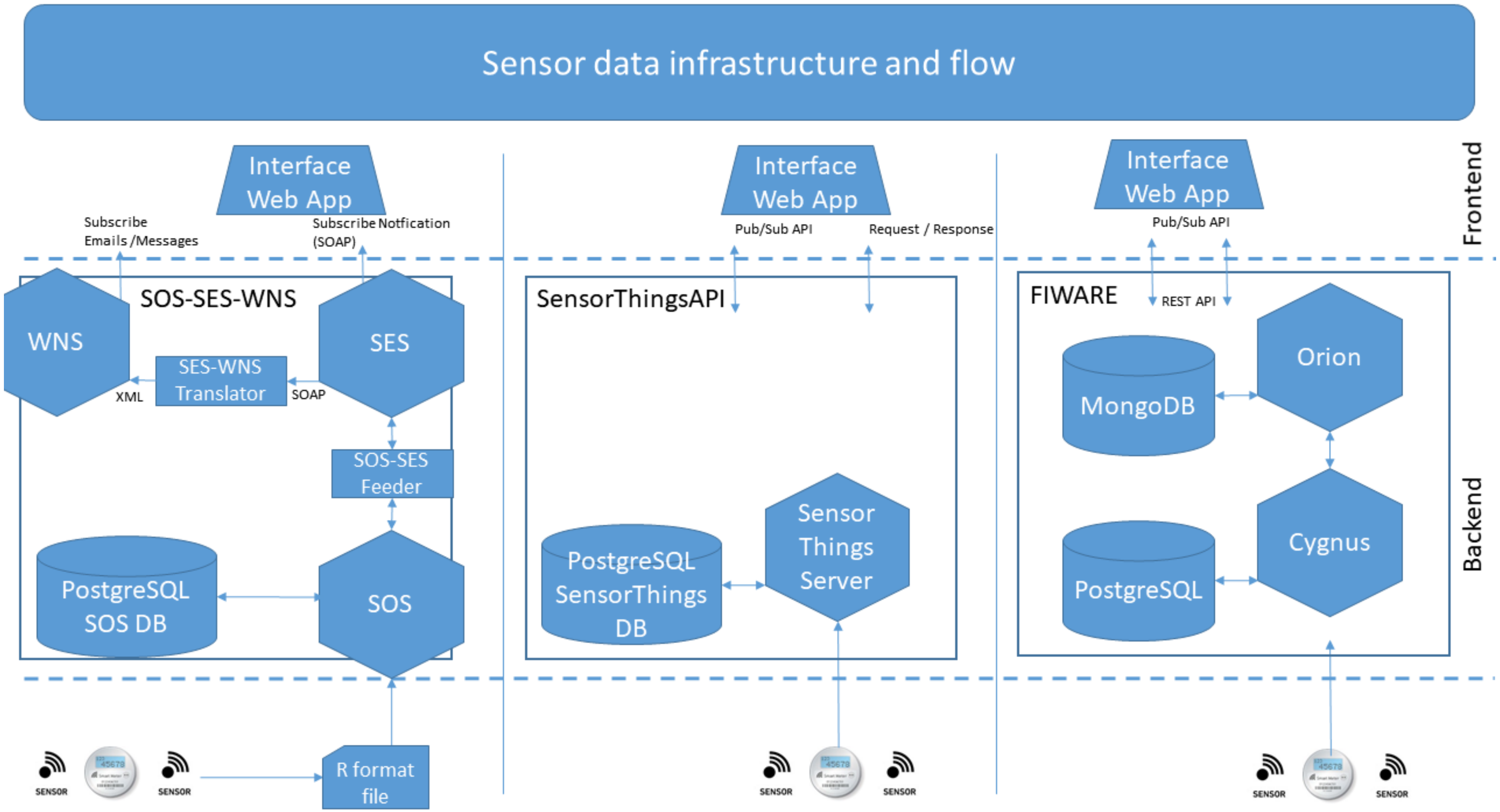


FIGURE 5.2: Sensor data infrastructure, left SOS-SES-WNS, used with AERO-TRAM, middle SensorThingsAPI used both in the Vertical solar installations testing and the Green roofs/façades monitoring, right FIWARE used in Vertical solar installations testing

5.2.3 3D spatial data flow via web services

Spatial models

Spatial data in 2D that is relevant for the energy chain can be shared via web services (such as the WMS, WCS or the WFS) in a standardized manner. These are presented in their dedicated section, 4.4.1. The data can be sent for further use once ETL operations are performed and the quality of data has been checked. This subsection narrows its focus on the flow of spatial 3D city models. In this thesis' case, these were always stored in CityGML, either as files or in a database version of the standard, the 3DCityDB. This data originated from open data or it was built from open data and is in general restricted to building models.

Three methods to produce CityGML data are presented in the section 4.3.5. These methods were used directly in this thesis to produce CityGML files. Four additional methods to obtain CityGML from open data are discussed in [231]. These are (1) using building footprints and remote sensing data, (2) converting Open Street Map (OSM) building data, (3) using LIDAR data sets and open data for building footprints and (4) using geoportals, in case a city has one available, see figure A.7 for a detailed presentation of each workflow. When CityGML samples could be directly obtained, extensive testing had to be performed in order to ensure data quality. A sample was provided by the public authorities of Karlsruhe, while for the building in Spain it was built from openly available data.

In the case of CityGML data transmission via web services, different working solutions were attempted. Figure 5.3 presents the different implemented solutions. Two of these were purely CityGML serving oriented (one used the 3DCityDB WFS, and another one used the GeoServer), while a third looked at a workaround for providing non-standardized spatial data, that can mimic CityGML data characteristics (and can later be used to generate 3D city models). All approaches aimed at testing the viability of such a solution in a completely open-source environment. From the previously presented web services, the WFS (Web Feature Service) is found by [186] to be the most suitable OGC standard for supplying real geometry data. Table 5.2 presents the implemented solutions with their associated advantages and disadvantages.

3DCityDB WFS

As CityGML is much more than geometry, semantics and topology are big part of the reason for which the standard has seen implementation and adoption by public authorities. The first implementation of a useful WFS service for CityGML came with [122], where the 3DCityDB library contained an OGC compliant WFS implementation working with the 3DCityDB schema. This implementation is referred to as the 3D City Database WFS. With its open source software version, simple online access to the 3D city entities is possible directly to the database, [163]. With the current version (at the time of this manuscript writing v. 4.1) the following operations are supported: GetCapabilities (metadata information regarding the server providing the WFS), DescribeFeatureType (returns a schema of the CityGML features that are invocable via the WFS), ListStoredQuerries (provides a list of the cached queries), DescribeStoreQuerry (makes available detailed information regarding the aforementioned queries) and GetFeature (returns a selection of CityGML features from the 3DCityDB using a query). A commercial solution, the Virtual City Systems Web Feature Service (VCS WFS), proposes a more complete solution where CityGML objects

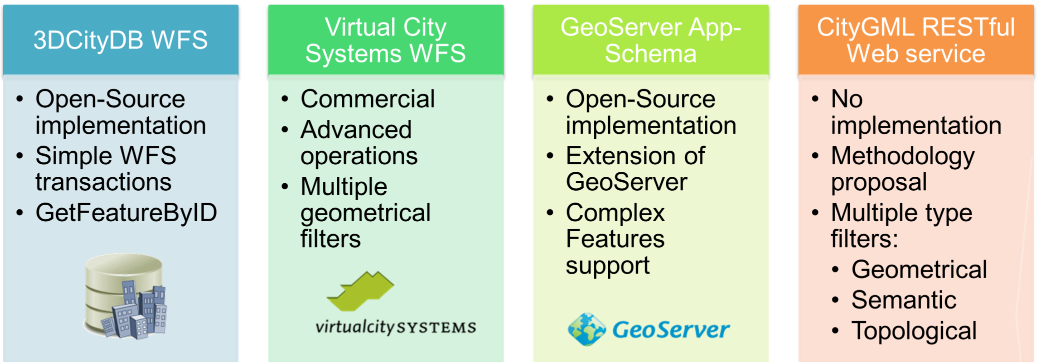


FIGURE 5.3: Four suitable approaches for CityGML web service, after [242]

can not only be queried from the database, as in the previously mentioned open-source version, but can be filtered (spatially or thematically), as well as inserted, edited and deleted, [215]. This first implemented solution provides only partial resolution to the research question and only focuses on the retrieval of CityGML data.

GeoServer App Schema

The second technical solution used, GeoServer, is an open source server, known for its purposely interoperable design, [48], [131]. It was investigated if serving CityGML via a WFS with advanced functionality is possible, with complete results presented in [242]. This would allow for data exchanges based on specific feature types and complex queries to be performed. The implemented solution used GeoServer (as it also respects the OGC WFS standard and provides all tools: discovery, query, locking, transaction and stored query operations). It was found that the tested solution can cover almost all CityGML classes with the exception of the Generic Object one. For modeling the urban energy chain this represents a significant downside as the Generic Object can handle all objects not described by the CityGML standard (in this thesis's specific case, energy relevant data, such as building characteristics). These issues in between CityGML 2.0 and GML encoding rules of GML refer to the fact that in the GeoServer Application Schema, all public GML application schema used must respect the GML "Striping" encoding rule (a complex type cannot be the direct property of another complex type). The generic.xsd schema does not obey the GML encoding rules. In addition CityGML models often use complex data types for properties and group features within feature collections (this generates very useful multi-level hierarchical database structures, see the proposed identifier structure in chapter 3). Another possible limitation stems from the types of geometry used with CityGML which are more complex than the limited geometry types supported in relational databases. As such this second tested solution is also incomplete.

A practical workaround

The third approach stems from a recognition of the data availability reality. This means that most often available data is made available in SHP file format, it is not standardized and incomplete, a fact debated broadly in chapter 3. Therefore, a solution was conceived to work within this reality: SHP file data that was subjected to

ETL operations and then stored in local repositories. From there it was either further shared via standard WFS, converted to CityGML with the use of Python scripts or converted to glTF (graphics language Transmission Format) using the CesiumJS application. Inside the shared data, an extra attribute was offered, that contains the number of floors for each building. Using this attribute, the interface mimics a CityGML LOD1 model in which an extrusion of the building is created. This is useful as it generally replicates the volumes of the buildings in the city without the shape of the roofs. In the background this data can be used to create 3D models and perform detailed spatial analysis relevant for UBEM, such as shading analysis.

CityGML RESTful web service

Following an attempt to identify other available solutions, similar in purpose, in literature, one such example was identified in the work of [186]. They compare SOAP and REST solutions and propose another methodological approach based on REST-style architecture. This allows for the acquisition of CityGML data based on semantics instead of following on the same track as the standard OGC WFS, which is designed to retrieve, visualize and modify data based on geometry characteristics. The authors of the aforementioned study focus their efforts on the building class, as it is the most commonly used with CityGML. They also provide a successful path for integration into a virtual data hub for all the other CityGML classes. In their proposed solution one can also query data beyond geometry, using topological and semantic rules and most importantly the standard extensions, namely the Energy ADE and the Utility Network ADE. This solution does fully cover the issue at hand, and offers significant advantages from a perspective of energy relevant data transfer. However, it is for the moment only a theoretical implementation. It does not include data management such as update or edit abilities.

Technical summary

TABLE 5.3: Web services used in BIM, CAD and GIS for geospatial data relevant in UBEM

Service standard	Current version	Content	Domain	Type
Web Map Service (WMS)	1.3	Serving georeferenced map images	GIS	Open
Web Coverage Service (WCS)	2.1	Defines web-based retrieval of coverages (digital geospatial information representing space/time-varying phenomena)	GIS	Open
Web Feature Service (WFS)	2.0	Serving georeferenced map images	GIS	Open
WFS for 3DCityDB	3.1	Querying and modifying CityGML objects	GIS	Proprietary
Web Service for CAD (WSC)	0.1/?	Serving georeferenced map images	CAD	Proprietary
BCF-API	2.1	Supports BCF issues exchanges between software applications	BIM	Open

Table 5.3 presents web services encountered in the energy chain while testing hypotheses included in this manuscript. These web services, either used in applications developed and tested by the author, either serving as input, were integrated in an infrastructure that was purposely designed to have a GIS centric perspective. This happens as most data delivered per web service in this domain tends to be stored in repositories that have GIS capacities. In addition, the funnel system developed to load data in the central repository of this thesis mainly imports 3D GIS data, which means that data is converted or obtained a priori to feed it into the system. Several possible parallel approaches are presented in the tight vs. loose coupling section, 5.5.

5.3 Experiments related summary and discussion

The present thesis was, in the knowledge of the author, the first trial that brings together CityGML, SOS, SensorThingsAPI and FIWARE. A rundown of the three sensor installations that delivered data within the thesis is presented in 5.1. It can be observed that the three different constellations allow for multiple configurations to be tested out with regards to location and quality of sensor installation. The measurements are different in scope with diverse end objectives. This creates a comprehensive perspective of the capabilities and implementations encountered. Further, live data feed was achieved, as much as the technical and logistical situation could allow for, in the second and third implementation that the author was completely in control of directly.

As a general rule, choosing the right framework to work with when dealing with sensors and city spatial models is always dependent on the specifications of the use case. In the OGC case, the main take away is that the SensorThings API and the SOS-SES-WNS pack differ significantly. The first follows specific design decisions that support connecting *things* with modest resources to the web following mainly REST principles (an efficient JSON encoding, the use of MQTT protocol, the flexible OASIS OData protocol and URL conventions). As such, it is easier to implement and use, but lacks significant abilities when compared to the three pack solution, which, if implementation difficulties occur, can be reduced depending on requirements. Within these experiments the SOS-SES-WNS was found to be the most difficult to implement, however one of the most potent solutions. SES supports critical rules that filter data and help derive new information, while WNS provides a variety of ways to help transfer data via notifications. Only FIWARE offers similar capabilities, and goes beyond as it involves in a normal fashion noSQL databases. With regards to documentation it was found that FIWARE offers the best solutions, while in terms of simplicity the SensorThings API wins, however *basic* its services may be.

Specifically to FIWARE, the ORION-Cygnus constellation provides useful capabilities for establishing live data streams links to CityGML objects. This solution can link objects described in the Energy ADE and the UtilityNetwork ADE with sensor data measured on site. It would facilitate monitoring situations (visualization purposes and event triggers). What makes it also very potent is its installation onto a Docker system, and the fact that the data is stored into a MongoDB, a database system specifically designed to handle large data sets, a so called NoSQL (not a relational) database. In this manner, FIWARE allows for big data friendly architecture to be used in combination with the standard RDBMS (Relational DataBase Management System). This data pursues the following flow: sensor installation, Orion, Cygnus, NoSQL DB. The author finds that the FIWARE solution provides a significant benefit for smart applications due to the very large capabilities bundled together. It is a conclusion also shared in the study of Salhofer, who finds that this statement holds true also for their tested use cases, [201].

What this manuscript's reader should take away is that a proper use case definition before any implementation is done is paramount. This will facilitate the choice with regards to the proper implementation based on the table 5.1.

For the data modeling of the energy chain, one of the main issue is related to the custom devices from the energy and utility domain. These experiments show that they can be modeled alternatively either using CityGML extensions (e.g. AbstractDevice inside the UtilityNetwork ADE) or even inside FIWARE (with the DeviceModel, its schema allows for the creation of custom devices that monitor ambient conditions or state of components within a system). Within SensorThings API, the data model contains a table entitled *FeatureOfInterest*. This entity was used to describe the building that housed the sensors. It was then further linked by way of an identifier to the CityGML object. This information can be then inherited to the CityGML object that describes the building.

The way in which data connectivity (in terms of the same object belonging to multiple data sources) was always assured in our implementations was via the unique identifier method presented in chapter 3.6, that then allows for the same object to exist in two systems and to simply be referenced from one system to another by means of XREF links. CityGML, FIWARE, SensorThings API and the SOS data model specifically allow for this to exist. Similar work in terms of scope to the one presented in this chapter was submitted by Chiang and his colleagues, [52]. This work approaches the integration of *things* and semantical city models by using the OGC SensorThings API and OGC CityGML via Semantic Web Technology. The authors present an integration ontology to connect data from these two standards that includes different perspectives and definitions of a *Thing*. To connect information from the CityGML and SensorThingsAPI, SPARQL queries were used. It represents another step forward and potentially a useful lead for developers implementing smart city apps.

In general, this proposed sensor - urban energy data integration by way of spatial links can facilitate the integration of IoT and/or energy data with 3D city models to achieve truly interoperable smart cities. However, this work represents only a step towards this goal. It can be regarded as a facilitator, if the distributed data/loose coupling solution is adopted, as presented in the dedicated section on tight vs. loose data coupling, 5.5.

The presented work includes no performance metrics, a potential drawback, even though in the experiment phase performance and load tests were conducted. These metrics cannot be used for comparison, as the services compared operated in different use cases. Only SensorThings API and FIWARE could be directly compared in the second tested constellation (Karlsruhe with the first and Las Palmas with the second), however the two services belong to different *weight* categories when it comes to the capabilities and support infrastructure and software. This is why the only technical comparison that is included in this manuscript is to be found in table 4.7. It should however suffice in helping practitioners orient themselves towards a preferred choice.

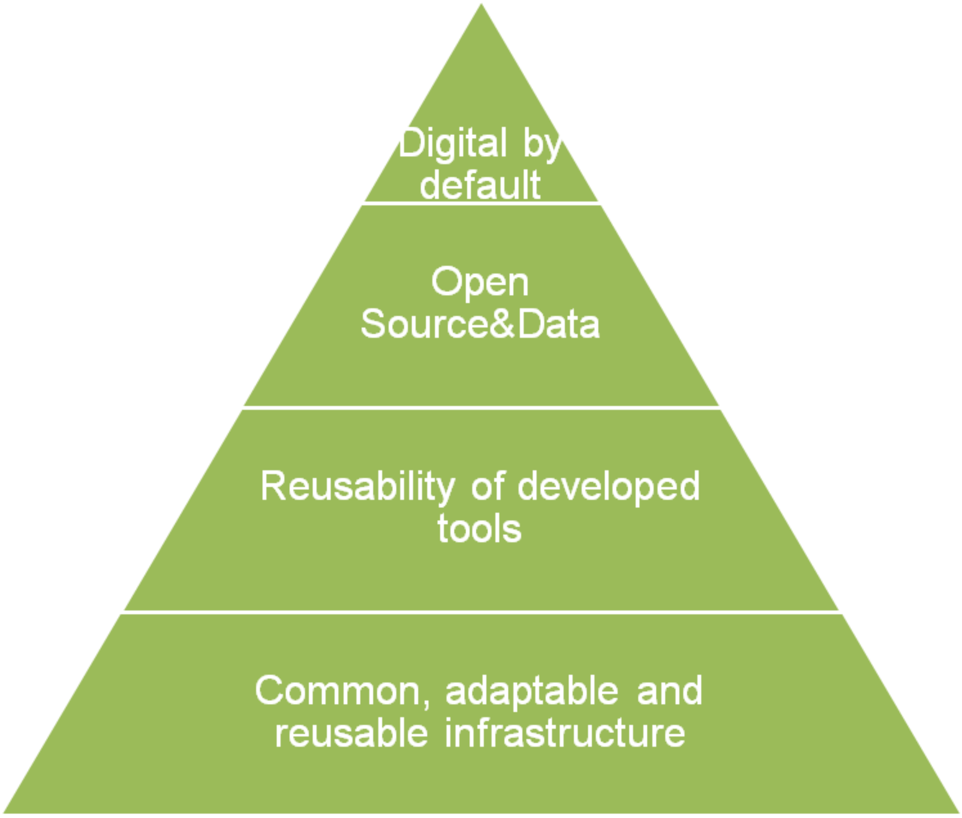


FIGURE 5.4: Good digital practices in modeling, simulating and monitoring the energy chain

As can be seen in the presented work GIS has a slight lead on BIM when it comes to standardized data sharing and availability. This is true when it comes to commercial use, web services and data availability. As such, it supported the idea that the main data infrastructure used in the applications of this thesis follows a CityGML centric approach. This hypothesis was put to the test in the thesis life. A similar proposed approach that stores all data in GIS and focuses on converting BIM to GIS was proposed and tested by [237]. The authors partake in a GIS focused experiment with limited conversions of BIM data and conclude that for meaningful city wide impact to occur communities wide planning units and city scales with building units are required. In many ways this work draws parallels to the one presented in this chapter.

5.4 Good practices and a check-list for next generation cities

In principle all smart city efforts, as described in 2, involve a higher degree of integration of existing infrastructure and data. This sub-chapter seeks to formulate requirements associated with this integration, by defining principles for next generation cities that facilitate integrated management of the energy chain, a sort of check-list that next generation cities can abide by in order for their energy chain to be properly managed by making use of spatial and thematic data. In figure 5.4 a check list is presented. The implementations and experiments that were made during the life of this thesis have indicated that these four points are facilitators towards integrated energy chain management and planning. The concepts are further detailed and explained in the following sub-sections, these are:

- Common, adaptable and reusable SDI,
- Reusability of developed tools,
- Open Data and Open Source code,
- Digital by default.

In this direction, a study of literature found the effort of [44], that assesses a multitude of independent smart cities initiatives working at different spatial scales. The authors conclude that in order to turn cities into smart cities, initiatives need a systemic and integrated approach, one that makes the best use of infrastructures and encourages interoperability and scalability of solutions. Seeking resolution to a similar issue, but aiming in scope only at the scientific domain, the FAIR initiative has taken shape, [235]. The authors define guidelines to improve digital assets' by improving the following characteristics:

- Findability,
- Accessibility,
- Interoperability,
- Reusability.

Wilkinson and his colleagues emphasise minimal to none human direct intervention or supervision of machine-actionability (i.e., the capacity of computational systems to find, access, interoperate, and reuse data). In fact, they go as far as to propose a reversal of historical roles, where humans accept the new reality, one in which society increasingly relies on computational support to deal with spaghetti like data (increased in volume, complexity, and its creation speed). This is however, less possible in the urban energy chain, outside of the academic context, especially as the given context includes many less flexible legacy systems. However, the list of proposals stemming from this author have certain similarities, as the SDI is what allows for accessibility and interoperability. Findability should be ensured by the urban ontology related identification system proposed in chapter 3, section 3.6.

Another important pillar in helping define such work principles for managing, modeling and simulating spatial-temporal data in the city wide energy chain are the United Nation sustainability development goals. Cities and states exist in very different *realities*, a fact debated upon in chapter 2, with significant geographical, legislation or cultural differences. The 17 goals that the UN has defined are valid for all humankind, as such, relevant also in our use case. Goals 6 to 9 (clean water and sanitation, affordable and clean energy, decent work and economic growth, industry, innovation and infrastructure) and 11 to 13 (sustainable cities and communities, responsible consumption and production, climate action) are relevant to this thesis. In a manual on sustainable development, [198], Sachs credits studies that confirm that across the world people feel happier and more satisfied with life when they trust their government. He continues to define good governance as something that not only applies to the public sector but also to the private sector, which is also valid in the energy chain use case, as public and private actors actions intertwine and depend on one another.

Another interesting find during the literature research was a state wide initiative, found to be contextually beneficial, in scope, to this work. It was found in the Baltic nation of Estonia. In order to provide a solution for a functional data sharing infrastructure the Baltic state of Estonia has built a framework with virtual government services called e-Estonia, [153]. Their service is built on block chain technology in order to uniquely identify and secure transactions and offers citizens and companies more than 4,000 virtual services. In fact 99% of their public services are accessible online via a one-time login gateway, [174]. All spatial data that are owned by the Estonian state, local governments and other legal persons governed by public law are published and made available in a single point of access, [145]. This reliable and secure way of exchanging information, based on transparency between citizens, companies and the government, seems to have fostered faith between those actors, with most IT framework projects started by the government in the late 2000's (the geospatial portal and the medical DB for example in 2008) and paying off in public confidence. According to the 2015 spring Euro-barometer, [180], 51 percent of Estonians trust their state, compared to the average of 29 percent of all Europeans. This is related to the so-called snowball effect, as once the standardized web services were put in motion and people started using them, there was an ever increasing dependency on these services and interest in their further development.

The previous literature references focus on public actors. A balancing of this perspective is offered by [197]. The author presents solutions for companies seeking to benefit from the ongoing digital transformation. They recommend that that digital conscious companies should take part in standardization efforts as this is a mutually beneficial endeavour. Participating, brings in needed know-how and models the further development of standards with the firms involvement.

In the energy chain specific spatial extent, all actors need to operate according to the law, with accountability and transparency guiding political and business practices. This should go further, using foresight to plan adequately to developing trends, such as the current energy market transformation. At the same time stakeholders need to be responsive to the needs of other stakeholders, and be proactive in critical issues such as land use, pollution, and the fairness and honesty of political and business practices.

5.4.1 Common, adaptable and reusable SDI

The four pillars, presented in figure 5.4 include as foundation step the common, adaptable and reusable SDI. Similar to a building, spatial data infrastructure is the corner stone of any attempt of a city to successfully push forward with planning and adapting it's infrastructure to the requirements of next generation cities, specifically with regards to the energy chain. As all topographers and building site managers know, a cornerstone is the first stone set within a construction of a masonry foundation. It is then used as a reference point in determining the position of the entire structure. This is also similar in impact to the SDI on the energy chain. There are numerous use cases (presented in detail in section 4.1), that are positively impacted when associated with a competent use of SDI.

A common SDI, that is shared by multiple groups prevents deprecated one of a kind solutions to continue to function. These legacy software are common in the energy and utilities domain due to the lack of standardized connectivity to other domain. These can be seen playing a detrimental role in [75] reducing the ability of the local utility company to upgrade it's software to emerging use cases. A common solution allows all actors in the energy chain to either store data on the public SDI or to link their own SDI to the publicly available one via web services. Further facilitating this links would be the unique identifiers proposed in chapter 3 and the use of spatial entity marking and scales (presented in the same chapter mentioned before).

With an adaptable (sometimes referred to as a generic) SDI a city can able to be modify it for a new use or purpose. Generally, an adaptive system is able to respond to changes in use or changes in the interacting parts, [5]. Software specific in a generic system is their ability to support a variety of architectural models and adaptation mechanisms. One of the features that commonly appears in such systems is a feedback loop. It ensures that the system can deal with the current requirements. In addition such systems often present hierarchical structure.

Lastly, the reusable aspect, a key note in software architecture, as [157] testifies. Software reusability is their ability to be reused without requiring further development and should not be confused with the adaptable characteristic mentioned above. In the case of a city SDI, whatever development is done of it, it should be made with the direct intent of it being used more than once, not a one off solution that potentially wastes the organizations resources.

Even though it would be a major advantage, using only open source SDI is not a condition per se. It would however facilitate integration with other tools and actors. What is required of the SDI, is that it uses open web services, can work (import/export) with open standards and has non proprietary entry/exit ways and data storage capabilities. Architectural choices are further discussed in their dedicated section 5.5.

5.4.2 Reusability of developed tools

In the IT (Information Technology) domain, reusability of software is the general norm in product development. This is the second pillar in the present recommendations. It relates to the use and re-use of existing products of the software development life cycle (software components, test suites, designs and documentation) in one form or another within the software product development process.

Considerable resources are invested by all the actors involved in our spatial context, be they public or private, to develop or purchase software tools that control, manage or facilitate use of the assets they control in the energy chain. This tools need be redeployed without large cost and/or time impacts. At the same time this tools need to work from one city to the next, so that software companies can re-deploy them.

Within the energy chain, multiple tools are developed to tackle specific use cases. A first step in ensuring reusability of a tool is the connectivity to an SDI that respects the first pillar. Dual flow should be also part of this criteria. It means that Get capabilities should be complemented by Insert into the same SDI so that the data is not hosted on the side of the software. This is a major impediment with legacy software, as [240] find.

Development of energy models to assist in the management of the energy chain is highly dependent on the implementation of such good practices. This is shown in the publication of the Open Energy Modelling Initiative, [185]. Enforcing reusability reduces duplication of work and makes high-quality data and planning tools accessible to all actors with and without the funds for commercial options.

5.4.3 Open source and open data for open governance

As cities are the main contributors of greenhouse gas emissions, [212], Annex 63, an effort of the International Energy Agency looked at the steps needed to tap into the great potential of energy efficiency measures and emission reduction in this spatial context. Their conclusion is that the effort revolves around implementation and usage of technologies, but also at their adaptation to the local context. Among the requirements, processes and frameworks identified, access to open data from the energy domain to efficiently support the modeling and simulation of energy scenarios in communities has a crucial impact on the success rate of such initiatives, an idea also supported in [189]. This large community effort supports the idea that open data has an impact in local success of methods used and recommended in other spatial contexts.

Another benefit of sharing open energy data is found in the study of [12]. It showed that sending energy consumption data to an end-users, as well as the data of their neighbors was a low-cost and durable way to reduce demand. The sharing of information created awareness of the consumer's own behaviour. It also increased intra-community communication on energy use patterns and successful local energy consumption mitigation measures.

Of course, as with other actions and phenomena, open data initiatives have multiple facets to a potentially positive action. For example, such applications raise the question of privacy, and whether or not such practices are desired. [154] investigated and evaluated this aspect, by balancing privacy concerns with potential benefits of exploiting smart meter data. The manuscript presents the principle concerns from end-users regarding smart meters, including the risk of illegal uses (e.g. burglars knowing when households are unoccupied), commercial uses (e.g. targeted advertisement) or legal and personal surveillance. The counter-balance to these concerns is presented by the use of this data in modeling, simulation and operation in distributed system operation and planning that ultimately helps with demand reduction. The authors of the aforementioned study argue in favor of privacy-friendly measures that ensure beneficial use of open data without infringing on sensitive personal information. Other solutions proposed argue for the spatial and/or temporal aggregation of such data or for the sharing and displaying of data locally, where such data is generated and finally their aggregation into energy use profiles, [121].

Practice however, shows that data owners are reluctant to share data and have a certain *inertia* about their current *modus operandi* way of doing things. This is to be expected with public entities and companies that have often existed in one form or another for a significant time duration. Nevertheless, entities seeking value creation in this environment are finding ways to do just that. One simple example are hackathon events - sprint like events. Sprint events, used in the IT domain with AGILE development methodologies have become the mainstay of software development. Hackathon events that gather significant participation lead to value generation with small data samples from otherwise inaccessible data sources, [218].

With regards to spatial data on energy infrastructure, open data initiatives have largely concentrated on public static data. Energy network data sets are starting to become freely available, for example being integrated in the OSM data sets, but also due to research projects, see [187] for examples. Many different public and private actors are making ever larger open repositories available by the use of web portals and spatial data infrastructure catalogues. Even if these data are present in one source, they still very often entail very different characteristics, with a lack of standardization presenting a large impediment. It creates the need for extensive ETL processes, as presented in chapter, 3.

Linked to the open data initiatives by *ideology* open-source software (OSS) is defined by [132] as a type of computer software whose source code is released under a license in which the copyright holder grants users the rights to study, change, and distribute the software to anyone and for any purpose. The success and impact of this software was being predicted as far back as 2002, [35] and [137] when already OSS packages were becoming available outside of the scientific community.

There are large variations of the licence models, however, as our spatial extent includes various actors with different interests. It is not the purpose of this thesis to evaluate which model suits best what situation but rather to express the fact that the entire process modeling and simulation of the energy chain can be performed with OSS, a position also confirmed by [86] with regards to open source energy system optimization tools. This approach enhances interoperability but does not exclude connections to other commercial tools due to the open specifications of data input/output mechanisms, a fact which is especially important as legacy software exists in many of the entities from the energy chain.

In general, open source licenses can be categorised in two categories: Copyleft and Permissive Licenses with the main differences occurring in compliance requirements and the license of products developed on top of the current products. Permissive licences promote re-use, including further assimilation into commercial products, without requiring improvements to the code to be released at all. A copy of the license text and the original copyright notice must be included in the newly licensed code. Copyleft licenses have the same requirement as the permissive licences, with the addition that any further development of the code must be made available under the same license as the original. Table 5.4 presents the most common licences for both data and code.

TABLE 5.4: Common licence models for code and data, after [185] and own work

Domain	Usage	Licenses
Software	Copy left	GNU GPL, LGPL, AGPL, Eclipse, Mozilla
Software	Permissive	BSD, MIT, Apache, European Union Public License, ISC
Data	Copy left	Creative Commons BY-SA, Open Data Commons ODbL
Data	Permissive	Creative Commons BY, CC0

In addition to file based standards, open (as in publicly available standards) web services are required to standardize the input and output of information from data owners or data managers to companies and citizens that can produce value out of it. Because ultimately data is only valuable when it is used.

During the modeling, simulation and monitoring of different use cases in the energy chain, this manuscript's author aimed to use open source software and open data throughout the experiments. Figure 5.5 presents at least one open source solution for each of the domains that intertwine in the energy chain: spatial analysis, modeling,

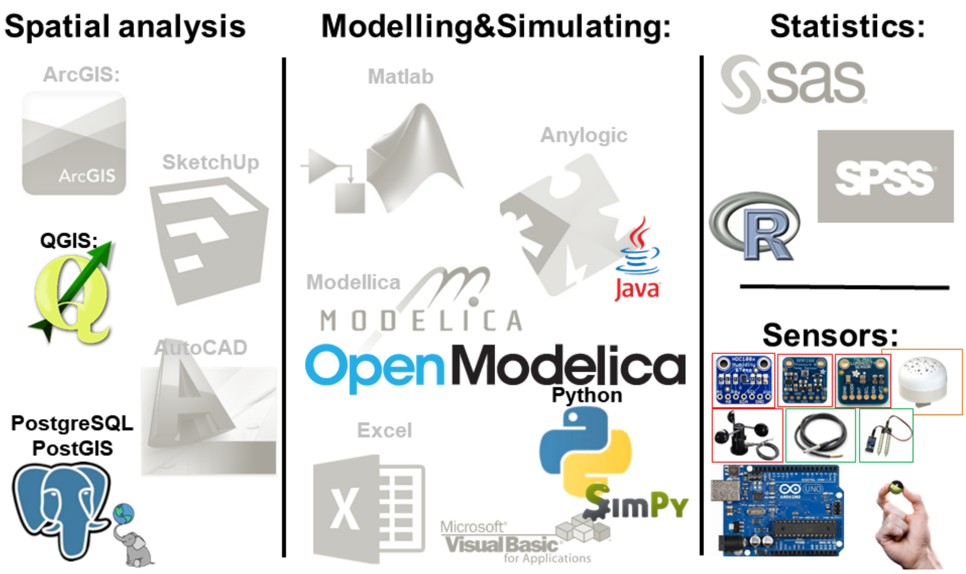


FIGURE 5.5: Open source solutions for spatial analysis, modeling, simulation, monitoring and statistical analysis of the energy chain, depicted with colors are all open source solutions, with grey tones commercial solutions in use encountered during this thesis, after [167]

simulation, monitoring and statistical analysis. This point is also made in [167] for the same given context. QGIS together with PostgreSQL and PostGIS were used for spatial analysis and data storage, Python for modeling, C++ for simulation, R for statistical testing and analysis and monitoring was performed with open hardware solutions built from scratch or on top of senseBox, an educational sensor toolkit, built on top of an Arduino (with an open source operating system).

The use of open standards and the publication of contiguous open data sets are prerequisites that, in addition to building public trust in authorities, increase services quality and create added value for cities and all other energy chain stakeholders.

5.4.4 Digital by default

The words “digitization”, “digitalization” and "digital transformation are often used interchangeably, as the work of [197] shows. There exist however important conceptual differences in between the three terms. Digitization regards the conversion of certain technologies, e.g. digitizing of paper maps, where digitalization is about making use of the newly digitized products and develop new procedures, business models, or commercial offerings. Digitalization provides context to digitization and explains the significance of a technology and its relevancy to the organization. Digital transformation is a “process that aims to improve an entity by triggering significant changes to its properties through combinations of information, computing, communication, and connectivity technologies". In figure A.1 the concepts and hierarchical relationship behind the three terms is described. It clearly depicts digital transformation encompassing all digitalization efforts, while the latter further incorporates all digitization works.

With regards to the energy chain specificity, an additional requirement should be set in for all new projects, digital by default, with all utility networks, energy meters, energy sources. Even though it may come as a surprise, one of the great issues with managing and modeling the city wide energy chain is the fact that very often, data is still present only in paper form or basic digital sharing formats, such as PDFs. As a consequence, there are many great efforts to digitize (creating a digital copy of) still existing analog content, [147, 83, 204]. However, digitization *per se*, is only a step towards the conversion of data and it requires to be done from a perspective of digitalization, where the digital products are to be integrated with other ones. This is where the previously mentioned open standards come to assist, a so called digitizing towards open standards. This would create the premises for a successful digital transformation of the energy chain.

Digital by default should ensure that the ETL operations discussed in Chapter 3 are greatly reduced. Even though the pyramid presented in figure 5.4 has *Digital by default* as a smallest step, it is the spearhead that can further all the other three pillars of good digital practices.

5.5 On data fusion - to merge or not to merge

Finally, with regards to the data of next-generation cities and their energy chain, this subsection of the thesis discusses data fusion, a concept required for bringing large data sets together (energy related and otherwise). These data are relevant for modeling, simulation and monitoring activities. The relentless fight for consistent *proper* data often has modelers integrating multiple data sources in their workflow. The research question on integration of this data should be taken to see how and how far this should be made. In the scientific domain it is referred to as data fusion. It involves vast allocation of resources that are needed to bring existing SDIs in line with the new requirement. Often, the spatial actors involved are reluctant in sharing their data or to host it onto third parties and legislation is required in order to move things over (*convince* them that such action is *beneficial*).

In the general IT domain data fusion brings together distinct data sources. The end purpose can be application dependent, however it is generally recognized as a procedure that produces consistent, accurate, and useful information than that provided by any individual data source, [117]. Similarly, data integration involves combining data residing in different sources and providing users with a unified view of these data, [53]. For experts working within the geospatial digital domains data integration and fusion are tantamount concepts and most applications combine spatial data sets with different thematic content into a single (fused) repository containing merged properties of the original sets.

5.5.1 Tight and loose coupling

An attempt was made to conceptualize the concepts surrounding data integration and standardization presented in this chapter and to separate those in tight and loose coupling. Often, the natural tendency of data owners is to accumulate data under one roof and data structure, something that is particularly true for the energy chain with the diverse actors and spatial scales described in the previous chapter. This *modus operandi* is generally referred to as tight coupling. Data is brought together from different repositories and connected on the side of the host, as can be seen in figure 5.6.

As with many other architectural choices it has advantages and disadvantages. The main advantage is the fact that the data is usually associated with a relational data

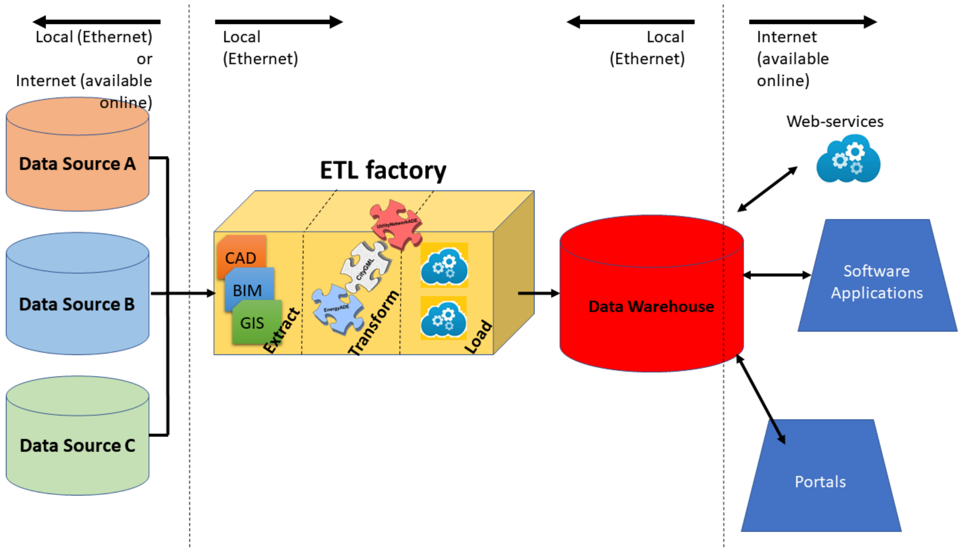


FIGURE 5.6: Tight coupling general architecture - theoretical implementation

structure. It facilitates the linking of all the previously mentioned categories of data and can be easily linked to standardized data structures. The disadvantage is that the relational data structure is more often than not lacking standardization. In addition, hosting many different types of data together in a single repository forces the data owner to use a *one size fits all* approach in terms of the databases and software. This bundles in together sensor data, geospatial data and different energy chain thematic data with various degrees of success.

An alternative to this approach is loose coupling. This method does not centrally store data. It fact it does the exact opposite, and that is to keep it within different repositories that have different data owners. A good practice within this approach is that data is often treated is that from the very beginning and classified, for example into databases with nonSQL or SQL characteristics. Figure 5.7 presents the architectural concepts behind loose coupling.

However, even though separate, this data remains accessible to the outside world via a virtual data hub. This hub ensures standardization of data available to the spatial actors. The owners of the actual data repositories data usually have different perspectives, energy thematic, spatial thematic or urban thematic. For this second method to be efficient and handled with ease, standardization implementation needs to be accepted on a wide scale.

As with the first method there are advantages and disadvantages. The main advantage is that data is handled by their thematic experts, who are well aware of existing standards in their field. Also, these repositories tend to be well maintained as they are the bread basked of many of these spatial actors. A disadvantage is the lack of control on the side of the virtual hub owner as to the content of data. This is however dependent on weather the access granted to the repositories is bi-directional, meaning Get and Write, or unidirectional, meaning Get Capabilities only. Figure 5.7 presents the architectural concept of loose coupling.

The study [197] finds this second perspective useful and states that data ownership *per se* should not be a priority, but rather data accessibility. Ensuring that information

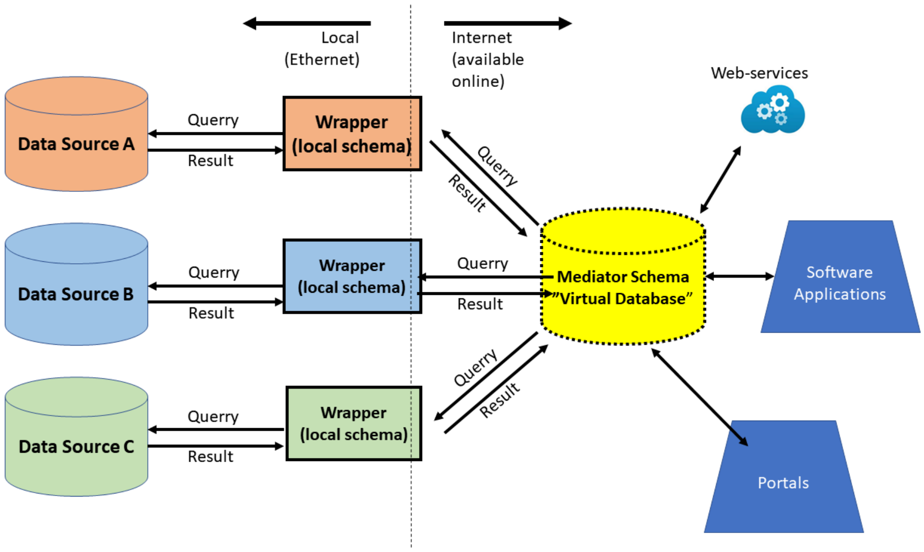


FIGURE 5.7: Loose coupling general architecture - theoretical implementation

is accessible to multiple actors, enforces successful management of the energy chain and often provides increased quality of energy services.

In between the two architectures there are many shades of grey, that is to underline the fact that one solution could implement technologies present in both frameworks. For example, at times, loose coupling solutions employ multiple databases storage, separating in between nonSQL and SQL structure where required, while still maintaining data on the premises of a single spatial actor. At others there may be parts of data that are allowed to stay with separate actors (due to legal requirements and/or technical implementation advantages) while other parts of the data are brought in together.

5.5.2 Energy chain architectures - discussion

With regards to the energy chain, figure 5.9 presents an implementation that makes use of tight coupling. Spatial data, building energy data, utility network and monitoring data is all bundled up in a single file. This is of CityGML type and exploits the extension mechanism presented in the dedicated ADE section of this manuscript, 4.3.3. Processes for ETL data handling have to be deployed throughout the workflow. The processes on the left hand side are not standardized and have to be re-adapted to each situation and data source as the input will differ and data quality is a major issue in geospatial data fusion. Once a standardized file has been produced the production pipeline can include standardized ETLs and provide data to another user, API, portal or interface.

A loose coupling implementation working for the entire energy chain is depicted in figure 5.10. In addition to spatial data, building energy data, utility network data and monitoring data, the proposed flow allows for integration of live data feeds. Together with the web services presented in their dedicated sub-chapter, 4.4 this architecture also allows for live feed data to be integrated. This solution is highly flexible, allowing for fast re-utilization and makes proper use of dedicated databases for specific data types (SQL and NoSQL).

Data	Spatial	Temporal	Solutions	
Producers	Yes	Yes	CityGML + EnergyADE (XML or DB)	
Consumers	Yes	Yes	CityGML + EnergyADE (XML or DB)	
Networks	Yes	Yes	CityGML + Utility Network ADE (XML or DB)	
Things data (Sensors e.g. Smart meters, Air quality)	Yes	Yes	Energy ADE*, Dynamizer ADE	SOS / SensorThings API / FIWARE

SENSOR SENSOR

*limited to energy related sensor data

FIGURE 5.8: Energy chain components and storage

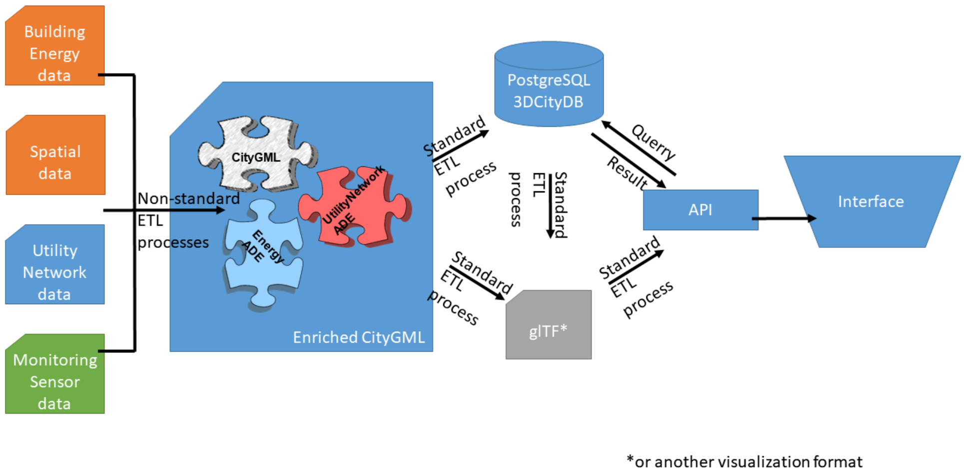


FIGURE 5.9: Tight coupling general architecture - practical implementation

Both of these solutions were implemented when dealing with the experiments presented in this manuscript, 5.2, and the UBEM modeling activities, 6. The loose coupling presents the most advantages as it shortens ETL processing times, and always provides access to the latest data available. However, the tight coupling implementation also provides a working solution, which can, in certain cases be the only viable way to follow.

The loose architecture also embraces REST principles, allowing for interaction with Uniform Resource Identifier (URI) available resources (in practical terms, over the internet). In addition, the loose coupling architecture creates the premises for good practices, as they include error handling in the inner APIs and prevent error affected data of spreading in the data produced further down the pipeline. Adding new components is a fairly straight forward procedure once the workflow has stabilized, making the architecture scalable and flexible. This is for example how the infrastructure evolved to include monitoring data, from a first use case oriented only towards UBEM, and then towards *Things* and energy monitoring for model validation. These statements are also in line with the findings of [159], in a specific energy chain use case, and [138] in a more general perspective.

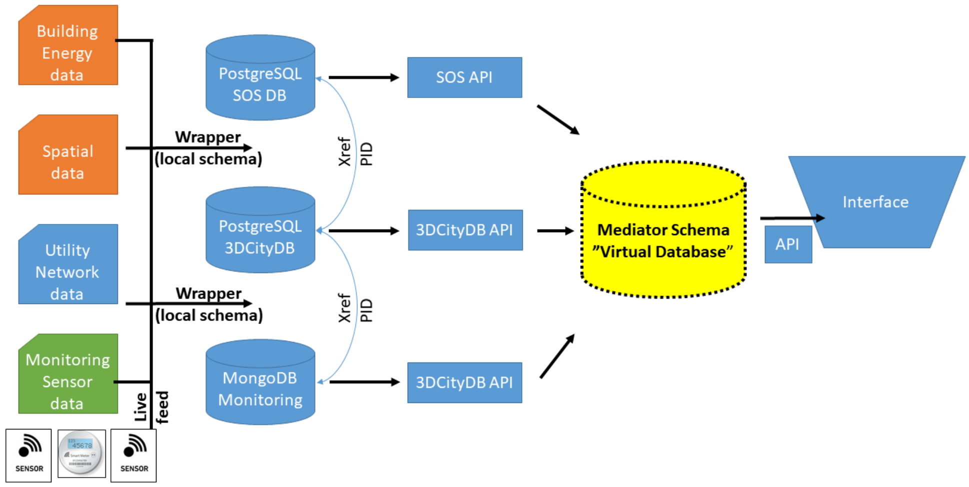


FIGURE 5.10: Loose coupling general architecture - practical implementation

To underline the fact that one most not simply choose among the two architecture views, a mixed file/web service solution based on CityGML and sensor web services is presented in figure 5.8. Similar situations are presented very often to practitioners, where a purely loose solution is not feasible due to certain restrictions regarding the data of one energy chain or another. Energy producer data is to be stored in CityGML and the EnergyADE. These producers can be large entities, but can also be a single family residence which owns a solar panel and is selling electricity in the network. Consumers can be represented using the same solution. Networks require the second extension, the Utility Network ADE. Lastly the monitoring data can be stored using the EnergyADE, the DynamizerADE or another database solution via the sensor specific web services mentioned in the dedicated section, 4.4. All types of data can be either stored as a file or in a database depending on the use case.

Finally, a choice in between the two architecture types should be made once use cases are defined, user categories have been properly identified and user stories have been outlined, following the AGILE development logic. Further, all potential data sources need to be identified, with entry/exit ports definitions and legal requirements established (data and code licenses).

Chapter 6

aEneAs, the scale aware urban energy modelling library

The odyssey (journey) refers to the implementation of a library of models used to model different components of the urban energy chain. Named after Aeneas, a mythical hero, it provides support to the claim that semantic city models can act as data hubs for this kind of modeling. It mainly uses buildings as data units and allows and can be classified in terms of scope as part of the UBEM field that uses a mixture of spatial-physical models to model the real urban world. The library was developed either directly by this author or under his supervision during his employment at EIFER within the scope of this PhD as well as master and bachelor thesis student overseer.

This thesis scope lies within the spatial and temporal side of the urban energy modeling field. The following chapter focuses on the actual task of modeling the urban energy chain, as previously defined. This effort is ever more significant today in the context of high energy prices and a conscious audience with regards to human induced climate change in search of mitigation solutions. This public interest also advances political and managerial decisions in both state and private entities that favor using both human and climate friendly approaches in urban planning and management with different means to this ends. To facilitate such calculations energy indicators are used as prime KPIs in decision making. This KPIs are fed directly from results provided by UBEM modeling and simulations However important the task may be, converting different stakeholders political and public interest in successful policy integration in our own houses, districts and cities is no trivial task and this is what urban energy domain modeling is here to facilitate.

6.1 Desideratum - why UBEM

Energy modeling is the computer simulation of a spatial object/entity used to determine or estimate building energy usage, [43]. Modelling and simulation are the most common tools used to assess the technological and policy impacts of smart solutions mentioned in 2, as well as to plan the best ways of shifting towards next generation (smart) cities and classifies energy-related work on planning and operation models within the smart city into five main intervention areas: generation, storage, infrastructure, facilities, and transport.

At the interaction of the GIS, urban planning and energy modeling worlds lies a nascent field, [191], called Urban Building Energy Modeling (UBEM). It is main objective is the modeling of dozens to thousands of buildings using bottom-up approaches and is somewhat opposite to detailed individual Building Energy Models (BEM). In BEM, practitioners perform energy use analysis of the building inventory at the level of a single building in high detail. UBEM tools, standards, and paradigms have been identified in the previously cited manuscript and are often simplified BEM tools that use streamlined workflows. In addition, UBEM also provides hybrid methods, merging single building energy models and broad scale building stock modeling, that survey the energy analysis of districts and cities. In this context multiple studies show that the emerging science of UBEM (Urban Building Energy Modeling), can successfully provide support for different design or policy decisions with acceptable error ranges, [121], [164].

UBEM belongs in general to the Urban Modeling field, together with its sister field, Urban Energy System Modeling (UESM), [202]. UESM is a discipline, that focuses more on the energy solution in the energy chain. As depicted by Samadzadgegan and colleagues, UBEM provides information on the energy demand to UESM, and together provide for complete solutions in the modelling of urban energy chain.

6.1.1 UBEM classification

UBEM modeling approaches can be classified depending on the scale at which modeling is performed, [76]. These can be separate using the perspective of the building stock. Modeling at large scale bundles of buildings is referred to as top-down. At the opposite end, the bottom-up method calculates energy consumption at a single building unit.

The methods developed in UBEM help guide the energy management of multiple buildings (dozens to thousands, [191]) and avoid the classical top-bottom approach which extrapolates from the status-quo. This makes for imprecise results at building unit but alleviates the large data gathering efforts required for city or district level studies, as can be seen in [43]. The authors of the aforementioned study also identify the multiple degrees of complexity behind modelling a complete urban energy system and present a classification of the analytical, iterative, and hybrid methods used.

Energy models used in BEM and UBEM can also be classified as black and white box models, [103]. Traditionally, white box models are models that try to describe buildings as close as possible to reality. Opposite to that, black box models use historical data for modelling and are of great value when exact information on the buildings is missing. In between white and black box models there exist a large span of *shades of grey* situations when one, two, or more parameters are missing and are replaced with historical data or is generated from sources of statistical data by methods presented in chapter 3. Most, if not all, UBEM models experience *grey* situations in which the data input that models require is missing. Hong and his colleagues perform a review of UBEM tools, restricted only to those that have associated scientific publications, and go on to classify the methodological approach into:

- physics-based dynamic simulation (white box),
- reduced-order calculation method (grey box),
- data-driven method (black box).

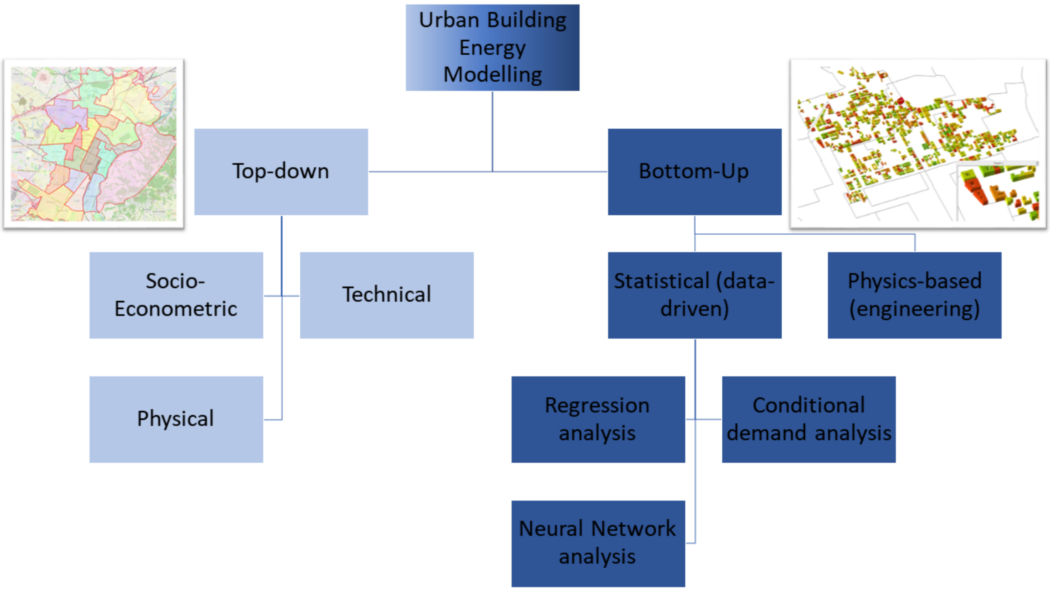


FIGURE 6.1: UBEM model classification, after [76], [191] and [103]

The first type, physics-based dynamic simulation, tries to mimic a building's energy behaviour up to the smallest possible resolution, both spatially as well as temporally, identifying all sub-components responsible for energy flow, in and out as well as within of a building. The second type of models, reduces the complexity of the first models, and thus requires less data input. They are able to offer information on building energy behaviour for large building numbers in a relatively short time span. This category is also where the library of models proposed within this chapter belong to. Lastly, data-driven methods rely on large quantities of statistical and/or measured data. These usually include standards (ISO, DIN) based classifications of buildings per different categories with associated occupancy and thermal characteristics. Sensors then provide the ability to determine building characteristics and match it to existing database types.

An integrated presentation of the two methods for classifying UBEM models is found in the study [77]. This is presented in figure 6.1. What can be observed in the figure is that ultimately both top-down and bottom-up methods can be physics-based, reduced order approaches or data drive. It all depends on the availability of data and the intended results use.

6.1.2 Variables and UBEM general workflow

With regards to the variables used by BEM and UBEM for a simulation, these parameters, as classified in [80], are largely similar in both fields. They have significant differences only with regards to the accuracy of their determination (BEM using in general more precision for its estimations) and fall in the following categories:

- Geometrical parameters,
- Building physics parameters,
- Meteorological parameters,
- Usage profiles,

- Monitoring data.

The first category refers to building geometry and its physical location in the real world, while the second one regards the physical properties of the materials that form the building, as heat capacity and thermal transmittance of exterior building elements (walls, slabs, roofs) and optical properties of windows. In the third category all weather related variables are included, outside temperature, solar radiation, wind speed, humidity, cloud coverage. The fourth category includes variables that help classify the behavior of building occupants and profiles for heating, lighting equipment and ventilation. The solutions proposed in this thesis fall in the field of UBEM, as large city-wide spatial models were produced and combined with building physics and socio-economic indicators to provide a detailed approximation of the city. Lastly, monitoring data is used for calibration and validation of results.

When the building inventory is described using the previously mentioned parameters a virtual (or digital) twin is created. This can also be referred to as a virtual prototype - a dynamic digital representation of a physical system and of the spatial entity, [146]. These twins are generally created using a software package, or a combination of multiple software.

The UBEM workflow can be summarized in five steps, graphically depicted in figure 6.2. These are as following:

(1) First, the objective definition occurs. This entails the narrowing down of large general national, regional and local policies down to local objectives. This ultimately drive the energy model and scenario definition. In general results are presented as *Business as usual*, where the *status quo ante* is quantified, and other scenario, where different implementation measures are applied and their impact is quantified, relative to the *status quo ante*.

(2) In the second step, the spatial description of city entities is created. This step involves extensive ETL processes that clean and validate data. Multiple data sources (GIS, CAD, BIM) are combined to create uniform spatial data sets that represent buildings and networks, components of the urban energy chain. Further, UBEM relevant data (energy parameters) from multiple sources that can be used for energy modeling is identified. Again, comprehensive ETL processes are employed, that finally allow this data to then be merged to the previously mentioned spatial data. These steps are presented in figure 4.9 and correspond to the section dedicated to the data funnelling system section of this manuscript, 4.3.5. In this thesis manuscript and in the presented experiments, this data is stored and linked to CityGML building objects that contain hierarchical relations based on the identification system defined in section 3.6. Once this step is complete the previously mentioned digital twin is obtained.

(3) In the following step, simulations are performed on the data, generating demand estimates under different assumptions (scenarios) and other energy use related KPIs. This step is performed with certain given assumptions over a predetermined time duration. The given graphical depiction of figure 6.2 assume that a model library is already in place that can be used with minimal effort.

(4) After the results are obtained, calibration and validation is performed, either by comparing against statistical data or by use of the sensor data, see section 5.2.3 with regards to the types of data, services and storage solutions. Finally the results are

describes using a report or publication. Often, data is made available on open data portals.

(5) These results are then used to achieve the goals of the UBEM modeling process, to manage and improve the urban energy chain. These tasks, required for urban energy modeling, are prerequisites in all energy planning and management activities. This is done in order to understand the processes and the relevant scales at which they happen, as [99] indicates. This is due to the fact that optimal energy decisions in city/district planning and design help optimize whole-system performance (city and district level).

6.2 aEneAs - Trojan hero's odyssey

The research questions posed in chapter 1 required the testing of the suitability of semantic city models to act as data hubs for modeling of the urban energy chain. During this dissertation work energy models with a focus on thermal energy demand and supporting spatial calculation tools were implemented with the end goal of providing UBEM functionalities. This sub-chapter presents tools and work, bundled up in a model library, that were developed and implemented during the thesis, and their corresponding description.

The library received the name aEneAs, which stands for urban Energy Assessment. The name of the library bears the name of the mythological founding father of Rome (and the grandfather of Romulus and Remus), Aeneas, who is a mythological hero survivor from the ancient city of Troy. According to Roman and Greek literature, Aeneas was a member of the royal line at Troy and played a prominent part in defending his city against the Greeks during the Trojan War. He brought know-how regarding city structure, buildings and institutionalized organization along with him as he made his way from Troy on to Rome and contributed to a successful and improved city.

The author's personal contribution to aEneAs is that of conceptual and code development, scientific supervision and general management of the library package. In addition to the authors work, the code development of the library was made with the support of HiWi's (Wissenschaftliche Hilfskraft or assistant researchers). They were writing their master or bachelor thesis under the author's direct supervision at EIFER. The purpose and vision of this manuscript's author was to develop aEneAs as a proof of concept for the validity of city wide energy and data modeling methods, with most tools being developed as prototypes. However, three of the components were further developed and became an integral part of software products used in production workflows.

The library is aimed at using existing energy calculation methodologies and knowledge and help replace unorganized or proprietary data structures behind energy modeling. As far as possible, this functionalities make portability and reusability from city to city possible. The components of the library presented in the thesis have all been developed at least up to Technology readiness level (TRL) 4, and a maximum of 7. Figure A.2 presents the TRL levels as described by the EU for Horizon 2020 projects, [69]. Corresponding levels marked with yellow correlate with the levels that components of the library have reached during their development.

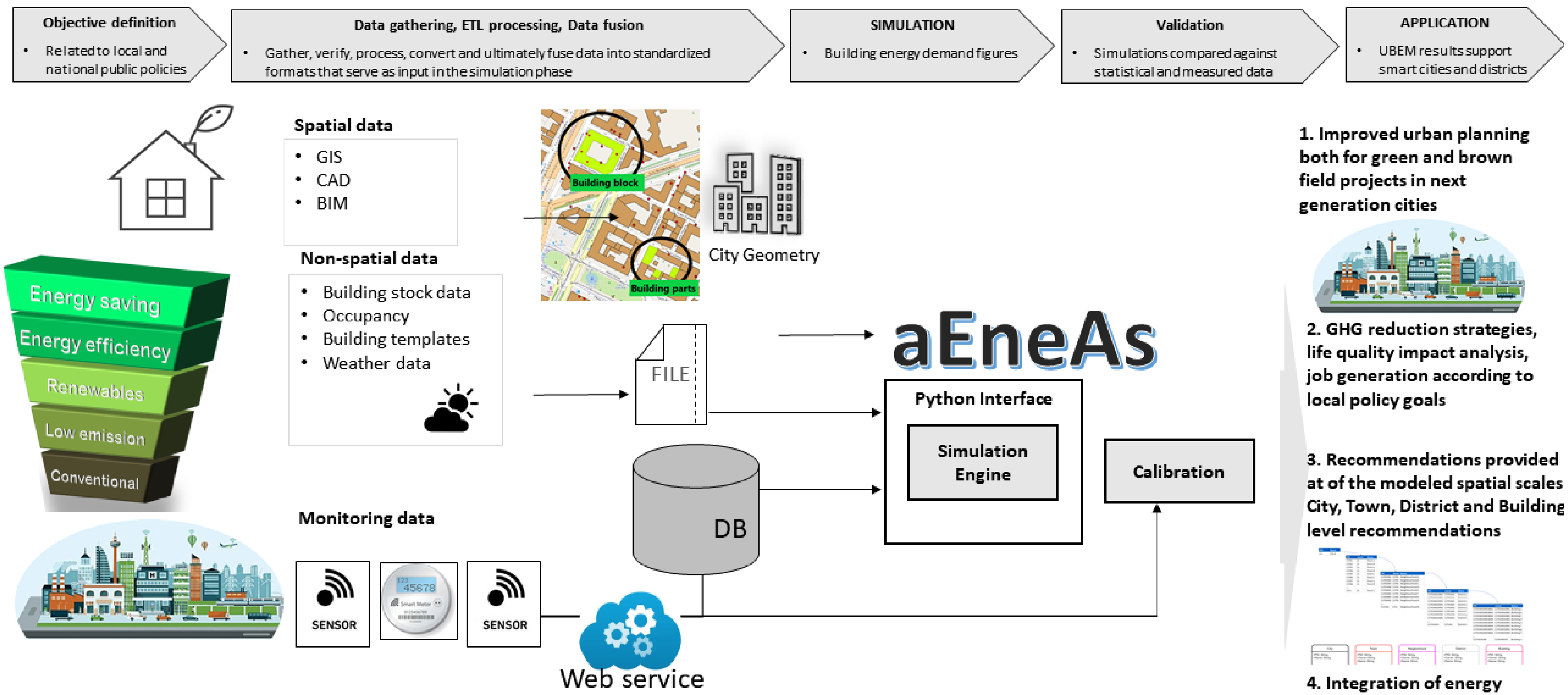


FIGURE 6.2: UBEM general steps, own work with elements from [6]

Based on the results presented in [149], it has been observed that a fifth of all UBEM publications encountered work with EnergyPlus as the simulation tool, a physics-based dynamic simulation engine. This happens even though the tool cannot make use of CityGML data models, the most often used data set in UBEM, according to [103], which means that ETL operations are involved. Chapter 4, section 4.3.5 presents and discusses data sets stemming from GIS, CAD, BIM and other relevant data sources and methods of converting these into CityGML. In addition to these methods, the ever greater availability of open data sets stored in this format made it the ideal choice to support UBEM activities. This is why a library working exclusively on such data input was proposed.

As such, it became part of the goal to show the viability of 3D semantic city models as a data hub for energy modeling. As it is standard with software development efforts, certain decisions related to the environment of development were made in the very beginning. In short, CityGML allows for the storage of standardized 3D semantic city models and has the ability to contain standardized extensions that are built on top of it. One of the key things that needs to be mentioned here is that beyond its imperfections and difficulties with implementations **there exists, still, no viable alternative open standard that competes in terms of scope with CityGML**, at least none that this author has interacted with. In addition there is a community of practitioners that stands behind it and offers support and develops open-source tools. Beyond the standard, aEneAs library makes use of the work of standardization within two consortia that the author of this thesis has been working in, one developing the Energy ADE and a second one the Utility Network ADE. Both the standards and the extension concepts are further detailed in the aforementioned chapter.

6.2.1 The backbone

The aEneAs library's first requirement was the establishment of an SDI (Spatial Data Infrastructure). This current section briefly describes the actual support infrastructure (or backbone) used for the testing of hypothesis in this thesis. For use cases, standards, and web services encountered and tested during the modeling of the energy chain please refer to chapter 4. These use cases were considered for the development of aEneAs.

In order to efficiently handle CityGML, which is written based on XML, the decision was made to use a database version of it. This XML to UML converted version, is suited to handle large data sets and allows the handling of data by using SQL queries and PostgreSQL databases. There are alternative workflows, such as XML based parsing, which other libraries have been using. However, given the general institutional context (experience with tools, institutional IT conditions, availability of support for the chosen tools and given trends), and the requirement that the development is able to run on both large cities and also on smaller districts with ease the decision was made to handle spatial data as a whole in spatial databases.

Figure 6.3 presents the support infrastructure for the aEneAs library. On the left hand side the funneling system that converts data into CityGML can be observed, this is described in greater detail in section 4.3.5. The data is then stored in a PostgreSQL DB that uses the 3DCityDB data structure. Additional data sources are represented by sensor data and other UBEM relevant data (discussed extensively in chapter 4 and 3). The same PostgreSQL DB and the 3DCityDB data structure is used for data storage and spatial operations performed in the aEneAs library.

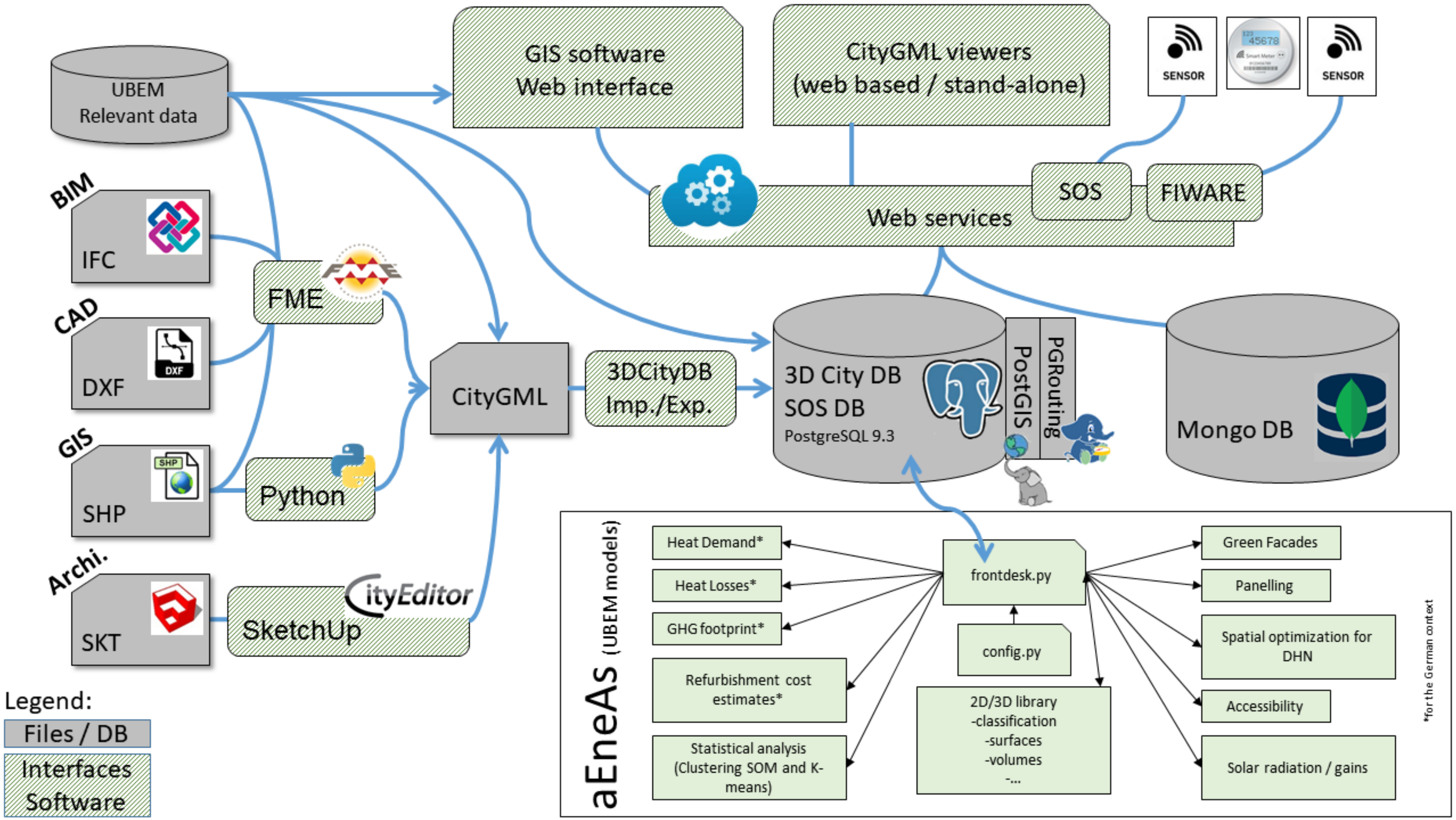


FIGURE 6.3: aEneAs and the support infrastructure

Earlier versions of this infrastructure were presented and discussed in [166] and [231]. When comparing to the previous versions of the infrastructure, the one presented in this manuscript includes additional components required for the coupling of sensors with 3D spatial models and web services for data flow. The storage of sensor data is made inside two databases, one PostgreSQL (for SOS, FIWARE and SensorThingsAPI data) and another in MongoDB (a NoSQL database specifically integrated in the workflow for use with FIWARE).

On the upper side of the same graphic, 6.3, visualization tools can be observed. These access data stored in the two DBs via web services that at times permit even data management inside the repositories. For more details on the web services please see the dedicated section in 5.2.

The following tools and extensions were used extensively in the work for aEneAs:

- PostgreSQL, an open database system,
- PostGIS, a database extension that enables spatial data storage and calculations,
- Python, a programming language with a very large user base.

PostgreSQL is a free and open-source relational database management system. It emphasises extensibility (an IT concept that provides space for further growth) and SQL compliance. The extensible database management system was originally developed at the University of California and uses a relational data model (data is organized with hierarchies, and structures), [195]. The choice for an SQL database in a city environment where data can be provided in very large quantities, may seem striking from that perspective, however, it is the urban context that provides hierarchies itself, with cities owning towns, who in turn contain neighborhoods, that are composed of multiple districts with dozens to hundreds of buildings.

Enabling the storage of spatial data, PostGIS is a database extender for PostgreSQL. It provides more than a thousand geospatial functions that work with both vector and raster data, [15]. It supports all the open 2D and 3D spatial data types as defined by the OGC. Additionally, a second extension made the modeling of networks possible. This is PGRouting, [130], which was used extensively with all network models developed in this library (integrated water and electrical networks, iterative)

The coding of algorithms, as well as a functional coding interface, is programmed in Python. This is a programming language that has a very large base of users, [209] and support libraries. The large number of users means that it is attractive to other potential users and developers. Additionally, it can connect to the PostgreSQL database, easily import different types of file standards and provides for mechanisms that allow it to be connected to other stand alone software and platforms.

6.2.2 Components - models

The focus of this section lies in the inputs, outputs, algorithms and their contribution to the aEneAs urban modeling library. The library contains 14 models that were developed on top of the same infrastructure, figure 6.3. These models are all serviced on the same urban spatial model, which is stored in the PostgreSQL 3DCityDB database schema. Buildings and other spatial objects referred to in the models, are identified using the PID system presented in chapter 3, the scale dependent identification system (3.6). In addition, two extensions to the 3DCityDB relational database were used, representing a relational schema of the EnergyADE and the UtilityNetwork ADE, explained in their dedicated section 4.3.3, and documented further in [232] and [33]. A series of network analysis functions, implemented in SQL, are stored within the PostGIS extension. These functions allow for the reading of semantic properties of elements, calculating composite physical parameters of a network, and performing simple topological routing.

The sections below provide only a bare description of each component, as this thesis focuses on binding this models into an UBEM capable library, and not on one particular model per se. It provides further support to the hypothesis of this thesis, that spatial models can provide a solid base for, and act as bridging elements, in modeling the urban energy chain. The main goal of these models is to assist in thermal building simulation. As a secondary objective the modeling of urban energy networks provides a means to the transportation of energy.

In addition to the models presented in this section further development has taken place. Models were developed for statistical analysis of the building stock and refurbishment measures, optimal electric mobility stations placement, PV potential, city services accessibility and iterative expansion of utility networks using the same infrastructure and data principles. These models have been used in the production workflow of the EDF city simulation platform and are not presented in this section. As these models often required distinct spatial calculations the library was also extended so as to allow for the calculation of spatial functions. These functions are: volume calculation, surface classification and CityGML data quality checks. The development of these functions was also a time costly endeavour in itself, however, the work falls outside of the boundaries of the present manuscript and will not be presented in the section below. See [210], [166], [233], [234], [208], [33], [34], [224] and [223] for further references with regards to these spatial operations.

Heat losses

Developed in the context of a master thesis [210] under the direct supervision of this dissertation's author, the model was partially published in [166]. This model is designed to calculate heating losses for a building. The depending on the building geometry and external temperature. The mathematical background is provided by the Passiv Haus Method developed by the Institute for Passive House, [72], also available in the form of a software package, [73].

Figure 6.4 depicts a fact sheet of the model with the most important characteristics and graphic results. On the left hand side of the fact sheet information regarding the model is depicted. The model's purpose is defined as estimation of the thermal energy losses in a building. The modeled resolution is a building or a group of buildings when no distinction could be made with regards to building thermal envelope among a group of buildings. Further, the fact sheet specifies that the model is connected to geometry. This means that certain input variables are calculated by making use of the 3D spatial model. It also implies that the model uses the unique identifier system allowing the model to extract information straight from the spatial database. In the next lines the input, output, relevance and accuracy of the model is defined. In the lower left hand corner, the spatial scales involved are presented. The model resolution is presented with red, while yellow depicts the optimal design scale. When these two coincide, then the colour used is orange. On the right hand side of the illustration graphical results from the model are depicted. These are heating losses estimates for buildings in the study area of Weidenweg, Karlsruhe and the distributed relative heating demand for five IWU building types. This fact sheet information pattern is repeated with following models in order to provide a standardized description of aEneAs's components.

The spatial input data required by this model are represented by the building model. All the subsequent work assumes that the model was corrected in previous ETL operations from a topological, semantic and syntax perspective. Spatial calculations are performed that allow for the building's inner volume and the further classification of external surfaces to be performed (into roof, wall and ground types). Further, the areas of each type are calculated. Classifying the previously mentioned surfaces of the building envelope allows for the further energy relevant calculation of gains and losses. Each of the corresponding building surfaces is allotted a thermal transmittance **U** value depending on the building age and type. These U-values originate with the IWU (ger.: Institut Wohnen und Umwelt) institute, [112], which provides a German national database for building types. The previously calculated parameters provide input to the calculation of the transmission and ventilation heat loss. Using additional parameters, namely the temperature zone factor (an indicator that adds climatic influences to the calculation) and air capacity coefficient the calculation of the building heat loss ($\mathrm{W\,K^{-1}}$) is performed using the following formula:

$$P_{hl} = \sum (A_i * U_i * f_t) + (V_A * n_A * C_a ir)$$

Where:

A_i	– is the component surface area, ($\mathrm{W\,m^{-2}}$)
U_i	– thermal transmittance, ($\mathrm{W\,m^{-1}\,K}$)
f_t	– factor for the temperature zone
V_A	– inner building volume, ($\mathrm{m^3}$)
n_A	– building air exchange rate, (ACH)
C_{air}	– heat capacity of air, ($\mathrm{J\,K^{-1}}$)

Hourly values of temperature data allow for the calculation of a Heating Degree Day (HDD) value (the difference between the set inside temperature of a building and the daily average outside temperature). These values are calculated only when the average day temperature was below 15°C for three days in a row (so that the thermal inertia of the building allowed for the captive heat to be dissipated). Aggregating all daily values produces the HDD annual value, HDD_a (K h):

$$HDD_a = \sum_{1}^{12} ((H_D * D_M) * (T_i - T_0))$$

Where:

H_D	– heating hours a day
D_M	– number of days per month
T_i	– set temperature inside the building, (°C)
T_0	– daily average temperature per month, (°C)

Finally the energy demand corresponding to the heat loss of the building is calculated (W h):

$$ED_{hl} = P_{hl} * HDD_a$$

SolarB - Solar radiation

This model calculates the amount of solar radiation that a building receives throughout the year. It is largely based on a method developed in the context of a master thesis [233], under the direct supervision of this dissertation's author, with the model being partially published in [234]. This component of the library received the name SolarB. The theoretical background was provided by the method presented in [244], which remains largely unchanged even in this thesis's implementation, as well as in other software tools, such as GRASS GIS, [102].

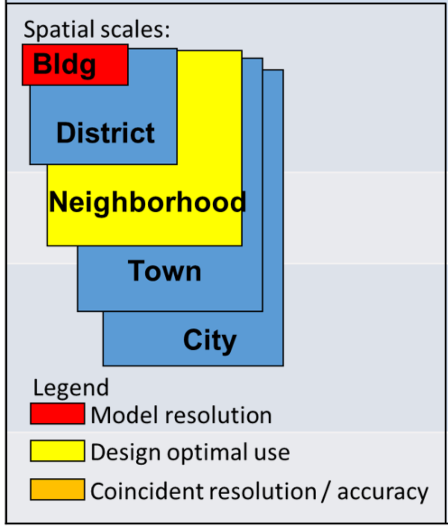
In figure 6.5, the model fact sheet can be found portraying the most important characteristics and graphic results. This model is designed to focus on the largest contributors to solar radiation on a building, the direct and diffuse radiation. Reflected radiation, as a minor contributor (<2%), [244], is ignored in this model. In order to calculate the solar radiation a point grid is first devised on the buildings included in the study (this varies in size, testing showing that a grid size of 0.5 meters was sufficient for reaching high accuracy at the level of a building, [234]). The model also has the ability to vary the time resolution of the calculation. All building surfaces are analyzed, with shared and ground surfaces excluded from the calculation. For each time point the corresponding sun position is calculated. Further, the beam and diffuse irradiance are computed, similarly for both horizontal and inclined surfaces B_{hc} and D_{hc} (W m^{-2}), only once per time step. Shading calculations make use of extensive PostGIS functions and are the most computational intensive step of the algorithm. Converting from irradiance values per point to monthly radiation per surface and then further to a building estimate is done with the following formula:

$$G_P = A_P * \sum_i^n \Delta t * (B_{Pi} + D_{Pi})$$
$$G_S = \sum(G_P)$$
$$G_B = \sum(G_S)$$

Model fact sheet

Graphical results

Name: **Heat loss**	
Purpose	Thermal energy losses
Modeled resolution / spatial object	Building / Group of buildings
Method	Passiv Haus Method
Connected to geometry	Yes
CityGML Level of detail	2
Relevant spatial scale	Building → City
Input	External building envelope surface, building volume, corresponding U-value, weather parameters, energy-efficient air exchange, air heat capacity
Output	Heat loss for building hull HDD (Heating Degree Day)
Relevance	District and Neighbourhood scenario impact estimate when used in brown-field refurbishment planning as well as green-field energy use amelioration
Accuracy	Low at building unit level, optimal at aggregated neighbourhood unit



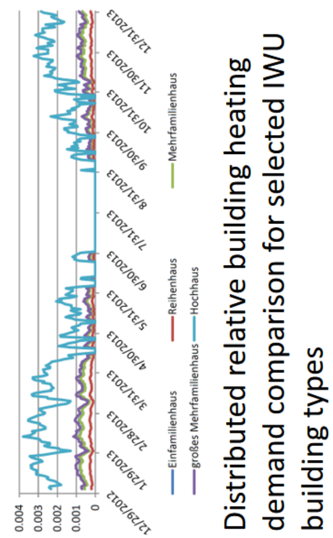


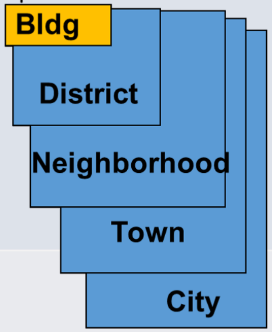
FIGURE 6.4: Heat loss model fact sheet and associated graphic results, developed after [72], and documented in [166] and [210]

Model fact sheet

Graphical results

Name: Solar radiation	
Purpose	Building solar radiation calculation
Modeled resolution / spatial object	Building / Group of buildings
Method	Wieland et al, 2015, Šúri and Hofierka, 2004, Wieland, 2014
Connected to geometry	Yes
CityGML Level of detail	2
Relevant spatial scale	Building → City
Input	External building envelope (surface elevation, aspect and slope), geographical parameters (latitude), meteorological (Linke turbidity index), measured ammount of global radiation, derived ratio of diffuse/global radiation
Output	Global, beam and diffuse irradiance per time step
Relevance	Energy balance for building and district, solar energy installations
Accuracy	High at building unit level

Spatial scales:



Legend
- Model resolution
- Design optimal use
- Coincident resolution / accuracy

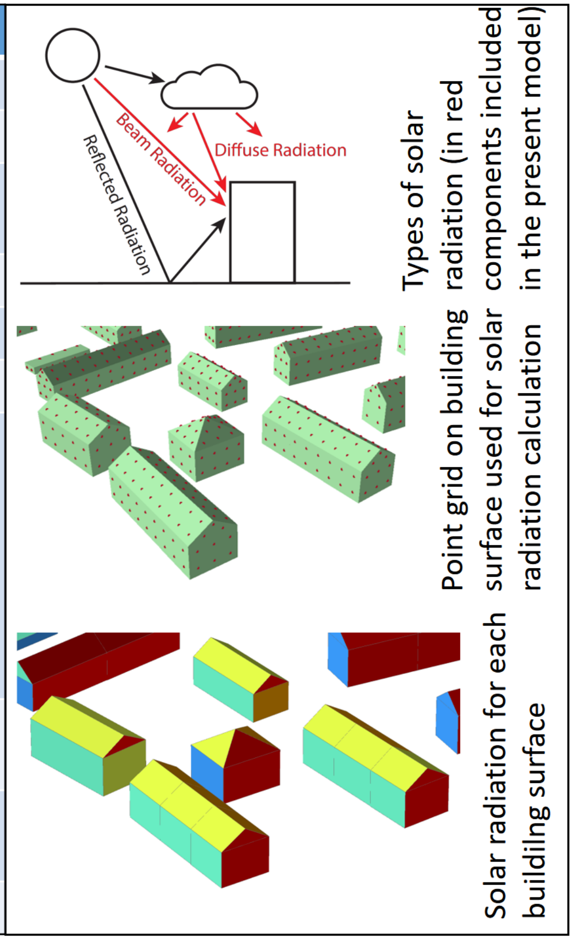


FIGURE 6.5: Solar radiation model fact sheet and associated graphic results, developed after [244], and documented in [234] and [233]

Where:

G_P, G_S, G_B	– are the global radiation on a point, surface and building, (W h)
A_P	– area represented by each surface point, (m^2)
Δt	– time step length, (h)
B_{Pi} and D_{Pi}	– Beam and diffuse components of solar irradiance for a point P at time step i, (W m^{-2})

The accumulation of the resulting irradiance values for the total time steps produces the corresponding radiation values per square meter. Radiation values are interpolated linearly for each surface point for each day of the year. Further, the summation of global radiations value can be performed at the level of month, trimester and yearly. These summations can be performed at the level of a surface or building depending on user requirements.

A further application of the solar radiation model is the calculation of photo-voltaic (PV) potential with different photo voltaic technologies. A model allowing this calculation was proposed for aEneAs, to be built on top of the solar radiation model presented in paragraph 6.2.2. However, no implementation in this library was made that past the technical proposal stage. A model developed by colleagues at EIFER, that follows in the steps of the SolarB and further developed parts of the code, was published in [160].

Solar gains

Once the amount of solar radiation has been calculated it is then possible to calculate the gains from solar radiation in the buildings heating demand equation. The summary of this model is depicted in figure 6.6. Most often, the available spatial data is missing window relevant information, as can be seen with public CityGML in the federal states of Germany, [81]. It was decided to approach this issue from a statistical perspective. The actual window surface is estimated using ratios provided by the IWU institute, in their systematic survey of the German building stock, [112]. All external building surfaces receive corresponding window ratios according to the year of construction and building type. In this way, the nominal value of the window surface is obtained.

The IWU institute also provides a method to calculate window solar gains, [143]. This method is only applied on heating days (as defined with the Heating losses model, 6.2.2). The value for solar gains is obtained at the relevant resolution desired and can be determined per surface or per building. The following formula was used:

$$Q_s = r_i * g_{normal,i} * \sum_i G_i * A_{F,i}$$

$$r_i = r_{Frame} * r_{Shadow} * r_{Dirt} * r_{not-normal}$$

Where:

Q_s	– Solar gains, (kWh/a)
G_i	– Global radiation during heating days, (kWh/m^{2}a)
$A_{F,i}$	– Window area, (m^2)
$g_{normal,i}$	– Transmittance value of window glazing (single, double or triple), (W m^{-1} K^{-1})
r_i	– Reduction factor
r_{Frame}	– Reduction factor due to non-transparent window parts
r_{Shadow}	– Reduction factor due to shading
r_{Dirt}	– Reduction factor due to dirt on the transparent window area
$r_{not-normal}$	– Reduction factor due to the angle of contact in between light and window surface

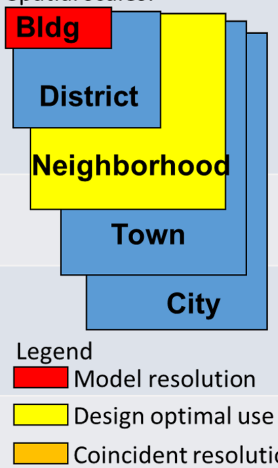
PV Paneling

This model is used to provide optimal positioning of the photo-voltaic panels that are to be installed on the building roof tops. The model was first documented and published in [208]. The main goal of the implementation is to seek the optimum spatial position of PV panels on the roof of a building structure. In order to propose such an installation, the place and orientation of the panel needs to be determined on the roof of the building. In order to provide this estimate with regards to the electricity produced by a PV panel, solar irradiation values for a typical year were produced for the building rooftops. Typical years are meteorological average years produced with a dedicated software solution, called Meteonorm, that uses stochastic generation, [192]. For the technical details of the standard PV panel dimension, multiple manufacturer technical specifications were found and averaged to obtain an estimated surface size of 1.5m^2 (1*1.5m).

With regards to the actual positioning and filling of the roof surface two methods were proposed and tried. In the first method, given limitations to the weight of the panels or to the cost of the panels are assumed. These are provided by giving a maximum number of panels that can be installed on a building roof. The solar irradiation is then calculated for a point grid resolution of 0.5m using the solar irradiation model, 6.2.2. The roof grid points are sorted depending on the recorded value of total solar irradiation. In the next step, a PV panel is created on the roof top, in portrait or landscape, around the point with the highest radiation value. The algorithm then iterates through the following points to determine the optimal position for the panel, and if need be restarts the process with a lower ranking point. The process is repeated with the remaining points so as to include the maximum number of PV panels.

The second method is designed for use at district, neighborhood, town and city level. It divides the building's roof surface into PV panel sized pieces, and then calculates the amount of solar radiation received by each piece. This implementation adjusts the grid size definition from the solar irradiation model, 6.2.2 to 1*1.5m, thus removing the need for further processing time solely for the paneling model. The point

Model fact sheet	
Name: **Solar gains and panelling**	
Purpose	Building solar gains and Panelling
Modeled resolution / spatial object	Building
Method	IWU, 2005, Wieland, 2014, Loga and Imkeller-Benjes, 1997, Sexton, 2017
Connected to geometry	Yes
CityGML Level of detail	2, if available 3
Relevant spatial scale	Building → Neighbourhood
Input	External building envelope (total surface), building typology parameters (window glazing thermal transmittance, reduction factor), global radiation, HDD
Output	Solar gains per time step, per surface, per building, PV pannel distribution
Relevance	Energy balance for building and district, solar energy installations, building optimal energy balance
Accuracy	Low at building unit level, optimal at aggregated neighbourhood unit



Graphical results

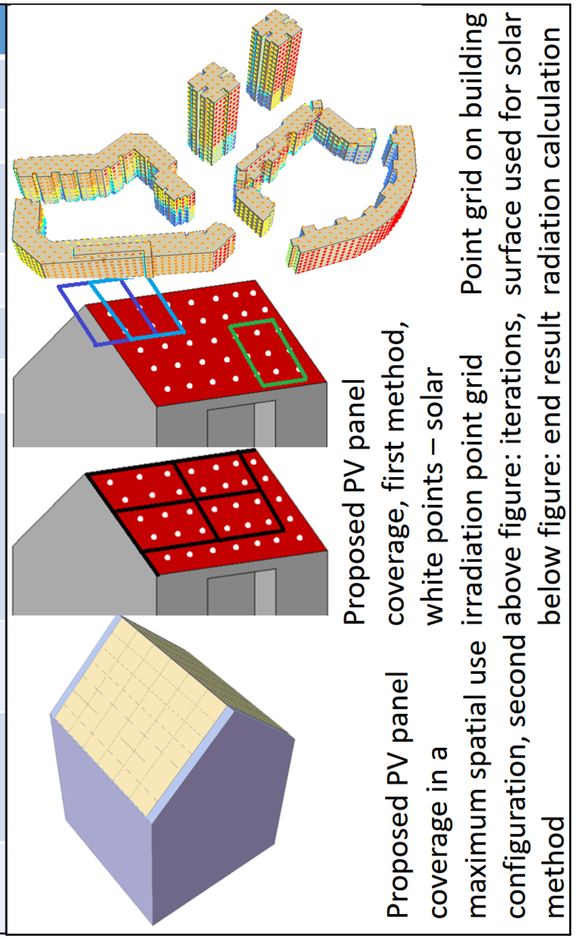


FIGURE 6.6: Solar gains and PV paneling models fact sheet, developed after [112], [143], and documented in [233], [200] and [208]

grid dimensions are adjusted for portrait or landscape orientation, in order to ensure maximum coverage of the roof surface. For performing city wide PV panel estimates, this second method was found optimal, even as it is less precise at the level of a building (lower resolution point grid and less flexible positioning of the PV panel).

Heating demand and associated greenhouse gas emissions

This model uses the input provided in the first and the third models, the heating losses and the solar gains based on the buildings external hull, to calculate the total heating demand. Heating demand is the difference in between the losses and the gains, as defined and implemented in [199]. This method makes assumptions with regards to the other components, that are shown in [81] to be otherwise large influencers at the scale of a single building, occupancy behaviour, infiltration rate, wall to window ratios. This assumptions inflict a large reduction in the accuracy of the results at the level of a building and even each other out at significant larger scales, neighborhoods and towns, [121]. The process flow that allows for this this calculation to be performed, as well as the model fact sheet are depicted in Figure 6.7. The following formula describes the calculation of energy demand for heating:

$$ED_{heating} = ED_{hl} - Q_s$$

Where the two components are described in the Heat losses (6.2.2) and Solar gains (6.2.2) models.

GHG emission is a prevalent term used to describe greenhouse gases (CO2, CH4, and N2O). In this model's case, the results provide an estimate with regards to the amount of GHG emitted with the generation of thermal energy for building heating. The method used is presented in [200], and estimates the total amount of CO_2 equivalent gasses. It converts other greenhouse gases into CO_2 by using the conversion ratios of the GWPV (Global Warming Potential Values) presented in [161].

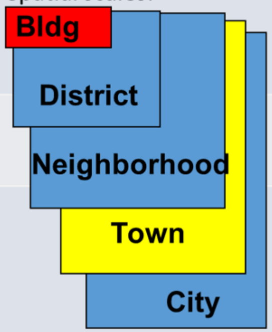
The method was applied in Karlsruhe, Germany where information on the heating sources was made available at the level of a town (Stadtteil) with the help of a commercial database solution called INFAS, [108]. This solution has been used previously with other scientific publications, see [107]. Statistical information with regards to the average GHG emissions (CO_2 equivalent was obtained from the IWU institute, [88]. The formula used in the calculation, presented in [200], is as follows:

$$CO_2Ee = \sum_i (ED_i * EF_i)/1000$$

Where:

CO_2Ee	– Energy related GHG emissions in Carbon Dioxide equivalent, (tonsCO_2e)
ED_i	– Annual heat energy demand for building i, (kWh)
EF_i	– Emission Factor for heating source of building i, (kg/KWh)

Model fact sheet

Name: **Heating demand and GHG emissions**		
Purpose		Building heating demand
Modeled resolution / spatial object		Building / Group of buildings
Method		Nichersu, 2014, Wieland, 2014, Loga and Imkeller-Benjes, 1997, IWU, 2005
Connected to geometry		Yes
CityGML Level of detail		2
Relevant spatial scale		Building → City
Spatial scales: 	Input	Heating losses Solar gains GHG emissions from heating
	Output	Heating demand for building GHG emissions from heating
	Relevance	Energy balance for building and district, Scenario testing for brown and green field projects
Legend Model resolution Optimal use by design Coincident resolution / accuracy	Accuracy	Low at building unit level, Optimal at neighbourhood or town unit

Model flow

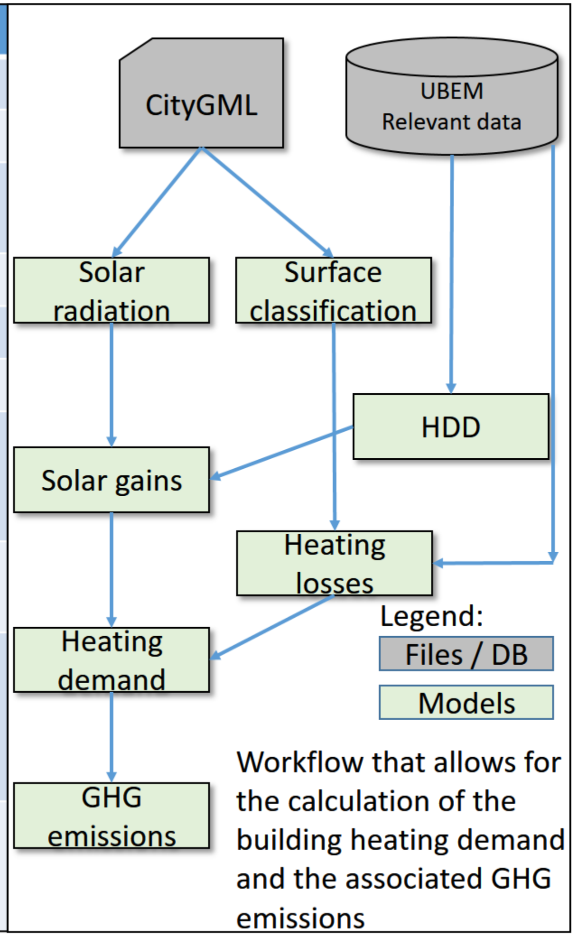


FIGURE 6.7: Heating demand and associated GHG emissions fact sheet and computation workflow, developed after [72], [112], [143] and documented in [210], [166], [200] and [233]

Refurbishment cost estimates

This component of the library allows for the approximate calculation of the refurbishment cost as well as the potential improvement to the buildings heating demand by making use of the spatial and energy relevant parameters of the building. The potential improvement to the building heating demand makes use of the heating demand model presented in paragraph 6.2.2 and presents improvements to the building by calculating the differences in between the two calculations (before and after refurbishment). It uses a method described by the Center for Construction Costs of the German Architect Order (Ger: Baukosteninformationszentrum Deutscher Architektenkammern - BKI) in their documentations, [32], [28], [29], [30], [31]. These publications include 20 building types, that can be linked (as they are one and the same) to the IWU building types, [112], mentioned before with the heat loss model, 6.2.2. The implementation of this model was facilitated by the work of an assistant researcher, and was described in [113].

Refurbishment is a precondition defined by the EU in the New Green Deal, [71] so that the climate protection goals, agreed under the Paris agreements in 2015, can be achieved. As such, it was the next target for the aEneAs modeling work, as building physical parameters are large influencers on thermal energy demand, as shown in [71]. There are four types of refurbishment, as defined by the Directorate-General for Energy of the EU, [70], defined according to the different renovation depths, relating them to non-renewable primary energy savings achieved in a specific calendar year:

- Below threshold ($x < 3\%$ savings)
- Light renovations ($3\% \leq x \leq 30\%$ savings)
- Medium renovations ($30\% < x \leq 60\%$ savings)
- Deep renovations ($x > 60\%$ savings)

According to [70], in 2019, the vast majority of renovations fall in the first category and are performed in a step-by-step manner, with gradual improvements over a long period of time. At the opposite end, deep renovations are performed in less than 0.1 to 0.3% of all cases in a relatively short time span. The BKI costs indicator for refurbishment, [32], largely focuses on bringing buildings to the current energy use standards. This construction works would fall under two categories: medium and deep renovations, depending on the buildings original construction year. They bring a significant improvement to the building in terms of primary energy savings.

With regards to the workflow of the model, the aEneAs frontdesk is responsible for logging and bookkeeping of the program running process, [113]. The cost estimate provided by the BKI fluctuates in value depending on the Federal State in which the building is located. As such, the ownership relation (building to Bundesland) has to be performed first, making it the first parameter used in the price estimation. The construction indices for different states and the historical indices from the table were first normalized so as to provide mean values for each indicator used in the refurbishment (cost per square meter of useful surface and per building volume). Further parameters used in the classification of price levels for refurbishment were the refurbishment type, which distinguishes in between deep and medium renovations, and spatial indicators coming from the 3D city model (building useful area, volume and roof surface). To help classify the buildings according to the 20 IWU categories, the INFAS commercial database solution was used, [108].

As per [30] and [32] the formulas below were used for the calculation. The formula for the normalization of building refurbishment costs with help of the indices for different states (example provided for BW, Baden-Würtemberg, where the city of Karlsruhe, used in this application is located):

$$K_{new}^{Land} = \frac{K_{old}}{k_{m^2 Land_i}} * k_{m^2 Land_{BW}}$$

The following formula was used for the normalization of building refurbishment costs with help of the historical indices:

$$K_{new}^{Year} = \frac{K_{old}}{index_{RefurbYear}} * index_{CurrentYear}$$

Finally, the formula for the cost estimation for new buildings provides the estimated refurbishment costs in € at the level of a building:

$$RC_i = K_{new}^{Land} * K_{new}^{Year} * UA_i * NF_i * Price_{m^2}^{type}$$

Where:

RC_i – Refurbishment cost for building i, (€)

UA_i – Usable area for building i, (m^2)

NF_i – Number of floors for building i, -

$Price_{m^2}^{type}$ – Refurbishment price per square meter for the same building type, $(€/m^2)$

Different refurbishment strategies can be tested for medium and deep refurbishments, and the application can be reduced in implementation for historical buildings, whose street facing façades have restrictions with regards to the type of refurbishment that can be performed.

Greenery as building installation

This model estimates the impact of green roofs or green façades on a building's heating demand. The model was first documented in the frame of a master thesis, [223]. Greenery solutions are part of a group of solutions referred to in the academic environment as *nature-based solutions*. Green façades have been shown to be easier to install on a building. At the same time they present large ecological benefits with low associated costs. Green roofs require more investment in time and money but are also effective, [127]. With regards to their impact on building energy demand, these installations help reduce building solar gains in the summer time and reduce heating losses in wintertime, [224]. Figure 6.8 depicts the model fact sheet, concept and the 3D SketchUp model of the Illingen test site used in the validation of the model.

This model is formed from two separate implementations, one for green façade and another one for green roofs. Both implementations follow the description provided by Susorova et al. in [214]. The main difference in implementation stems from the

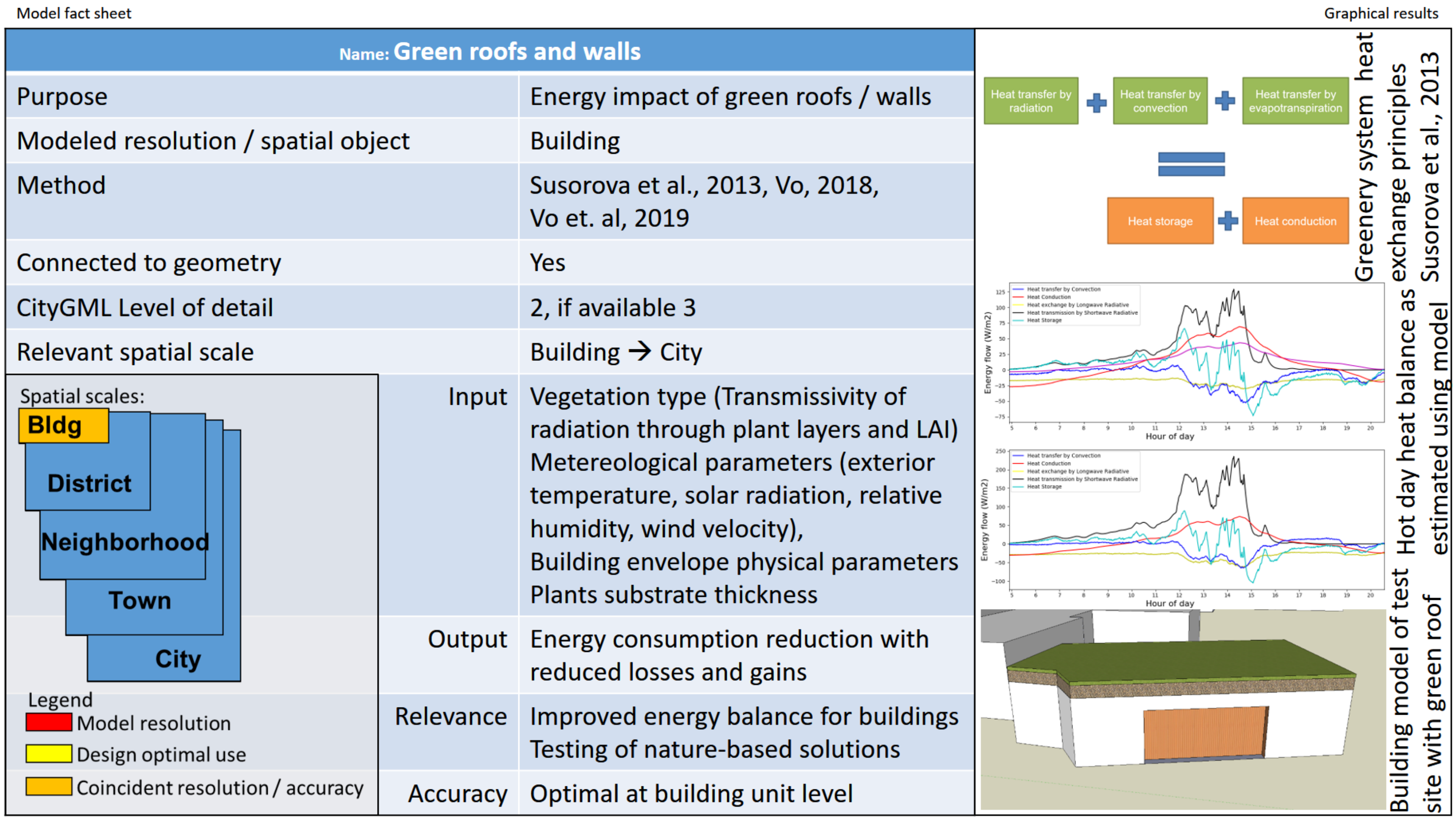


FIGURE 6.8: Green roof and green façade, fact sheet, concept and graphical results, developed after [214] and documented in [223] and [224]

fact that for the green façade no soil layer is included in the simulation. The primary thermal exchange processes of the green façade model is presented in figure A.3. The following formula represents the heat balance equation and, most importantly, the heat stored in the wall behind the vegetated façade (S_{vw}) used in the model:

$$SR_{vw} + LR_{vw} + C_{vw} + XR_{vw} = Q_{vw} + S_{vw}$$

$$S_{vw} = L * c_{pstr} * \rho_{pstr} * (\frac{dT_{S_{vw}}}{dt}) + D * (c_{psub} * \rho_{psub} + c_{pwater} * \rho_{pwater}) * (\frac{dT_{S_{sub}}}{dt})$$

Where:

SR_{vw} – Shortwave radiation to/from the vegetated façade (W/m2)
LR_{vw} – Longwave radiation to/from the vegetated façade (W/m2)
C_{vw} – Convection to/from the vegetated façade (W/m2)
XR_{vw} – Radiative exchange between the bare façade/soil and plant layer (W/m2)
Q_{vw} – Heat conduction through the wall behind the vegetated façade (W/m2)
S_{vw} – Heat stored in the wall behind the vegetated façade (W/m2)
L – structural thickness (m)
c_{pstr} – Specific heat of structural material (J/kg K)
ρ_{pstr} – Structural material density (kg/m3)
D – Subtract thickness (m)
c_{sub} – Specific heat of substrate material (J/kg K)
ρ_{sub} – Substrate material density (kg/m3)
c_{pwater} – Specific heat of water (J/kg K)
ρ_{pwater} – Water density in substrate layer (kg/m3)
t – Time (s)

The green roof model includes additionally the behaviour of the substrate layer (the soil layer of the plants), which can be observed in the previous formula. S_{vw} is computed for both implementations using the temperature behind the vegetation and the substrate layers using a numerical bisection method, also presented in [214].

The model was validated for using monitoring data from real-time experiments during summer measurements at three locations in Germany (see 5.2 and [223] for more details with regards to the monitoring installations and validation of results).

Integrated water and electrical networks management

The previous models take a building centered approach, where energy use is analyzed within the *micro cosmos* of a single construction and is then replicated district and city wide. With regards to the energy utilities aspect, a model was developed on top of the same infrastructure that illustrates the capabilities with regards to networks and related spatial calculations. Figure 6.9 present a fact sheet for the integrated water/electrical networks management model.

The model was developed within the framework of a master thesis, [34], and partly documented (with regards to modeling of the water network) in [33]. This dual network behavioural model simulated 24 hours of network operation at a time step of 30 minutes. The networks each have associated usage profiles sets with higher measurement rates, however, for the model to run smoothly and show the downward effect of a network event this time step was considered sufficient. It also included cascading effects, which can be triggered by modifying the set of input parameters. The integrated dual network acts as a demonstrator for the capabilities of modeling multi-utility networks using the UtilityNetwork ADE and semantical city models stored with CityGML.

In order to be able to perform such complex spatial operations a suitable data set was found in the city of Nanaimo, Canada, where a coupled electric/water network exists. There, an energy reclamation facility converts the incoming flow of freshwater and generates electricity. Thus, the two networks and their data models are linked at the water turbines. This is presented conceptually in figure A.4. The object class used to describe the turbine is the *CityFurniture* from CityGML. This solution can be replaced by making use of the *GenericCityObject* class.

The city of Nanaimo with their local Public Works company (and very benevolent staff) provided open data for their water and electrical networks and diagrams of the mixed plant, freshwater delivery/power generation facility. The water network pipes and appurtenances could be produced from the data provided by the city, while the electrical network had to be fabricated by using the road network (which was available as open data). The original water network data had to undergo a lengthy ETL process in order to produce a UtilityNetwork ADE data set. This work is documented in [33].

The behavioural model represents a typical day of operation on the water and electrical network sample. It includes both water demand and electrical production from the energy reclamation facility. Water demand is associated with the buildings in the nearby suburb from the water reservoir. As water from the water reservoir is being depleted by (building) usage it is replenished. The usage follows a consumption pattern provided by the City of Nanaimo Public Works staff, [34]. As the re-filling of reservoir is activated, the water pushing through the turbines produces electricity. Additional electrical production in the network stems from PV panels installed on the roofs of some of the buildings. This occurs only when the production exceeds the electrical demand in the respective building. The electrical output from PV production is replicated from that measured in a real local installation that provides open source data, the Nanaimo Food Share PV power production.

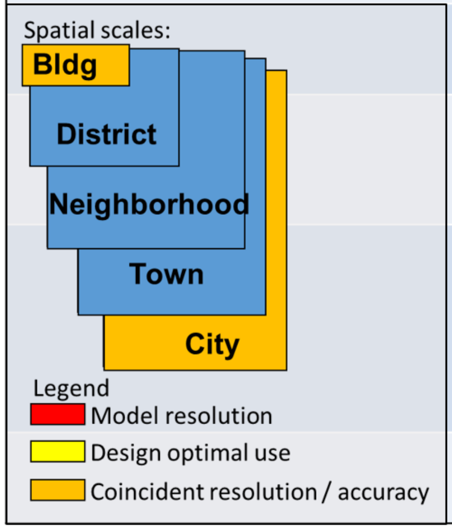
The behaviour of the water and electrical networks is mimicked using a linear time series model that is stored in a Python script. Values for properties of network elements are calculated at given locations (usually at appurtenances, bifurcations and buildings). These values impact the state of other elements in the network, and are able, for example, to open the valve for re-filling of the water reservoir at the Nanaimo water and energy reclamation facility.

In order to simulate cascading effects, the topology of the model is altered. This is depicted in figure A.5. If a water pipe belonging to the network is undergoing maintenance or an electrical cable fails, the topological link in between the elements is set to "broken". This change will impact further (connected) nodes, in what would be a typical cascade effect. The behavioural model that models this behaviour does not interfere with the data model, stored in the DB and keeps a separate log to the changes that occur to elements and their properties, thus maintaining a separation between the logic and data tiers.

Model fact sheet

Graphical results

Name: Integrated water and electrical networks management	
Purpose	Utility network interdependency behavioural model, cascading effects
Modeled resolution / spatial object	Network and buildings
Method	Boates et al, 2018, Boates, 2018
Connected to geometry	Yes
CityGML Level of detail	Not applicable for networks
Relevant spatial scale	Building → City
Input	Water and electrical network Event trigger
Output	Behavioural model with water and electrical flow Cascading impact of reduced output
Relevance	Open source data model for network interdependency (with topological connectivity and topographical dependency) Cascade mittigation testing
Accuracy	Optimal at building (for commodities user) and city level (network operator)



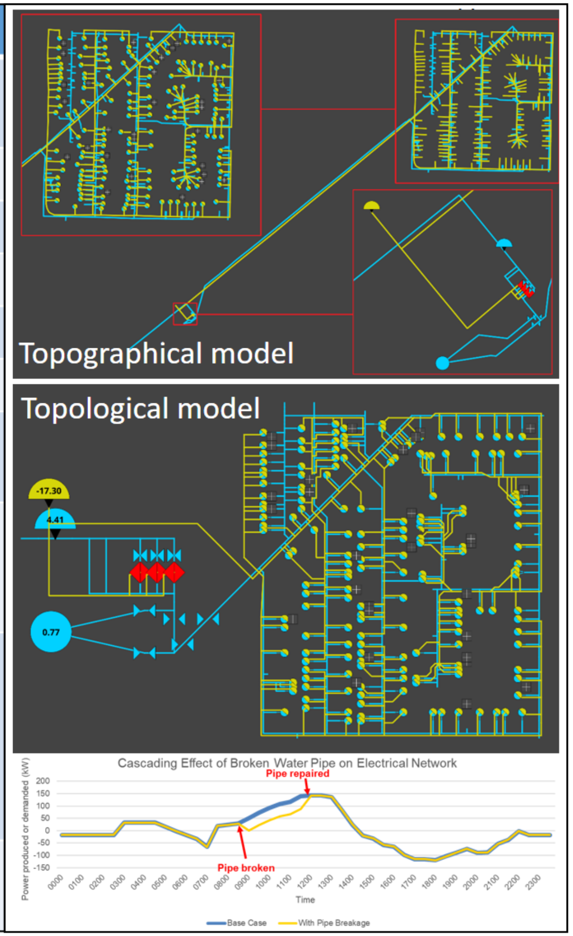


FIGURE 6.9: Fact sheet for integrated water/electrical networks management, documented in [34] and [33]

Chapter 7

Outcome - Conclusion and Outlook

In modeling the urban energy chain, there is no variable as potent as the *fabric* of space and time. It brings the multiple components of the energy chain together and allows for spatial and temporal dependencies to be correctly modeled collectively. This embraces Tobler's first law of geography, [217]: "[E]verything is related to everything else, but near things are more related than distant things."

Geographers are bound to observe the world while engineers help build it. To embrace both sciences, a facilitator is found in modeling and data integration, supported by spatial data in the form of 3D semantic city models that act as data hubs. These are stored in spatial databases and use web services that ferry data to and from one component to another without time-costly ETL processes.

7.1 Conclusion

This thesis' results endorse the idea that city-wide energy modeling is most efficient when performed at a neighborhood scale while using buildings as data units for consumers and production entities. Utility networks require modeling under new paradigms where both energy and information flows are bi-directional. End consumers are no longer just that; they are transformed into prosumers. This is especially true for the *Ger.: Energiewende* concept where precise estimates of local energy demand and production are required, and local energy production is preferred.

Modeling urban energy problems facilitates solving complex questions and equations as the spatial models become central linking elements of the different types of energy-related data. Testing includes urban simulation, executed to provide quantitative and qualitative energy site management assessments. Using scenarios, both green and brownfield urban development can be compared, and the influence of future contextual parameters, such as energy prices, warmer climates, and upgraded technologies, on new designs can be explored. This is also the primary use case of the EDF city simulation platform project, which has partially financed this thesis, for which the author is very grateful. These scenarios and their corresponding impact/results are then provided via visually intelligible/comprehensible predictions to multiple stakeholders. The proposed open style architecture enables joint decision making, a participatory approach, in which all actors can work together to understand the impact of proposed actions.

All buildings and their associated energy systems are unique. Therefore, their digital models must reflect this reality. Since the 1950s, there has existed a trend in creating standardized building components, for example, the use of identical utility network cabinets or shower/bath modules in sky-scrapers. These come pre-fabricated and are only installed on-site, as in large hotels. This is very similar to the proposed digital standardization approaches taken for modeling and further supports the use of standardized buildings data models.

The city wide energy chain

During this thesis, it was found that the term 'urban' in the GIS domain has a substantial variation in meaning and scope. As such, a description was provided that can be used to define urban spaces in the context of energy modeling using buildings as well as energy behavior and energy networks in this domain. Additionally, a detailed list of energy networks was defined that provides further input to the use cases described in chapter 4.

Urban energy chain modeling is part of the more significant trends in urban management. Important trends were also identified to provide context and explain the impetus for this thesis. These were smart or next-generation cities, smart districts, climate change mitigation solutions, open data, open governance, citizen observatories, standardization efforts, and the concurrent IoT/Big data revolution. A definition of a smart city was provided in the given context and a description of what it entails at the level of a district.

Scale and UBEM space

One of the puzzling research questions of this thesis relates to the right scale for urban energy chain modeling. The evidence presented in this chapter suggests that buildings, as defined in chapter 2, acting as energy chain data units, provide an optimum balance of scale and data accuracy, in the context of data revolutions, a significant step forward for better modeling platforms. However, for meaningful project implementation of modeling results, the most favorable size is neighborhood and district scale. To support its use in UBEM, data treatment methods and an ontology were presented that were successfully used in multiple cities.

Scale is defined in chapter 2 as a byproduct of resolution and extent with ranges provided for both parameters. To help solve this question, a standardized spatial scales definition was proposed. Further, an extensive survey was carried out to identify the actors who interact in the energy chain. The district scale, also referred to as Precinct, Building block, or Quarter in different contexts, containing between a dozen and a hundred buildings, was identified as the scale at which all actors, their actions, and their data meet.

The concepts presented in this manuscript make use of digital twins. Three domains provide spatial data to create such replicas, the so-called elephants: CAD, BIM, and GIS. A detailed survey of the scales used for these three domains was performed. It was observed that the three domains converge at the district scale. This coincides with the minimum scale deemed optimal for meaningful energy transition projects and gives further credit to UBEM modeling performed at this scale. This convergence of the digital domains and the spatial actors further facilitates the implementation of recommendations stemming from this thesis' modeling activities.

In the scope of this thesis, data was collected and partitioned to building spatial units and, with the help of ontology-based hierarchical relationships, amassed them further on to larger units. Analysis of the spatial relevance of data unit sample size in the case of UBEM models used for thermal energy demand provided further insight into the appropriate scale in terms of results' significance. Calculations were made using building units that are then aggregated at a neighborhood or a district level. This method facilitated the provision of accurate data at their respective spatial scales. With the current availability of monitoring data and its use in model validation and revision, there is a noticeable trend towards achieving high accuracy at lower scales, going as far as the building unit and the residential unit.

Two significant issues with urban energy modeling are data consistency concerning spatial scales and urban fabric contiguity. Spatial scale mismatches require extensive ETL operations. An exhaustive list of scale-related ETL operations was presented and their impact and the potential use of the produced data. These operations should not be underestimated, as they can occupy a significant amount of time in an urban energy model project's life span. Regarding contiguity, the non-interrupted urban fabric of building envelopes and neighborhoods helps to ensure energy savings and facilitates the creation of planning units. This issue plays a vital role in volume calculations, a major regressor in thermal energy calculations. Buildings, per se, constitute contiguous entities, and their use as spatial units allows for continuous urban pattern analysis.

A taxonomy of spatial standards and scales is presented in chapter 4. Formats can be classified by using four attributes: format/content related (standard type, data model availability, data content), design scale, energy thematic (energy chain position, energy commodity type), and entity type (public/private/citizen). These multiple characteristics illustrate and provide depth and dimensions to the complexity of the task at hand, modeling and monitoring the entities, their attributes, and interactions at play in the urban energy chain. Again, the scale appears as a major *separator*, as the lack of standardized spatial units in the design scale hinders further efforts of data integration and adoption of existing standards.

Standardized spatial scales and persistent identifiers

Holistic, integrated energy systems work on different spatial levels, from residential and tertiary use buildings, which are now active players in energy production, up to district, neighborhoods, city, regional and national scale. Modeling and simulation are the sole means of exploring the performance of new designs and concepts in this rapidly changing urban context. In section 3.3 a proposal of standardized spatial scales for use in UBEM is made. These unit scales are energy meter, building, district, neighborhood, town, and city. The scales were tested within different urban energy chains regarding economic, cultural, and geographical characteristics and are shown to be adaptable to new locations.

A scale-dependent urban data ontology with persistent identifiers was proposed to solve the identified problems. This open ontology avoids issues related to proprietary ontologies and cascading effects from their use and is built on top of the proposed standardized spatial sales. This human-readable code allows for differently sized cities to be modeled and mitigates associated issues, such as data heterogeneity and inconsistency in urban modeling domains. In this manner, all urban energy chain stakeholders are bound, within a non-biased perspective, by persistent means of identification. For a successful implementation of this proposal, it is not necessary that all energy chain actors use this system, but rather that the data is linked, in a central repository or a virtual hub, see 5.5. This repository needs to contain a mapping system that allows for objects to be identified depending on their ownership status (What is the entity in the chain that they belong to?). This will fix the missing spatial data issue and help increase the quality of recommendations that can be made on top of the modeling results.

On urban energy chain data

Data is only valuable when it is used.

Data in the urban energy chain is often stored and maintained in older client software and on locally isolated server systems. This is why integration of legacy and third-party software systems is a crucial step in the direction of smart applications, as other authors concur, [45], [134], [78]. As was presented in this manuscript, the energy chain contains multiple entities with inconsistent data structures and software. This also often involves legacy standards and software whose architecture was never designed to support the integration or sharing of data hosted by them with third parties. At times, this is a solution chosen for the sake of data privacy; at others purposely made by software developers. Part of the solution is found in ontologies that include mapping the overall perspective. These offer support for semantic heterogeneity and interoperability with external systems. They can be used as a conceptual scheme to represent multiple repositories that allow seamless integration. This is where the spatial scale-based ontology proposed in this thesis can help.

It stands to reason that data and thematic experts don't necessarily need visualization tools to understand the raw data presented to them. It was found that plenty of the tools used by geomatics, urban, and energy thematic experts are often a stumbling block, a hindrance of sorts, as they treat data without informing the user of such an action. For example, this can be seen with a well-known tool that repeatedly came up in user surveys, mentioned in chapter 4, for use cases behind the UtilityNetwork ADE, the Microsoft Excel, and CSV (Comma Separated Value) files. Users are left unaware that simply by opening a file, certain syntax elements are changed. Such events transform the original data without the user's intention and create issues within the large community of actors in the urban energy chain that have to share data among themselves and for the urban energy modeler to gather data from all of them. This is why an essential result of this work lies in the fact that for a data expert to understand the data, it does not necessarily mean visualizing it. Rather, it should be done by analyzing samples outside of any client software, usually directly inside a database repository, behind a web service, or by using a programming interface.

Performing data analysis without any treatment is the first step toward data integration in urban energy modeling. This conclusion should impact the way geomatics, urban, and thematic energy experts are trained, as it should be the first exercise in data handling. This outcome can also be observed in discussions in the data handling community where the ETL vs. ELT (Extract Load Transform) debate is ongoing. This author stands clearly on the ETL side, as it reduces data treatment times further down the pipeline. It also ensures private and other confidential data exposure risks are mitigated from the onset of a project by complying with strict data handling directives and laws pertaining to data handling.

During this thesis' experiments, certain situations showed that even when the same file standard or API was used by a data owner and a data user, discrepancies in data structure and representation can appear. This is due to misinterpretation as to the standard implementation or even *corner cutting* by data creators, resulting from insufficient quality checks or even incomplete implementation of standards. For example, this can be seen in CityGML data where solids are sometimes present, sometimes not, even in the same file with the same LOD. This can also occur, specifically regarding building thermal energy, with CityGML Building parts and CityGML Energy ADE building zones. These can be entirely differently interpreted, even inside the same country, as is the case in Germany, in two separate federal states, as shown in [81]. Similar situations occur with sensor or simulated data, as due to inconsistent metadata, attributes can be misunderstood, misspelled, or simply be missing. These bring about *tomato-tomahto* issues, the most often encountered example being kWh and kW, which stand for the total amount of energy used and rate of energy use. Therefore, standardization is not a universal panacea, but a step forward, one that requires constant implementation and verification as to the correctness of the work performed and proper communication with all stakeholders so that value can be created from public data.

Semantic city models as integration hubs

In this thesis, the hypothesis that urban spatial data in the form of semantic city models, and its use, by way of data modeling and analysis, provides the binding feature for enabling next-generation or smart cities was tested. To facilitate this work, this author oriented himself towards current IT work methodologies. The most used one currently, the AGILE method, requires as a first step the identification of use cases and users before any planning and coding can begin. For this purpose, twenty-five use cases were identified that relate to the urban energy chain and require the modeling of both spatial data in the form of semantic city models and thematic energy data together. Combined with an inventory of the spatial actors and their existing *modus operandi*, they provide the AGILE method with the required input. These use cases involve the components presented in the energy chain components survey and show the powerful impact of data with spatial and thematic energy characteristics within the hypothesis of this thesis. These surveys further facilitate future standardization efforts stemming from the ADE consortium.

Spatial and energy thematic data sets originating from single data sources describing entire urban energy chains are rarely available. The experiments presented a funneling system that, with the aid of time costly ETL processes, stored or linked all relevant information in/to semantic city models. Storage of urban energy chain data used semantic city models as a data hub for integration. These city models were represented using the CityGML standard and its extensions, the Energy ADE, and the UtilityNetwork ADE, as well as different web services for sensor data. Whenever possible, web services were used as an alternative to file handling. Their use allows for less time-intense ETL processes as web services often come with a high degree of standardization and available support documentation. This combined system allows data to be successfully fused using CityGML database structures.

CityGML remains the only spatial and GIS-related standard with capabilities that support modeling and monitoring the energy chain at the city scale using buildings as the smallest unit. Closed semantic models *hold the promise of interoperability in name only* and have been shown, in this manuscript, as well as in [101], to be the cause of further issues in horizontal and vertical interoperability. The fact that CityGML includes graphical, semantic, and thematic properties, taxonomies, and aggregations supports this manuscript's claim to the suitability of CityGML for modeling the entirety of the city-wide energy chain. These design choices allow for both UBEM input data and the urban interactions, be they spatial or energy thematic, to be modeled and simulated.

Monitoring: Energy relevant coupling of sensors and 3D models

Web services to the rescue

Digital twins, a mainstay of this thesis, also involve the use of monitoring equipment, or *things*. Three experiments were carried out using sensors. In the author's knowledge, this is the first trial that brought together CityGML, SOS-SES-WNS, SensorThingsAPI, and FIWARE. Section 5.2 provides a comprehensive understanding of the capabilities of these solutions. SensorThingsAPI turned out to be the easiest to implement, but it is also the lightest in terms of capabilities. Even though SOS-SES-WNS packs a stronger *punch*, the complex implementation, as well as the fact that there is currently no connection to NoSQL database solutions in the sense of Big data storage, means that FIWARE was the favored all in one solution.

The successful binding of spatial objects with the spatial models stored in semantic city models through the identification system proposed in chapter 3 allowed for the integration of sensors and energy data with 3D city models and represents a step forward towards truly interoperable smart cities. The results will aid readers of this manuscript choose a wise course of action for their applications. A significant aid in this thesis's experimentation with UBEM models, monitoring, and infrastructure was the previously mentioned identification of use cases and user types. Their proper definition is paramount before any implementation can take place, as it will help choose between different solutions and then lighten the actual system implementation.

The convenient alternative to file-based solutions lies in the different harmonized APIs presented. Data owners using a set of pre-defined operations allow for data querying. This approach considerably simplifies the time-heavy ETL operations mentioned repeatedly in this work. However, this approach also requires the hierarchical inter-object mapping to be performed using the spatially related ontology mentioned in chapter 3. *Something must bring urban energy chain objects and phenomena together, and that is the fabric of urban space.*

aEneAs, the scale aware modelling library

Due to the significant contribution of the building stock to human-induced climate change, UBEM can play an essential role in the planning phase. Impact estimation of urban refurbishment measures combined with monitoring are viable tools in the toolbox of urban planners and managers before greenfield or brownfield investments. Specifically, primary thermal energy demand is one of the most significant untackled topics in the energy domain, as nothing similar to the lighting improvement revolution has yet been seen for this part of the urban energy equation. This is the transition that UBEM seeks to help. aEneAs, the modeling library developed in the framework of this thesis starts by first quantifying primary thermal energy demand and then the impact of refurbishment measures. Lastly, it estimates the potential of renewable energy production. aEneAs also includes components that consider energy distribution in the given context, showing a path toward data modeling and simulation required for distributed energy production at the neighborhood and district level.

Holistic, integrated energy systems work on different spatial levels, from buildings, which are now active players in energy production, to district and city-level designs and national-scale infrastructure. The technologies identified in chapter 2 with regards to the urban energy chain meet at the level of a building. This scale is also where they will have maximum impact, as this is the scale at which people live and work and around which they usually plan their life. This is precisely why aEneAs was developed to work using building units. It performs complex energy modeling tasks using semantic city models stored in CityGML inside a spatial database.

Context is king

The technical solutions and implementations in the present thesis are designed to work in an urban setting and allow for the harnessing of the power of the spatial and temporal character of data in energy modeling. Standardizing the input of models and binding it to open standards allows for the re-utilization of advanced tools. It shortens the time to move one solution from one urban center to another.

aEneAs performs urban wide modeling and presents results at a single building unit level. The model results offer varying accuracy at the standardized spatial scales used in the thesis. Overall, good results are achieved at the district's scale and in some models at the level of a single building. The surrounding environment and its influence, such as shading and collected solar gains, are also included. This information can be used for integrated energy system planning and precise local renewable energy production estimates.

Good practices

The experience gathered from this thesis' experiments is provided as a synopsis in the form of good practices that can be recommended to cities that want to evolve towards next-generation cities. These are: the application of a common, adaptable, and reusable SDI, the development of reusable tools, the creation of open data and use of open source code, and the use of a digital by default policy. Once applied, these work principles should improve future work on the energy chain, both in terms of modeling and monitoring. They will also help facilitate intelligent cities initiatives, enhance public trust and promote value creation from public data.

In the thesis' experiments, a sturdy case is made for monitoring and modeling the urban energy chain using exclusively open-source software solutions and open standards for files and web services. A good example lies in the implemented data storage solutions: PostgreSQL and MongoDB. MongoDB is fast becoming an industry-standard while PostgreSQL is already one, being the top choice for spatial repositories. Furthermore, PostgreSQL is a 30-year running open-source effort that shows that open source is not just a trend but a phenomenon that is here to stay. It has jumped from one generation to the next and supports cities and modelers alike in many use cases. In addition, public data is a public good, and its availability, even limited, creates trust through open standards. This is a fundamental requirement for a well-governed city.

What could be concluded from this work is that there is no perfect one size fits all tool. Instead, the proposed good practices should be followed and implemented as much as possible to allow the entire ecosystem to be made interoperable and expanded in a sustainable manner.

Limitations

The findings in this thesis should lead to enhanced data handling and data quality in the urban energy chain. However, this is very much dependent on local know-how and the willingness to invest in such adaptation. The author of this thesis recognizes that, for the moment, such complex modeling and monitoring, even though central to becoming next-generation cities, is an expensive feature. This is currently only available to first-tier cities with enough capital for dedicated spatial departments and skilled personnel to deliver such integrated data sets. On the other hand, the willingness and political tenability to redirect substantial funds is also linked to political, cultural, and public recognition of environmental phenomena. Modeling of urban heat island impact on cooling demand, green facades, and green surfaces can easily find financial support in cities already impacted by climate change where the general public and state officials follow scientifically supported evidence.

Another significant limitation of the proposed solution is that entities are often moving in the ever-changing city. The presented modeling focused on buildings alone. However, a substantial part of the energy chain is not stationary. The modeling of spatially moving entities, such as buses, trams, and metros, is also possible within this methodology. This thesis has received basic testing in the AERO-TRAM use case where moving trams were visualized, and the moving measurement flow performed well while using the SensorThings API. Additionally, maintaining the city model and its versioning are not easy tasks in the current versions of CityGML, though its newest version promises to improve this capability.

7.2 Outlook - the future of semantic city models and the urban energy chain

Digital transformation

Cities, in general, are going through a significant transformation. This shift is part of a society-wide change and is generally referred to as the digital transformation. As shown in chapter 5, certain cities and countries are embracing it and will emerge fundamentally changed out of it. The world pandemic context has accelerated this change as it forced people and enterprises alike to use more digital services and enhanced the focus on interoperability.

This shift means that standardization organizations are asked to provide standards that work in an interconnected world of big data. At the same time, these standards need to embrace legacy software and data structures that were never designed for interoperability. This contrast requires a show of patience for the developing actors and support for the overachievers - the ones who set new paths for others to follow.

This author expects live or near-real updates of 3D semantical city models in first-tier cities in the next decade. This will be possible with the advent of very potent, universally available mobile sensing platforms.

This ever more extensive interoperable infrastructure will require data storing capabilities in organized and non-organized data structures. This is a strong incentive for the further development of the existing standards, particularly the UtilityNetwork ADE, which at the time that this manuscript was written is still primarily a work in progress. Furthermore, links to non-organized data repositories need to be created and maintained, such as FIWARE already foresees to MongoDB. High data volume will help improve modeling results and, at times, reduce the need for modeling as monitoring will help directly illustrate urban energy phenomena.

Initial objections towards XML as a solution in the IT domain regarding its size and complexity have lost steam as computers kept acquiring more computational power and internet bandwidth increased. At the same time, wide adoption of the CityGML standard remains an elusive goal, with such data sets only being part of the spatial repositories of first-tier cities, or better said, cities with available capital. At times, as in the case of the EU, national or regional spatial data agencies take it upon themselves to provide such data sets. This is a step forward, however, as shown in [81], data quality is not a given, and even though the format is the same, the content can differ.

In parallel to this work, the CityGML 3.0 has been developed, [124]. This thesis focused solely on the second version of the standard for which large open data sets exist. The new version proposes solutions to some of the points raised in this thesis. It remains to be seen how successful it will become, an indicator easily measured with the adoption ratio. However, the discussion surrounding CityGML v3.0 raised at the standard adoption by the OGC shows that the spatial community at large is very much interested in the further use and improvement of this standard, regardless if an improvement of 2.0 or adoption of 3.0 is the *correct* way to push ahead. The circumstances in which CityGML 3.0 came to be are similar, in a way, to the evolution of the Python programming language to version 3.0. Both Python 2 and Python 3 continue to exist. Still, slowly the community seems to be inching towards the later version, which presents a likely pattern for what will eventually happen with CityGML semantic city models.

Sharing is caring

Specifically, in the energy chain, the residential, tertiary, and industrial urban sites will be able to communicate and provide energy as a service to one another in the same way that the car-sharing concept is operating today. Cities and local decision-makers need to aim for energy sharing concepts at the neighborhoods and district level first, for a plethora of reasons, some presented in this very thesis. Standardization and data sharing among spatial actors will help locally produced energy be consumed locally, significantly reducing the need for large distribution networks and their associated losses, which follow the so-called distance decay law - related to Tobler's first law of geography. In a second stage, this can be shifted towards the following spatial scales of the urban divide so that the urban fabric contiguity discussed in chapter 3 is respected, and artificial boundaries are not an impediment. To achieve this, communication standards are required and improved and widely available utility networks.

While electrical networks are a given in urban settings, a far-reaching effort to develop heating and cooling networks is still required, with limited, yet significant, exceptions in mainly major planning-based economies. Existing infrastructure spaces, such as gas networks, need to be used to provide existing installations with energy and, potentially, make place for new utility infrastructure as space is a scarce and precious resource in the urban domain. Development of the various energy networks of the urban energy chain presented in chapter 2 can be an expensive long-term burden. This is why they should be included in comprehensive urban planning and management, where the broad horizontal and vertical contributions to the urban energy system can be used to their full potential.

This means that the planning and management of the urban energy chain need adjustment, a balancing act of sorts between multi-faceted perspectives. Semantic city models and web services are there to bolster this balancing act by enforcing interoperability, providing links between different actors and their data, and helping to model and monitor the many types of spatial interactions that cramped urban spaces see plenty of.

FIN

Appendix A

Figures from other sources

The figures included in this appendix are referenced to in different sections of the thesis and belong to other publications.

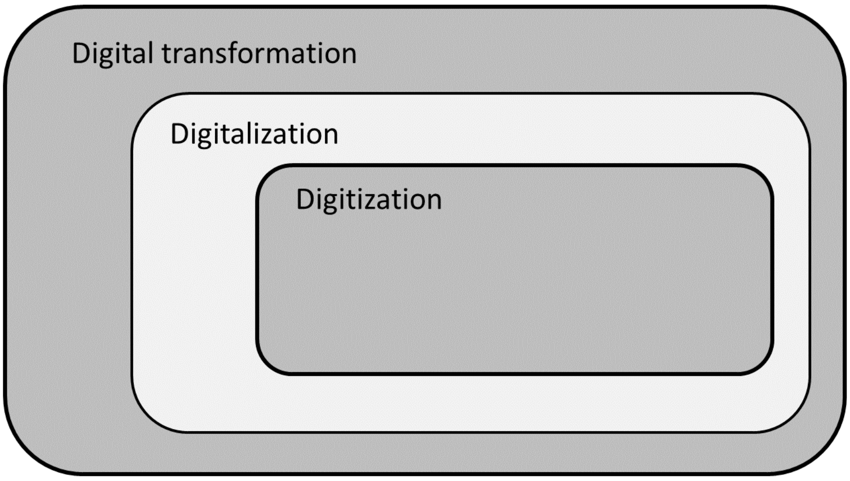


FIGURE A.1: Digital transformation, Digitalization and Digitization hierarchical relationship, discussed in section 5.4.4, from [197]

TRL 1 – basic principles observed
TRL 2 – technology concept formulated
TRL 3 – experimental proof of concept
TRL 4 – technology validated in lab
TRL 5 – technology validated in relevant environment
TRL 6 – technology demonstrated in relevant environment
TRL 7 – system prototype demonstration in operational environment
TRL 8 – system complete and qualified
TRL 9 – actual system proven in operational environment

FIGURE A.2: Technology readiness levels, marked with color are the levels at which the models produced in this thesis have been implemented, after [69]

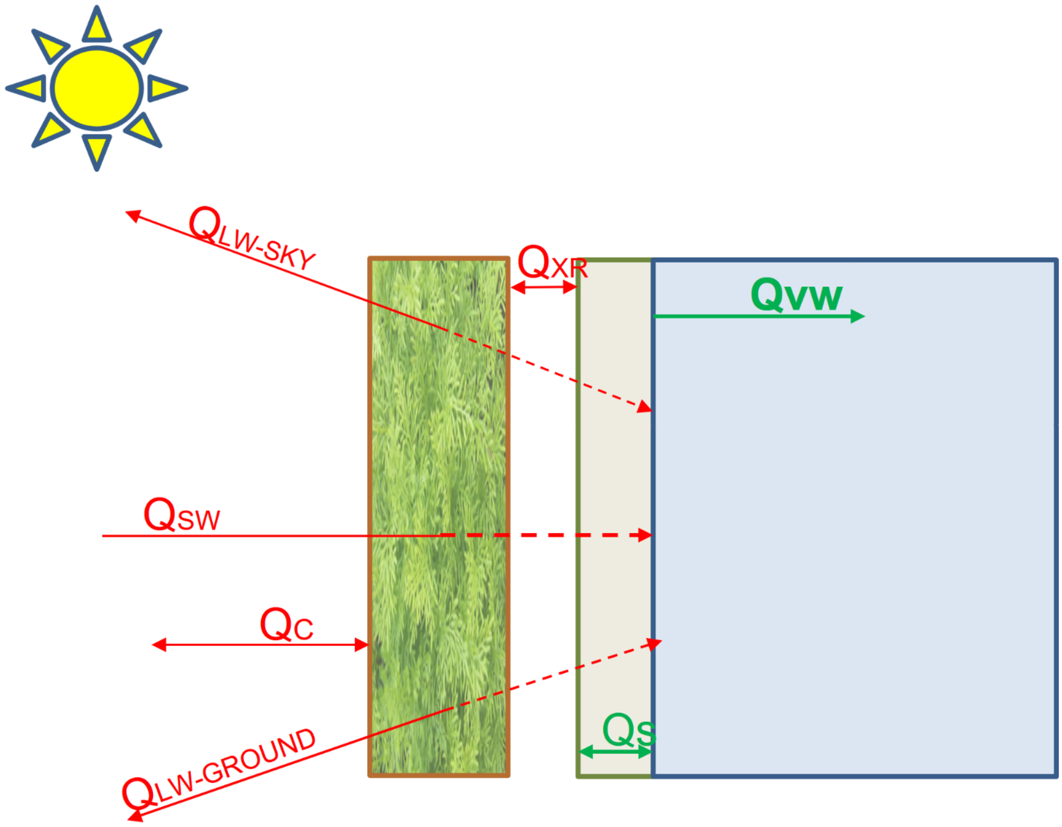


FIGURE A.3: Thermal components of the green façade model, as presented in section 6.2.2, from [223]

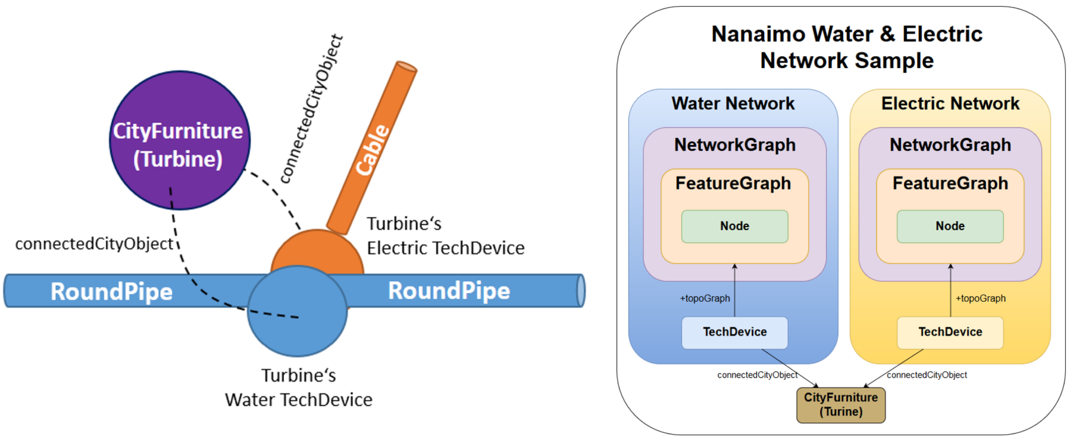


FIGURE A.4: Solution for coupling the water and electrical networks at a turbine, as presented in section 6.2.2, from [34]. On the left a conceptual representation of the means of coupling between the electric and water networks via the *connectedCityObject* attribute and a *CityFurniture* object. On the right a graphical representation of the data models.

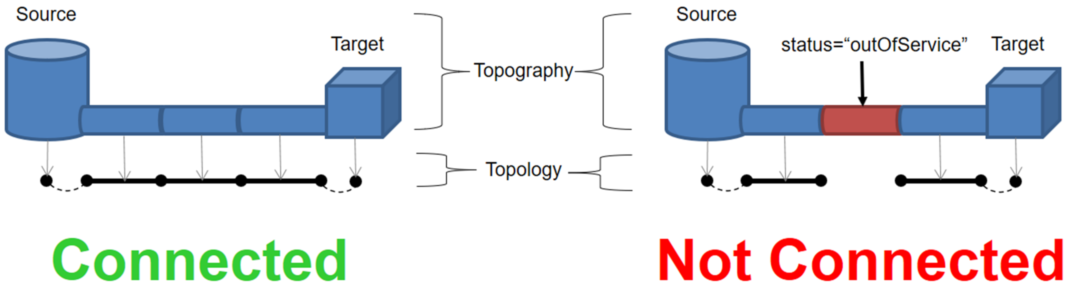


FIGURE A.5: Topological and topographical connectivity, conceptual representation, as presented in section 6.2.2, from [34]. On the left an uninterrupted pipe segment with both topographical and topological connectivity. On the right an interrupted pipe segment with topographical connectivity and topological break.

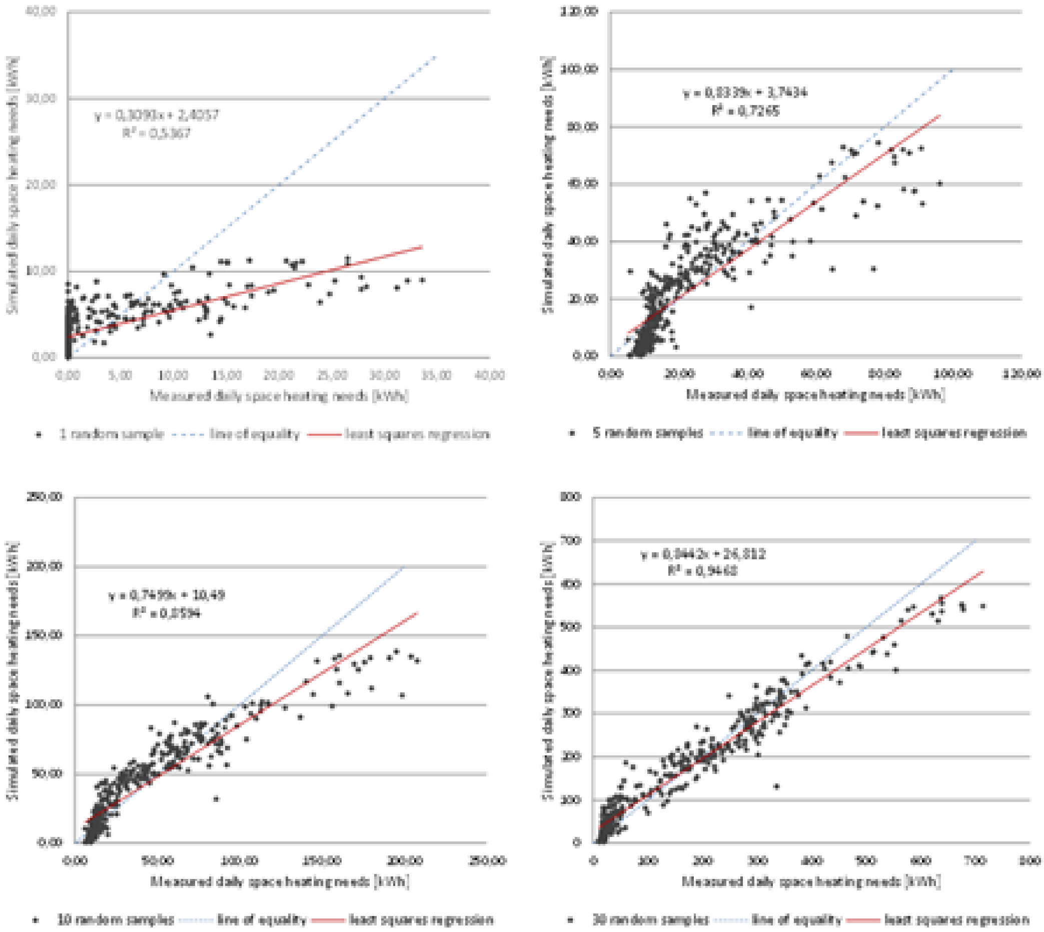


FIGURE A.6: Regression analysis of simulated against measured daily space heating needs for random samples of one, five, ten and thirty apartments (top left to bottom right) with least square regression function and line of equality, from [121]

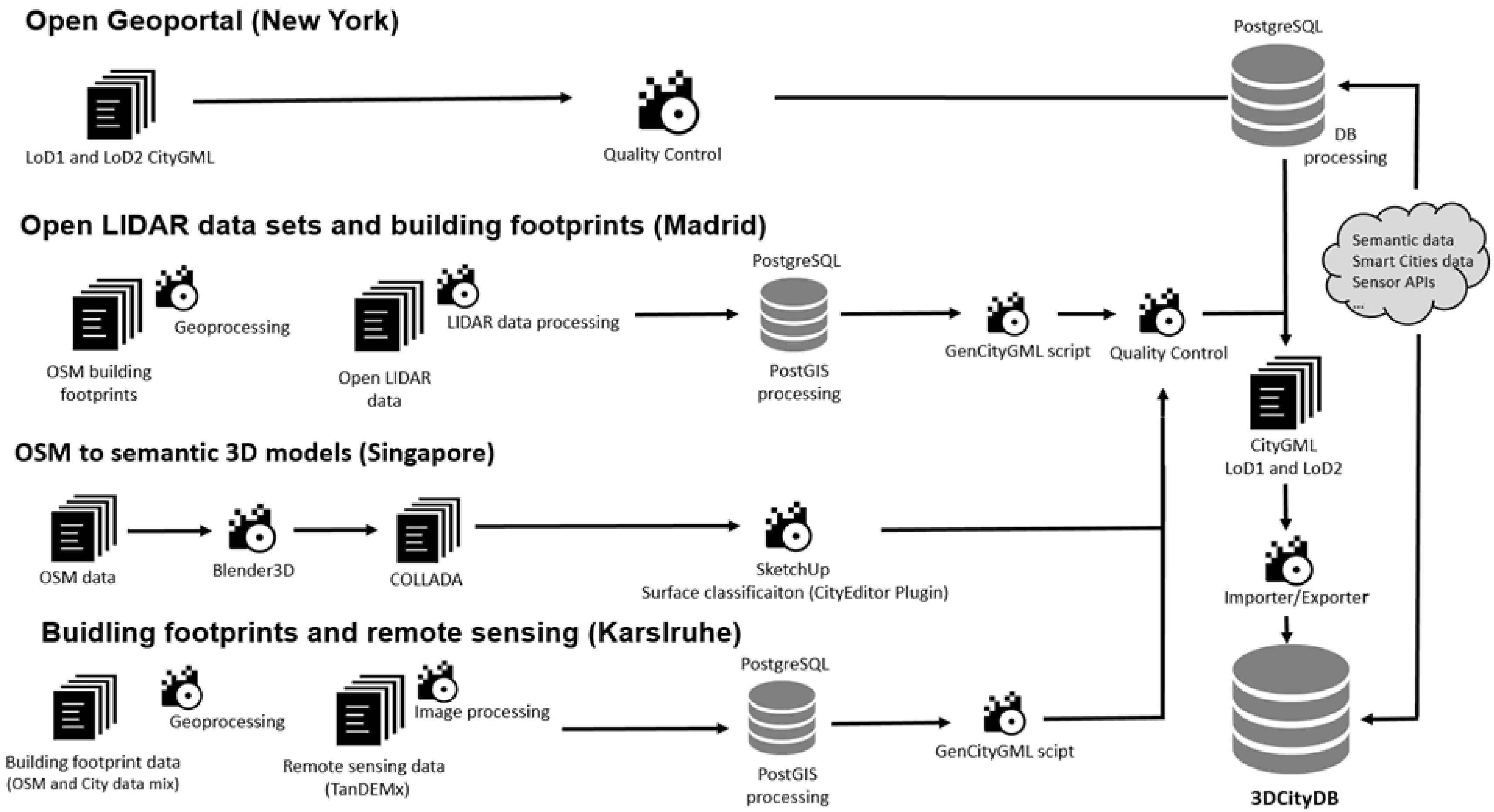


FIGURE A.7: Four methods in which CityGML data was obtained/produced in this thesis, from [231]

Bibliography

[1] Utility Network ADE. *CityGML UtilityNetworkADE - CityGML Wiki*. 2015. URL: https://tinyurl.com/yc24h4ch (Accessed on: 09/19/2019).

[2] Giorgio Agugiaro, Joachim Benner, Piergiorgio Cipriano, and Romain Nouvel. "The Energy Application Domain Extension for CityGML: enhancing interoperability for urban energy simulations". en. In: *Open Geospatial Data, Software and Standards* 3.1 (12/2018). ISSN: 2363-7501. DOI: 10.1186/s40965-018-0042-y. (Accessed on: 12/01/2018).

[3] Usman Ali, Mohammad Haris Shamsi, Cathal Hoare, Eleni Mangina, and James O'Donnell. "A data-driven approach for multi-scale building archetypes development". In: *Energy and Buildings* 202 (2019), p. 109364. ISSN: 0378-7788. DOI: https://doi.org/10.1016/j.enbuild.2019.109364.

[4] McKinsey Analytics. *THE AGE OF ANALYTICS: COMPETING IN A DATA-DRIVEN WORLD*. Executive summary. McKinsey Global Institute, 2016. URL: shorturl.at/vEF49 (Accessed on: 05/22/2021).

[5] Katja Andresen and Norbert Gronau. "An Approach to Increase Adaptability in ERP Systems". In: *Proceedings of the 2005 Information Recources Management Association International Conference*. San Diego, CA, USA: Idea Group Publishing, 2005, pp. 883–885.

[6] Yu Qian Ang, Zachary Michael Berzolla, Samuel Letellier-Duchesne, Violetta Jusiega, and Christoph Reinhart. "UBEM.io: A web-based framework to rapidly generate urban building energy models for carbon reduction technology pathways". In: *Sustainable Cities and Society* 77 (2022), p. 103534. ISSN: 2210-6707. DOI: https://doi.org/10.1016/j.scs.2021.103534.

[7] Kevin Ashton. "That 'Internet of Things' Thing". en. In: (2009), p. 1.

[8] Godfried Augenbroe. *Urban Energy modeling - Fidelity and purpose*. English. 10/2019. (Accessed on: 11/05/2019).

[9] Autodesk. *AutoCAD DevBlog*. 2021. URL: https://adndevblog.typepad.com/autocad/ (Accessed on: 05/24/2021).

[10] Autodesk. *Autodesk Vasari – For Building Performance Analysis – Insight 360*. 2013. URL: https://blogs.autodesk.com/insight/autodesk-vasari-for-building-performance-analysis-1/ (Accessed on: 05/06/2021).

[11] Autodesk. *Ecotect Analysis Discontinuation FAQ | Search | Autodesk Knowledge Network*. en-US. 2015. URL: https://tinyurl.com/5fefv4yr (Accessed on: 05/06/2021).

[12] Ian Ayres, Sophie Raseman, and Alice Shih. "Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage". In: *The Journal of Law, Economics, and Organization* 29.5 (2012), pp. 992–1022. ISSN: 8756-6222. DOI: 10.1093/jleo/ews020.

[13] Peter Barnes. "Index". In: *BIM for Project Managers*. 2020, pp. 179–183. URL: https://doi.org/10.1680/bimpm.65291.bm.

[14] D. J. Bartholomew. "Principal Components Analysis". In: *International Encyclopedia of Education (Third Edition)*. Ed. by Penelope Peterson, Eva Baker, and

Barry McGaw. Third Edition. Oxford: Elsevier, 2010, pp. 374–377. ISBN: 978-0-08-044894-7. DOI: 10.1016/B978-0-08-044894-7.01358-0.

[15] Dominik Bartoszewski, Adam Piorkowski, and Michal Lupa. "The Comparison of Processing Efficiency of Spatial Data for PostGIS and MongoDB Databases". In: *Beyond Databases, Architectures and Structures. Paving the Road to Smart Data Processing and Analysis*. Ed. by Stanis\law Kozielski, Dariusz Mrozek, Pawe\l Kasprowski, Bo\.zena Ma\lysiak-Mrozek, and Daniel Kostrzewa. Cham: Springer International Publishing, 2019, pp. 291–302. ISBN: 978-3-030-19093-4.

[16] Michael Batty. *The new science of cities*. English. 2013th ed. Massachusetts Institute of Technology, 2013. ISBN: 978-0-262-01952-1.

[17] BBC. *Finland makes broadband a 'legal right' - BBC News*. news. 07/2010. URL: https://www.bbc.com/news/10461048 (Accessed on: 11/08/2019).

[18] BDEW, VKU, and GEODE. *Abwicklung von Standardlastprofilen Gas*. Guide. 2016. URL: https://www.bdew.de/media/documents/Leitfaden_20160630_Abwicklung-Standardlastprofile-Gas.pdf (Accessed on: 01/24/2021).

[19] Jeff de la Beaujardiere. *OpenGIS Web Map Server Implementation Specification*. English. 2006. URL: https://portal.ogc.org/files/?artifact_id=14416 (Accessed on: 06/01/2021).

[20] T. Becker, C. Nagel, and T. H. Kolbe. "Integrated 3D Modeling of Multi-utility Networks and Their Interdependencies for Critical Infrastructure Analysis". English. In: *Advances in 3D Geo-Information Sciences, Lecture Notes in Geoinformation and Cartography* (2011). Editors: Kolbe, T. H., König, G. & Nagel, C. Publisher: Springer, Berlin Heidelberg, pp. 1–20. URL: http://mediatum.ub.tum.de/doc/1145740/358854.pdf.

[21] T. Becker, C. Nagel, and T. H. Kolbe. "Semantic 3D Modeling of Multi-Utility Networks in Cities for Analysis and 3D Visualization". English. In: *Progress and New Trends in 3D Geoinformation Sciences, Lecture Notes in Geoinformation and Cartography*. Editors: Pouliot, J., Daniel, S., Hubert, F. & Zamyadi, A. Quebec City: Springer, Berlin Heidelberg, 2012, pp. 41–62. URL: http://mediatum.ub.tum.de/doc/1145740/358854.pdf (Accessed on: 05/06/2021).

[22] Joachim Benner. *CityGML Energy ADE - CityGML Wiki*. 2020. URL: http://www.citygmlwiki.org/index.php/CityGML_Energy_ADE (Accessed on: 09/19/2019).

[23] Joachim Benner. *LandXML – Review with respect to the modelling of Utility Networks*. en. Standardization comitee workshop. Munich, Germany, 2016. URL: https://en.wiki.utilitynetworks.sig3d.org/images/upload/20160913_-_Benner_-_LandXML.pdf (Accessed on: 01/23/2022).

[24] Paul Bertheau, Catherina Cader, and Philipp Blechinger. "Electrification Modelling for Nigeria". en. In: *Energy Procedia* 93 (08/2016), pp. 108–112. ISSN: 18766102. DOI: 10.1016/j.egypro.2016.07.157. (Accessed on: 10/10/2019).

[25] Filip Biljecki, Kavisha Kumar, and Claus Nagel. "CityGML Application Domain Extension (ADE): overview of developments". In: *Open Geospatial Data, Software and Standards* 3.1 (08/2018), p. 13. ISSN: 2363-7501. DOI: 10.1186/s40965-018-0055-6.

[26] Filip Biljecki, Jantien Stoter, Hugo Ledoux, Sisi Zlatanova, and Arzu Çöltekin. "Applications of 3D City Models: State of the Art Review". en. In: *ISPRS International Journal of Geo-Information* 4.4 (12/2015), pp. 2842–2889. ISSN: 2220-9964. DOI: 10.3390/ijgi4042842. (Accessed on: 09/10/2019).

[27] Stefan Bischof, Christoph Martin, Axel Polleres, and Patrik Schneider. "Collecting, Integrating, Enriching and Republishing Open City Data as Linked

Data". In: *The Semantic Web - ISWC 2015*. Ed. by Marcelo Arenas, Oscar Corcho, Elena Simperl, Markus Strohmaier, Mathieu d'Aquin, Kavitha Srinivas, Paul Groth, Michel Dumontier, Jeff Heflin, Krishnaprasad Thirunarayan, and Steffen Staab. Cham: Springer International Publishing, 2015, pp. 57–75. ISBN: 978-3-319-25010-6.

[28] BKI. *BKI Objektdaten Energieeffizientes Bauen E2*. Verlagsgesellschaft Rudolf Müller, 2002. ISBN: 978-3-481-01885-6. URL: https://de.book-info.com/isbn/3-481-01885-1.htm (Accessed on: 03/17/2022).

[29] BKI. *BKI Objektdaten Energieeffizientes Bauen E4*. Baukosteninformationszentrum Deutscher Architektenkammern (BKI), 2011. ISBN: ISBN 978-3-941679-21-4. URL: https://bki.de/objektdaten-energieeffizientes-bauen/bki-objektdaten-energieeffizientes-bauen-e4.html (Accessed on: 03/17/2022).

[30] BKI. *BKI Objektdaten Energieeffizientes Bauen E5*. Baukosteninformationszentrum Deutscher Architektenkammern (BKI), 2014. ISBN: 978-3-941679-56-6. URL: https://bki.de/objektdaten-energieeffizientes-bauen/bki-objektdaten-energieeffizientes-bauen-e5.html (Accessed on: 03/17/2022).

[31] BKI. *Energiekonzept: Planung nach DIN 4108 / DIN 18599 | BKI Objektdaten Energieeffizientes Bauen E3*. Baukosteninformationszentrum Deutscher Architektenkammern (BKI), 2009. URL: https://tinyurl.com/ze6xvvnb (Accessed on: 03/17/2022).

[32] BMUV. *Bericht der Baukosten-senkungskomission*. German. 2015. URL: https://www.die-wohnraumoffensive.de/fileadmin/user_upload/pdf/buendnis-bezahlbares-wohnen-baukostensenkungskommission.pdf (Accessed on: 03/12/2022).

[33] I. Boates, G. Agugiaro, and A. Nichersu. "NETWORK MODELLING AND SEMANTIC 3D CITY MODELS: TESTING THE MATURITY OF THE UTILITY NETWORK ADE FOR CITYGML WITH A WATER NETWORK TEST CASE". en. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* IV-4 (09/2018), pp. 13–20. ISSN: 2194-9050. DOI: 10.5194/isprs-annals-IV-4-13-2018. (Accessed on: 10/12/2018).

[34] Isaac Boates. "Demonstrating Utility Network Interdependency Modelling using the UtilityNetwork Application Domain Extension for CityGML". English. Supervisor: Alexandru Nichersu and Giorgio Agugiaro. MA thesis. Karlsruhe: Hochschule Karlsruhe, 2018.

[35] Andrea Bonaccorsi and Cristina Rossi. "Why Open Source software can succeed". In: *Research Policy* 32.7 (2003), pp. 1243–1258. ISSN: 0048-7333. DOI: https://doi.org/10.1016/S0048-7333(03)00051-9.

[36] Mike Botts, George Percivall, Carl Reed, and John Davidson. "OGC® sensor web enablement: Overview and high level architecture". In: *International conference on GeoSensor Networks*. Springer, 2006, pp. 175–190. DOI: https://doi.org/10.1007/978-3-540-79996-2_10.

[37] Mike Botts and Alexandre Robin. *OpenGIS® Sensor Model Language (SensorML) Implementation Specification. Version 1.0.0.* en. Accepted: 2019-10-29. 2007. DOI: 10.25607/OBP-646. (Accessed on: 11/11/2021).

[38] Mike Botts, Alexandre Robin, and Eric Hirschorn. *OGC SensorML: Model and XML Encoding Standard*. English. Publisher: Open Geospatial Consortium. 08/2020. URL: http://docs.ogc.org/is/12-000r2/12-000r2.html (Accessed on: 11/28/2021).

[39] Roger Brackin and Pedro Gonçalves. *OGC OWS Context Conceptual Model*. English. 2014. URL: https://portal.ogc.org/files/?artifact_id=55182 (Accessed on: 06/01/2021).

[40] M. Breunig, A. Borrmann, E. Rank, S. Hinz, T. Kolbe, M. Schilcher, R.-P. Mundani, J. R. Jubierre, M. Flurl, A. Thomsen, A. Donaubauer, Y. Ji, S. Urban, S. Laun, S. Vilgertshofer, B. Willenborg, M. Menninghaus, H. Steuer, S. Wursthorn, J. Leitloff, M. Al-Doori, and N. Mazroobsemnani. "COLLABORATIVE MULTI-SCALE 3D CITY AND INFRASTRUCTURE MODELING AND SIMULATION". In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-4/W4 (2017), pp. 341–352. DOI: 10.5194/isprs-archives-XLII-4-W4-341-2017.

[41] Encyclopædia Britannica. *energy | Definition, Types, & Examples*. en. 2019. URL: https://www.britannica.com/science/energy (Accessed on: 10/06/2019).

[42] S. Cajot, M. Peter, J.-M. Bahu, F. Guignet, A. Koch, and F. Maréchal. "Obstacles in energy planning at the urban scale". In: *Sustainable Cities and Society* 30 (2017), pp. 223 –236. ISSN: 2210-6707. DOI: https://doi.org/10.1016/j.scs.2017.02.003.

[43] C. F. Calvillo, A. Sánchez-Miralles, and J. Villar. "Energy management and planning in smart cities". In: *Renewable and Sustainable Energy Reviews* 55 (2016), pp. 273–287. ISSN: 1364-0321. DOI: 10.1016/j.rser.2015.10.133. (Accessed on: 09/13/2019).

[44] M Casini. "Green Technology for Smart Cities". en. In: *IOP Conference Series: Earth and Environmental Science* 83 (08/2017), p. 012014. ISSN: 1755-1307, 1755-1315. DOI: 10.1088/1755-1315/83/1/012014. (Accessed on: 11/02/2019).

[45] Juan Antonio Ruiz Ceniceros, José Alfonso Aguilar-Calderón, Roberto Espinosa, and Carolina Tripp-Barba. "The external and data loose coupling for the integration of software units: a systematic mapping study". en. In: *PeerJ Computer Science* 7 (12/2021). Publisher: PeerJ Inc., e796. ISSN: 2376-5992. DOI: 10.7717/peerj-cs.796. (Accessed on: 01/22/2022).

[46] Carlos Cerezo, Julia Sokol, Saud AlKhaled, Christoph Reinhart, Adil Al-Mumin, and Ali Hajiah. "Comparison of four building archetype characterization methods in urban building energy modeling (UBEM): A residential case study in Kuwait City". In: *Energy and Buildings* 154 (2017), pp. 321 –334. ISSN: 0378-7788. DOI: https://doi.org/10.1016/j.enbuild.2017.08.029.

[47] Tania Cerquitelli, Evelina Di Corso, Stefano Proto, Paolo Bethaz, Daniele Mazzarelli, Alfonso Capozzoli, Elena Baralis, Marco Mellia, Silvia Casagrande, and Martina Tamburini. "A Data-Driven Energy Platform: From Energy Performance Certificates to Human-Readable Knowledge through Dynamic High-Resolution Geospatial Maps". en. In: *Electronics* 9.12 (12/2020), p. 2132. ISSN: 2079-9292. DOI: 10.3390/electronics9122132. (Accessed on: 02/07/2021).

[48] Debasish Chakraborty. "Web Based GIS Application using Open Source Software for Sharing Geospatial Data (IJARSG, pp 1124-1228)". In: *international journal of advanced remote sensing and gis* 4 (09/2015), pp. 1224–1228. DOI: 10.23953/cloud.ijarsg.109.

[49] Kam Wing Chan. "Misconceptions and Complexities in the Study of China's Cities: Definitions, Statistics, and Implications". English. In: *Journal of Economic Literature* (2007). URL: https://web.archive.org/web/20130115173048/http://courses.washington.edu/chinageo/ChanCityDefinitionsEGE2007.pdf# (Accessed on: 01/24/2021).

[50] Kanishk Chaturvedi and Thomas H. Kolbe. "InterSensor Service: Establishing Interoperability over Heterogeneous Sensor Observations and Platforms

for Smart Cities". In: *2018 IEEE International Smart Cities Conference (ISC2)*. 09/2018, pp. 1–8. DOI: 10.1109/ISC2.2018.8656984.

[51] Yixing Chen, Tianzhen Hong, Xuan Luo, and Barry Hooper. "Development of city buildings dataset for urban building energy modeling". In: *Energy and Buildings* 183 (2019), pp. 252–265. ISSN: 0378-7788. DOI: https://doi.org/10.1016/j.enbuild.2018.11.008.

[52] Yao-Hsin Chiang, Chih-Yuan Huang, and Masaaki Fuse. "The Integration of OGC SensorThings API and OGC CityGML via Semantic Web Technology". In: *Web and Wireless Geographical Information Systems*. Ed. by Sergio Di Martino, Zhixiang Fang, and Ki-Joune Li. Cham: Springer International Publishing, 2020, pp. 55–67. ISBN: 978-3-030-60952-8.

[53] Ping-Tsai Chung and Sarah H. Chung. "On data integration and data mining for developing business intelligence". In: *2013 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*. 05/2013, pp. 1–6. DOI: 10.1109/LISAT.2013.6578235.

[54] Paul Clarke, Rory V. O'Connor, and Murat Yilmaz. "In Search of the Origins and Enduring Impact of Agile Software Development". In: *Proceedings of the 2018 International Conference on Software and System Process*. ICSSP '18. event-place: Gothenburg, Sweden. New York, NY, USA: Association for Computing Machinery, 2018, pp. 142–146. ISBN: 978-1-4503-6459-1. DOI: 10.1145/3202710.3203162.

[55] Roberta Cocco. *Smart City? Milan is leading the game: here are its 4 secrets*. en. 04/2019. URL: https://www.morningfuture.com/en/article/2019/04/17/smart-city-milano-roberta-cocco/596/ (Accessed on: 11/03/2019).

[56] Javier Conde, Andres Munoz-Arcentales, Alvaro Alonso, Sonsoles Lopez-Pernas, and Joaquin Salvachua. "Modeling Digital Twin Data and Architecture: A Building Guide with FIWARE as Enabling Technology". In: *IEEE Internet Computing* (2021), pp. 1–1. DOI: 10.1109/MIC.2021.3056923.

[57] World Wide Web Consortium. *W3C XML Schema Definition Language (XSD) 1.1 Part 1: Structures*. English. 2012. URL: https://www.w3.org/TR/xmlschema11-1/.

[58] Simon J D Cox. *ISO 19156:2011 - Geographic information – Observations and measurements*. English. 01/2011. DOI: 10.13140/2.1.1142.3042.

[59] Andrew Crapo, Katrina Griffith, Ankesh Khandelwal, John Lizzi, Abha Moitra, and Wang Xiaofeng. "Overcoming Challenges Using the CIM as a Semantic Model for Energy Applications". In: *Proc. Grid-Interop Forum 2010*. 2010. URL: www.pointview.com/data/files/3/2433/1730.pdf.

[60] Drury B. Crawley, Linda K. Lawrie, Frederick C. Winkelmann, W.F. Buhl, Y.Joe Huang, Curtis O. Pedersen, Richard K. Strand, Richard J. Liesen, Daniel E. Fisher, Michael J. Witte, and Jason Glazer. "EnergyPlus: creating a new-generation building energy simulation program". en. In: *Energy and Buildings* 33.4 (04/2001), pp. 319–331. ISSN: 03787788. DOI: 10.1016/S0378-7788(00)00114-6. (Accessed on: 02/25/2020).

[61] Felix Creutzig, Giovanni Baiocchi, Robert Bierkandt, Peter-Paul Pichler, and Karen C. Seto. "Global typology of urban energy use and potentials for an urbanization mitigation wedge". en. In: *Proceedings of the National Academy of Sciences* 112.20 (05/2015), pp. 6283–6288. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1315545112. URL: http://www.pnas.org/lookup/doi/10.1073/pnas.1315545112 (Accessed on: 11/03/2019).

[62] S. Dalla Costa, E. Roccatello, and M. Rumor. "A CITYGML 3D GEODATABASE FOR BUILDINGS' ENERGY EFFICIENCY". en. In: *ISPRS - International Archives*

of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII-4/C21 (08/2011), pp. 19–24. ISSN: 1682-1777. DOI: 10.5194/isprsarchives-XXXVIII-4-C21-19-2011.

[63] Renata Dameri. "Searching for Smart City definition: a comprehensive proposal". In: *INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY* 11.5 (06/2012). DOI: 10.24297/ijct.v11i5.1142. (Accessed on: 10/05/2018).

[64] Fabian Dembski, Uwe Wössner, Mike Letzgus, Michael Ruddat, and Claudia Yamu. "Urban Digital Twins for Smart Cities and Citizens: The Case Study of Herrenberg, Germany". In: *Sustainability* 12.6 (2020). ISSN: 2071-1050. DOI: 10.3390/su12062307.

[65] DesignBuilder. *DesignBuilder Software Ltd - EnergyPlus Simulation*. 2021. URL: https://designbuilder.co.uk/35-support/tutorials/96-designbuilder-online-learning-materials (Accessed on: 05/07/2021).

[66] Lewis Dijkstra and Hugo Poelman. *A harmonised definition of cities and rural areas: the new degree of urbanisation*. en. Working paper. Bruxellles, Belgium: European Comission, 2014, p. 28.

[67] Ahmed Elmaraghy, Hans Voordijk, and Mohamed Marzouk. "An Exploration of BIM and Lean Interaction in Optimizing Demolition Projects". In: ISSN: 2309-0979. 2018, pp. 112–122. URL: https://iglc.net/Papers/Details/1506 (Accessed on: 05/13/2021).

[68] Comission European. *Smart Cities Information System (SCIS)*. English. Knowledge platform. 09/2019. URL: https://smartcities-infosystem.eu/low-carbon-technologies/smart-district (Accessed on: 09/12/2019).

[69] Comission European. *Technology Readiness Levels description*. English. 2014. URL: https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf (Accessed on: 01/03/2021).

[70] European Comission. *Comprehensive study of building energy renovation activities and the uptake of nearly zero-energy buildings in the EU : final report*. English. 2019. URL: https://data.europa.eu/doi/10.2833/14675 (Accessed on: 03/18/2022).

[71] Secretariat-General European Commission. *Communication from the Commission on The European Green Deal*. English. 2019. URL: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2019:640:FIN (Accessed on: 03/12/2022).

[72] Wolfgang Feist. "Energieeffizienz bei Gebäuden - Dargestellt am Beispiel Passivhaus". German. In: Darmstadt: PassivHaus Insitute, 2003, pp. 173–194. URL: https://tinyurl.com/yvpb7k6u (Accessed on: 04/28/2022).

[73] Wolfgang Feist. *Passivhaus Institut (Hrsg.): PHPP 2007: Passivhaus Projektierungs-Paket 2007*. 7. Darmstadt, 2007. URL: https://passiv.de/ (Accessed on: 03/12/2022).

[74] Pablo Fernández, José Miguel Santana, Sebastián Ortega, Agustín Trujillo, José Pablo Suárez, Conrado Domínguez, Jaisiel Santana, and Alejandro Sánchez. "SmartPort: A Platform for Sensor Data Monitoring in a Seaport Based on FI-WARE". In: *Sensors* 16.3 (2016). ISSN: 1424-8220. DOI: 10.3390/s16030417.

[75] Pablo Fernández Moniz, Jaisiel Santana Almeida, Agustín Trujillo Pino, and José Pablo Suárez Rivero. "A GIS-based solution for urban water management". In: *Water International* 45.6 (2020). Publisher: Taylor & Francis, pp. 660–677.

[76] Martina Ferrando and Francesco Causone. "An Overview Of Urban Building Energy Modelling (UBEM) Tools". English. In: *Proceedings of the 16th IBPSA*

International Conference and Exhibition. Vol. 16. Rome: IBPSA, 2019, pp. 3452–3459. DOI: 10.26868/25222708.2019.210632.

[77] Martina Ferrando, Francesco Causone, Tianzhen Hong, and Yixing Chen. "Urban building energy modeling (UBEM) tools: A state-of-the-art review of bottom-up physics-based approaches". In: *Sustainable Cities and Society* 62 (2020), p. 102408. ISSN: 2210-6707. DOI: https://doi.org/10.1016/j.scs.2020.102408.

[78] Giuseppe Fusco and Lerina Aversano. "An approach for semantic integration of heterogeneous data sources". In: *PeerJ Computer Science* 6 (03/2020), e254. ISSN: 2376-5992. DOI: 10.7717/peerj-cs.254.

[79] Aleksandra Gabryjalowicz. "BIM-GIS integration for energy performance studies". English. Supervisor: Alexandru Nichersu. MA thesis. Karlsruhe: KIT, 2016.

[80] A Geiger, A Nichersu, and V Hagenmayer. "Sensitivity of input data in building heating energy demand simulation". In: *BauSIM 2020 – 8th Conference of IBPSA Germany and Austria*. Graz, Austria, 2020, pp. 450–457.

[81] Andreas Geiger, Alexandru Nichersu, Karl-Heinz Häfele, and Veit Hagenmayer. "Usage profile enrichment of CityGML models for urban building energy modeling". In: *Proceedings of IBPSA BauSIM 2022*. Weimar, Germany: IBPSA, 2022.

[82] Christopher Gold. *Spatial context An introduction to fundamental computer algorithms for spatial analysis*. ISPRS Book Series 8. Taylor & Francis Group, 2016. ISBN: 978-1-138-02963-7.

[83] Rachel Gonsenhauser, Katy Hansen, Walker Grimshaw, Jannette Morris, Klaus Albertin, and Megan Mullin. "Digitizing a Statewide Map of Community Water System Service Areas." In: *Journal AWWA - American Water Works Association* 112.10 (2020), pp. 56–61. DOI: https://doi.org/10.1002/awwa.1595. (Accessed on: 12/01/2021).

[84] Michael F. Goodchild. "Scale in GIS: An overview". English (US). In: *Geomorphology* 130.1-2 (07/2011). Publisher: Elsevier, pp. 5–9. ISSN: 0169-555X. DOI: 10.1016/j.geomorph.2010.10.004.

[85] Niklas Grießbaum. "Coupling residential heating systems - Modeling networks of distributed micro co-generation units incorporating electrical heat pumps to satisfy residential heating needs". English. MA thesis. Karlsruhe: KIT, 2012.

[86] Markus Groissböck. "Are open source energy system optimization tools mature enough for serious use?" In: *Renewable and Sustainable Energy Reviews* 102 (2019), pp. 234–248. ISSN: 1364-0321. DOI: https://doi.org/10.1016/j.rser.2018.11.020.

[87] Michael Groves and Klaus Mundt. "Friend or foe? Google Translate in language for academic purposes". In: *English for Specific Purposes* 37 (01/2015), pp. 112–121. ISSN: 0889-4906. DOI: 10.1016/j.esp.2014.09.001.

[88] Marc Großklos. *Kumulierter Energieaufwand und CO2-Emissionsfaktoren verschiedener Energieträger und –versorgungen*. German. 2011. URL: https://www.iwu.de/fileadmin/tools/kea/kea.pdf (Accessed on: 03/13/2022).

[89] G. Gröger and L. Plümer. "How to achieve consistency for 3D city models". In: *GeoInformatica* 15 (2011), pp. 137–165.

[90] Gerhard Gröger and Lutz Plümer. "CityGML – Interoperable semantic 3D city models". en. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 71 (07/2012), pp. 12–33. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2012.04.004. (Accessed on: 10/07/2018).

[91] T. H. K. Gerhard Gröger, Angela Czerwinski, and Claus Nagel. *Candidate OpenGIS ® CityGML Implementation Specification (City Geography Markup Language)*. Published: OGC 06-057r1 version 0.3.0. 07/2006.

[92] Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, and Marimuthu Palaniswami. "Internet of Things (IoT): A vision, architectural elements, and future directions". In: *Future Generation Computer Systems* 29.7 (2013), pp. 1645–1660. ISSN: 0167-739X. DOI: https://doi.org/10.1016/j.future.2013.01.010.

[93] Annie Guerriero, Sylvain Kubicki, Fabrice Berroir, and Clément Lemaire. "BIM-enhanced collaborative smart technologies for LEAN construction processes". In: *2017 International Conference on Engineering, Technology and Innovation (ICE - ITMC)*. 2017, pp. 1023–1030. DOI: 10.1109/ICE.2017.8279994.

[94] R Guglielmetti, D Macumber, and N Long. "OpenStudio: An Open Source Integrated Analysis Platform". In: (2011). URL: https://www.osti.gov/biblio/1032670.

[95] Rowell Hagemann, Ulrich Corsmeier, Christoph Kottmeier, Rayk Rinke, Andreas Wieser, and Bernhard Vogel. "Spatial variability of particle number concentrations and NOx in the Karlsruhe (Germany) area obtained with the mobile laboratory 'AERO-TRAM'". In: *Atmospheric Environment* 94 (2014), pp. 341–352. ISSN: 1352-2310. DOI: https://doi.org/10.1016/j.atmosenv.2014.05.051.

[96] Zia Ul Haq, Gul Faraz Khan, and Tazar Hussain. "A Comprehensive analysis of XML and JSON web technologies". en. In: *New Developments in Circuits, Systems, Signal Processing, Communications and Computers* (2013), p. 8.

[97] Brad Hardin and Dave McCool. *BIM and Construction Management: Proven Tools, Methods, and Workflows*. 2nd Edition. Wiley, 2015. ISBN: 978-1-118-94276-5. URL: https://www.wiley.com/en-us/BIM+and+Construction+Management%3A+Proven+Tools%2C+Methods%2C+and+Workflows%2C+2nd+Edition-p-9781118942765 (Accessed on: 05/18/2021).

[98] Y. Q. He, Y. H. Chen, H. B. He, C. Liu, and X. B. Hou. "Research on the distributed GIS for distribution network based on CIM". In: *IOP Conference Series: Earth and Environmental Science* 186 (10/2018). Publisher: IOP Publishing, p. 012030. DOI: 10.1088/1755-1315/186/4/012030.

[99] Timothy L Hemsath. "CONCEPTUAL ENERGY MODELING FOR ARCHITECTURE, PLANNING AND DESIGN: IMPACT OF USING BUILDING PERFORMANCE SIMULATION INEARLY DESIGN STAGES". English. In: *Proceedings of BS2013*. Chambéry, France: IBPSA, 08/2013, pp. 376–384. URL: http://www.ibpsa.org/proceedings/BS2013/p_2015.pdf.

[100] I. H. Hijazi, T. Krauth, A. Donaubauer, and T. Kolbe. "3DCITYDB4BIM: A SYSTEM ARCHITECTURE FOR LINKING BIM SERVER AND 3D CITYDB FOR BIM-GIS-INTEGRATION". en. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* V-4-2020 (08/2020), pp. 195–202. ISSN: 2194-9050. DOI: 10.5194/isprs-annals-V-4-2020-195-2020. (Accessed on: 05/24/2021).

[101] Jack Hodges, Kimberly García, and Steven Ray. "Semantic Development and Integration of Standards for Adoption and Interoperability". English. In: *Computer* 50.11 (11/2017), pp. 26–36. ISSN: 1558-0814. DOI: 10.1109/MC.2017.4041353.

[102] Jaroslav Hofierka and Marcel Suri. "The solar radiation model for Open source GIS: implementation and applications". In: *Proceedings of the Open source GIS-GRASS users conference*. Vol. 2002. Trento, Italy, 2002, pp. 51–70. URL: https:

//citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.19.9831&rep=rep1&type=pdf (Accessed on: 03/13/2022).

[103] Tianzhen Hong, Yixing Chen, Xuan Luo, Na Luo, and Sang Hoon Lee. "Ten questions on urban building energy modeling". en. In: *Building and Environment* 168 (01/2020), p. 106508. ISSN: 0360-1323. DOI: 10.1016/j.buildenv.2019.106508. (Accessed on: 11/27/2019).

[104] Andreas Huber and Ines Mayer. "#215-Huber_Mayer-Smart Cities: an emerging city concept to frame sustainable transitions?" en. In: The Technical University of Denmark, Copenhagen, Denmark, 08/2012, p. 14. URL: shorturl.at/dBLQT (Accessed on: 10/01/2018).

[105] IEA. *Access to electricity, urban (% of urban population) | Data*. English. Statisical Database. 2018. URL: https://data.worldbank.org/indicator/EG.ELC.ACCS.UR.ZS?end=2017&start=1990 (Accessed on: 10/10/2019).

[106] IEA. *Energy Technology Perspectives 2017*. Tech. rep. Paris: IEA, 2017. URL: https://www.iea.org/reports/energy-technology-perspectives-2017.

[107] Ines Imbert, Marie Sevenet, Patrice Nogues, and Jean-Marie Bahu. "Messung von Energiearmut trotz Datenmangels." In: *Energie und soziale Ungleichheit: Zur gesellschaftlichen Dimension der Energiewende in Deutschland und Europa*. Ed. by Katrin Großmann, André Schaffrin, and Christian Smigiel. Wiesbaden: Springer Fachmedien Wiesbaden, 2017, pp. 263–290. ISBN: 978-3-658-11723-8. DOI: 10.1007/978-3-658-11723-8_10. URL: https://doi.org/10.1007/978-3-658-11723-8_10.

[108] infas. *infas360*. Commercial product, only accessible by paid subscription. 2013. URL: https://www.infas360.de/ (Accessed on: 03/16/2022).

[109] buildingSMART International. *openBIM*. en-GB. 2021. URL: shorturl.at/gwCL2 (Accessed on: 05/20/2021).

[110] ISO. *ISO 16739-1:2018*. en. 2018. URL: https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/07/03/70303.html (Accessed on: 05/20/2021).

[111] ISO. *ISO/AWI 37166 Smart community infrastructures —Specification of multi-source urban data integration for smart city planning(SCP)*. en. 09/2019. URL: http://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/92/69252.html (Accessed on: 09/19/2019).

[112] IWU. *Systematik und Datensätze Deutsche Gebäudetypologie*. German. 2005. URL: https://www.iwu.de/fileadmin/publikationen/gebaeudebestand/2003_IWU_Deutsche-Geb%C3%A4udetypologie-Systematik-und-Datens%C3%A4tze.pdf (Accessed on: 03/12/2022).

[113] Yao Jiacheng, Alexander Simons, and Alexandru Nichersu. *Refurbishment cost model*. English. 2015.

[114] Shaohua Jiang and Zheng Wu. "A BIM-Based Code Checking Approach for Green Construction". en. In: *International Conference on Construction and Real Estate Management 2017*. American Society of Civil Engineers, 11/2017, pp. 156–163. DOI: 10.1061/9780784481066.016. (Accessed on: 05/24/2021).

[115] Eike Juerrens. *SensorEventService < SensorWeb < Wiki*. Wiki page. 2015. URL: https://tinyurl.com/y7whc2rc (Accessed on: 11/28/2021).

[116] Byung Chul Kim, Duhwan Mun, and Soonhung Han. "Retrieval of CAD model data based on Web Services for collaborative product development in a distributed environment". In: *The International Journal of Advanced Manufacturing Technology* 50.9 (10/2010), pp. 1085–1099. ISSN: 1433-3015. DOI: 10.1007/s00170-010-2571-0.

[117] Lawrence A Klein. *Sensor and data fusion: a tool for information assessment and decision making*. English. SPIE press, 2004. ISBN: 978-0-8194-5435-5.

[118] Rosemarie Klemm and Dietrich Klemm. *Gold and Gold Mining in Ancient Egypt and Nubia*. English. Springer, 2013. ISBN: 978-3-642-22508-6. URL: https://tinyurl.com/3hpsey7c (Accessed on: 03/06/2021).

[119] K Klobut, T Mäkeläinen, A Huovila, J Hyvärinen, and J Shemeikka. "KPI framework for energy efficient buildings and neighbourhoods". en. In: *11th European Conference on Product and Process Modelling: eWork and eBusiness in Architecture, Engineering and Construction*. Limassol, Cyprus, 2016, p. 7. URL: https://www.researchgate.net/publication/317872245_KPI_framework_for_energy_efficient_buildings_and_neighbourhoods (Accessed on: 01/04/2021).

[120] Michele Knobel. *Digital Literacies: Concepts, Policies and Practices*. en. Google-Books-ID: doVQq67wWSwC. Peter Lang, 2008. ISBN: 978-1-4331-0169-4.

[121] Andreas Koch. "Continuous Simulation for Urban Energy Planning Based on a Non-Linear Data-Driven Modelling Approach". en. PhD thesis. KIT, 2017.

[122] Thomas Kolbe, Zhihang Yao, Claus Nagel, Felix Kunde, Philipp Willkomm, György Hudra, and Arda Müftüoglu. *3D City Database for CityGML v3*. English. 2015. URL: https://tinyurl.com/2sws5w4u (Accessed on: 09/25/2019).

[123] Smita Kumari and Santanu Kumar Rath. "Performance comparison of SOAP and REST based Web Services for Enterprise Application Integration". In: *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. Kochi, India, 08/2015, pp. 1656–1660. DOI: 10.1109/ICACCI.2015.7275851.

[124] Tatjana Kutzner, Kanishk Chaturvedi, and Thomas H. Kolbe. "CityGML 3.0: New Functions Open Up New Applications". In: *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 88.1 (02/2020), pp. 43–61. ISSN: 2512-2819. DOI: 10.1007/s41064-020-00095-z.

[125] Tatjana Kutzner and Thomas H Kolbe. *CityGML UtilityNetwork ADE – Scope, Concepts, and Applications*. en. New York, NY, USA, 2017. URL: https://123dok.com/document/y9gwlvrq-citygml-utility-network-ade-scope-concepts-applications.html (Accessed on: 04/04/2022).

[126] Tatjana Kutzner and Alexandru Nichersu. *Minutes of the UtilityNetwork ADE Workshop*. English. 03/2017. URL: https://en.wiki.utilitynetworks.sig3d.org/images/upload/Minutes_2017_03_UtilityNetworkADE_Workshop.pdf (Accessed on: 10/06/2019).

[127] Manfred Köhler. "Green facades—a view back and some visions". In: *Urban Ecosystems* 11.4 (2008). Publisher: Springer, pp. 423–436.

[128] Ruben de Laat. *opensourceBIM/BIMserver*. original-date: 2013-05-08. 05/2021. URL: https://tinyurl.com/yjy9heme (Accessed on: 05/24/2021).

[129] Ruben de Laat and Léon van Berlo. "Integration of BIM and GIS: The Development of the CityGML GeoBIM Extension". en. In: *Advances in 3D Geo-Information Sciences*. Ed. by Thomas H. Kolbe, Gerhard König, and Claus Nagel. Lecture Notes in Geoinformation and Cartography. Berlin, Heidelberg: Springer, 2011, pp. 211–225. ISBN: 978-3-642-12670-3. DOI: 10.1007/978-3-642-12670-3_13. (Accessed on: 05/24/2021).

[130] Julien-Samuel Lacroix. *PGRouting for dummies*. Nottingham, 2013. URL: https://www.mapgears.com/downloads/foss4g2013/foss4g2013-pgrouting-for-dummies.pdf (Accessed on: 03/24/2022).

[131] Pablo Landeta, Jorge Vásquez, Xavier Rea, and Iván García-Santillán. "Edition, Publication and Visualization of Geoservices Using Open-Source Tools".

In: *Information and Communication Technologies of Ecuador (TIC.EC)*. Ed. by Miguel Botto-Tobar, Lida Barba-Maggi, Javier González-Huerta, Patricio Villacrés Cevallos, Omar S. Gómez, and María I. Uvidia-Fassler. Cham: Springer International Publishing, 2019, pp. 266–280. ISBN: 978-3-030-02828-2.

[132] Andrew M. St Laurent. *Understanding Open Source and Free Software Licensing: Guide to Navigating Licensing Issues in Existing & New Software*. en. Google-Books-ID: 04jG7TTLujoC. "O'Reilly Media, Inc.", 08/2004. ISBN: 978-0-596-55395-1.

[133] Hugo Ledoux, Ken Arroyo Ohori, Kavisha Kumar, Balázs Dukai, Anna Labetski, and Stelios Vitalis. "CityJSON: a compact and easy-to-use encoding of the CityGML data model". In: *Open Geospatial Data, Software and Standards* 4.1 (06/2019), p. 4. ISSN: 2363-7501. DOI: 10.1186/s40965-019-0064-0.

[134] Philipp Lehsten, Alexander Gladisch, and Djamshid Tavangarian. "Context-Aware Integration of Smart Environments in Legacy Applications". In: *Ambient Intelligence*. Ed. by David V. Keyson, Mary Lou Maher, Norbert Streitz, Adrian Cheok, Juan Carlos Augusto, Reiner Wichert, Gwenn Englebienne, Hamid Aghajan, and Ben J. A. Kröse. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 126–135. ISBN: 978-3-642-25167-2.

[135] Samir Lemeš. "Blockchain-Based Data Integrity for Collaborative CAD". In: *Mixed Reality and Three-Dimensional Computer Graphics*. Ed. by Branislav Sobota and Dragan Cvetković. Section: 7. Rijeka: IntechOpen, 2020. DOI: 10.5772/intechopen.93539.

[136] Jennifer Lenhart, Bas van Vliet, and Arthur P. J. Mol. "New roles for local authorities in a time of climate change: the Rotterdam Energy Approach and Planning as a case of urban symbiosis". In: *Journal of Cleaner Production* 107 (11/2015), pp. 593–601. ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2015.05.026. (Accessed on: 09/13/2019).

[137] Josh Lerner and Jean Tirole. "Some Simple Economics of Open Source". In: *The Journal of Industrial Economics* 50.2 (2002), pp. 197–234. DOI: https://doi.org/10.1111/1467-6451.00174. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-6451.00174.

[138] Frank Leymann. *WSO2Con ASIA 2016: Keynote: At the Heart of Services – Loose Coupling...* English. Keynote presentation. Colombo, 2016. URL: https://www.slideshare.net/wso2.org/wso2con-asia-2016-keynote-at-the-heart-of-services-loose-coupling-and-its-architectural-benefits (Accessed on: 01/23/2022).

[139] Steve Liang, Chih-Yuan Huang, and Tania Khalafbeigi. *OGC SensorThings API Part 1: Sensing, Version 1.0*. en. Accepted: 2019-04-15T19:45:07Z. 2016. DOI: 10.25607/OBP-455. (Accessed on: 11/15/2021).

[140] Chenfeng Liu. "Near Real-time Management of Energy Relevant Sensor Data via Web Services under the SWE Framework". English. Supervisor: Alexandru Nichersu. MA thesis. Karlsruhe: Hochschule Karlsruhe, 2017.

[141] Zhe Liu, Jianmin Bao, and Fei Ding. "An Improved K-Means Clustering Algorithm Based on Semantic Model". In: *Proceedings of the International Conference on Information Technology and Electrical Engineering 2018*. ICITEE '18. event-place: Xiamen, Fujian, China. New York, NY, USA: Association for Computing Machinery, 2018. ISBN: 978-1-4503-6352-5. DOI: 10.1145/3148453.3306269.

[142] Ziwen Liu, Yujie Lu, and Lu Chang Peh. "A Review and Scientometric Analysis of Global Building Information Modeling (BIM) Research in the Architecture, Engineering and Construction (AEC) Industry". In: *Buildings* 9.10 (2019). ISSN: 2075-5309. DOI: 10.3390/buildings9100210.

[143] T. Loga and U. Imkeller-Benjes. *Energiepaß Heizung/Warmwasser - Energetische Qualität von Baukörper und Heizungssystem*. German. 1997. URL: https://www.iwu.de/fileadmin/tools/ephw/1997_IWU_LogaImkeller-Benjes_Energiepass-Heizung-Warmwasser-EPHW.pdf (Accessed on: 03/13/2022).

[144] S. Logothetis, E. Karachaliou, E. Valari, and E. Stylianidis. "OPEN SOURCE CLOUD-BASED TECHNOLOGIES FOR BIM". en. In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-2 (05/2018), pp. 607–614. ISSN: 2194-9034. DOI: 10.5194/isprs-archives-XLII-2-607-2018. (Accessed on: 05/24/2021).

[145] Maa-amet. *Estonian geoportal*. en. 11/2021. URL: https://geoportaal.maaamet.ee/eng/INSPIRE/index.php?lang_id=2&page_id=742 (Accessed on: 11/30/2021).

[146] Azad Madni, Carla Madni, and Scott Lucero. "Leveraging Digital Twin Technology in Model-Based Systems Engineering". en. In: *Systems* 7.1 (01/2019), p. 7. ISSN: 2079-8954. DOI: 10.3390/systems7010007. (Accessed on: 09/17/2019).

[147] Prodromos Makris, Dimitrios J. Vergados, Ioannis Mamounakis, Georgios Tsaousoglou, Konstantinos Steriotis, Nikolaos Efthymiopoulos, and Emmanouel Varvarigos. "A Novel Research Algorithms and Business Intelligence Tool for Progressive Utility's Portfolio Management in Retail Electricity Markets". In: *2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*. Bucharest: IEEE, 2019. ISBN: 978-1-5386-8218-0. DOI: 10.1109/ISGTEurope.2019.8905515.

[148] Avichal Malhotra, Julian Bischof, James Allan, James O'Donnell, Thomas Schwengler, Joachim Benner, and Gerald Schweiger. "City Quarter Information Modelling for Building Energy - A review". English. In: *BauSIM 2020 : 8th Conference of IBPSA Germany and Austria, 23-25 September 2020, Graz University of Technology, Austria : proceedings /*. Graz, Austria: Graz University of Technology, 2020, pp. 543–549. DOI: 10.18154/RWTH-2020-10859. (Accessed on: 01/31/2022).

[149] Avichal Malhotra, Julian Bischof, Alexandru Nichersu, Karl-Heinz Häfele, Johannes Exenberger, Divyanshu Sood, James Allan, Jérôme Frisch, Christoph van Treeck, James O'Donnell, and Gerald Schweiger. "Information modelling for urban building energy simulation—A taxonomic review". In: *Building and Environment* (2021), p. 108552. ISSN: 0360-1323. DOI: https://doi.org/10.1016/j.buildenv.2021.108552.

[150] Vida Maliene, Vytautas Grigonis, Vytautas Palevičius, and Sam Griffiths. "Geographic information system: Old principles with new capabilities". In: *URBAN DESIGN International* 16.1 (01/2011), pp. 1–6. ISSN: 1468-4519. DOI: 10.1057/udi.2010.25.

[151] Prateek Manke, Yogesh K. Garg, and Vinay M. Das. "Energy simulation tools and CAD interoperability: A critical review". In: *2013 International Conference on Energy Efficient Technologies for Sustainability*. 2013, pp. 191–196. DOI: 10.1109/ICEETS.2013.6533381.

[152] Tathiane Martins, Luc Adolphe, and Marion Bonhomme. "Building Energy Demand Based on Urban Morphology Analysis:" en. In: *PLEA2013*. Munich, Germany: TUM, 2013, p. 6.

[153] Priit Martinson. *Estonia – the Digital Republic Secured by Blockchain*. en. Estonia, 2017. URL: https://www.pwc.com/gx/en/services/legal/tech/assets/estonia-the-digital-republic-secured-by-blockchain.pdf (Accessed on: 04/05/2022).

[154] Eoghan McKenna, Ian Richardson, and Murray Thomson. "Smart meter data: Balancing consumer privacy concerns with legitimate applications". In: *Energy Policy* 41 (2012), pp. 807–814. ISSN: 0301-4215. DOI: https://doi.org/10.1016/j.enpol.2011.11.049.

[155] Boisse Meagan. "Ursula Eicker is Concordia's new Canada Excellence Research Chair in Smart, Sustainable and Resilient Communities and Cities". English. In: *Concordia University News* (04/2019). URL: https://www.concordia.ca/content/shared/en/news/stories/2019/04/17/ursula-eicker-is-concordias-new-canada-excellence-research-chair-in-smart-sustainable-and-resilient-communities-and-cities.html (Accessed on: 10/23/2019).

[156] Michael W. Mehaffy, Sergio Porta, and Ombretta Romice. "The "neighborhood unit" on trial: A case study in the impacts of urban morphology". In: *Journal of Urbanism: International Research on Placemaking and Urban Sustainability* 8.2 (2015), pp. 199–217. DOI: https://doi.org/10.1080/17549175.2014.908786. (Accessed on: 01/24/2021).

[157] Nabor C. Mendonça, David Garlan, Bradley Schmerl, and Javier Cámara. "Generality vs. Reusability in Architecture-Based Self-Adaptation: The Case for Self-Adaptive Microservices". In: *Proceedings of the 12th European Conference on Software Architecture: Companion Proceedings*. ECSA '18. event-place: Madrid, Spain. New York, NY, USA: Association for Computing Machinery, 2018. ISBN: 978-1-4503-6483-6. DOI: 10.1145/3241403.3241423.

[158] Christopher D. Michaelis and Daniel P. Ames. "Considerations for Implementing OGC WMS and WFS Specifications in a Desktop GIS". In: *Journal of Geographic Information System* Vol.04No.02 (2012). Type: 10.4236/jgis.2012.42021, p. 7. URL: //www.scirp.org/journal/paperinformation.aspx?paperid=18782.

[159] Ralf Mosshammer, Friederich Kupzog, Mario Faschang, and Matthias Stifter. "Loose coupling architecture for co-simulation of heterogeneous components". In: *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*. ISSN: 1553-572X. 11/2013, pp. 7570–7575. DOI: 10.1109/IECON.2013.6700394.

[160] S.M. Murshed, A. Lindsay, S. Picard, and A. Simons. "PLANTING: Computing High Spatio-temporal Resolutions of Photovoltaic Potential of 3D City Models". In: *The Annual International Conference on Geographic Information Science*. Geospatial Technologies for All. Springer International Publishing, 2018. ISBN: 978-3-319-78207-2. URL: https://doi.org/10.1007/978-3-319-78208-9.

[161] Gunnar Myhre, D Shindell, FM Bréon, W Collins, J Fuglestvedt, J Huang, D Koch, JF Lamarque, D Lee, B Mendoza, and others. *Anthropogenic and natural radiative forcing, climate change 2013: the Physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 659–740*. 2013.

[162] Renáta Máchová, Miloslav Hub, and Martin Lnenicka. "Usability evaluation of open data portals". In: *Aslib Journal of Information Management* 70.3 (01/2018). Publisher: Emerald Publishing Limited, pp. 252–268. ISSN: 2050-3806. DOI: 10.1108/AJIM-02-2018-0026. (Accessed on: 05/23/2021).

[163] Claus Nagel. *wfs – 3DCityDB Database*. 2021. URL: https://www.3dcitydb.org/3dcitydb/wfs/ (Accessed on: 10/20/2021).

[164] P. Nageler, A. Koch, F. Mauthner, I. Leusbrock, T. Mach, C. Hochenauer, and R. Heimrath. "Comparison of dynamic urban building energy models (UBEM): Sigmoid energy signature and physical modelling approach". In: *Energy and Buildings* 179 (11/2018), pp. 333–343. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2018.09.034. (Accessed on: 12/01/2018).

[165] Nawari O. Nawari and Michael Kuenstle. *Building Information Modeling Framework for Structural Design*. CRC Press, 2015. ISBN: 978-1-4822-4043-6.

[166] Alexandru Nichersu and Alexander Simons. "Building a CityGML Infrastructure for Energy Related Simulations". English. In: *Geographic Information Science 2014 Proceedings*. Wien, Austria: Springer, 09/2014, p. 4.

[167] Alexandru Nichersu and Jochen Wendel. "Open-source solutions for 3D geodata visualization web services". English, German. In: *Geomatik aktuell 2016: Seminar zum Thema: Geodaten in der Cloud.* Karlsruher Geowissenschaftliche Schriftenreihe B: Vermessungswesen und Photogrammetrie. Hochschule Karlsruhe, 2016, p. 76. ISBN: 978-3-89063-108-0. URL: https://silo.tips/download/geomatik-aktuell-2016 (Accessed on: 10/12/2018).

[168] Iulian Nichersu, Iuliana Nichersu, Eugenia Marin, Florentina Sela, Marian Mierla, and Cristian Trifanov. "Expert Judgement Assessment & SCENT Ontological Analysis". English. In: Tulcea, Romania: DDNI, 05/2017. ISBN: 2344-3766. URL: http://ddni.ro/wps/wp-content/uploads/2017/05/DELTAS-AND-WETLANDS_vol_4_2017.pdf (Accessed on: 11/05/2019).

[169] Hanna Niemi-Hugaerts. *About the Projects – CitySDK*. en-US. Forum. 11/2019. URL: //www.citysdk.eu/about-the-project-2/ (Accessed on: 11/05/2019).

[170] Francesca Noardo, Ken Arroyo Ohori, Filip Biljecki, Claire Ellul, Lars Harrie, Thomas Krijnen, Helen Eriksson, Jordi van Liempt, Maria Pla, Antonio Ruiz, Dean Hintz, Nina Krueger, Cristina Leoni, Leire Leoz, Diana Moraru, Stelios Vitalis, Philipp Willkomm, and Jantien Stoter. "Reference study of CityGML software support: The GeoBIM benchmark 2019—Part II". In: *Transactions in GIS* 25.2 (2021), pp. 842–868. DOI: https://doi.org/10.1111/tgis.12710.

[171] Romain Nouvel, Maryam Zirak, Volker Coors, and Ursula Eicker. "The influence of data quality on urban heating demand modeling using 3D city models". In: *Computers, Environment and Urban Systems* 64 (2017), pp. 68–80. ISSN: 0198-9715. DOI: https://doi.org/10.1016/j.compenvurbsys.2016.12.005.

[172] Nurzhan Nurseitov, Michael Paulson, Randall Reynolds, and Clemente Izurieta. "Comparison of JSON and XML Data Interchange Formats: A Case Study". en. In: *Proceedings of the ISCA 22nd International Conference on Computer Applications in Industry and Engineering*. San Francisco, California, USA, 2009, pp. 157–162.

[173] Daniel Nüst and Thomas Bartoschek. "Open Environmental Data Analysis (EGU 2018 PICO)". In: Vienna, Session ESSI3.5: Zenodo, 04/2018. DOI: 10.5281/zenodo.1217912.

[174] OECD. "Case Study 8: Estonia e-government and the creation of a comprehensive data infrastructure for public services and agriculture policies implementation". en. In: *Digital Opportunities for Better Agricultural Policies*. OECD Publishing, Paris, 09/2019, pp. 207–213. ISBN: 978-92-64-45984-7 978-92-64-88785-5 978-92-64-64129-7 978-92-64-56341-4. DOI: 10.1787/510a82b5-en. (Accessed on: 11/30/2021).

[175] OECD. *The Metropolitan Century: Understanding Urbanisation and its Consequences*. OECD Publishing, Paris, 2015. DOI: https://doi.org/10.1787/9789264228733-en. (Accessed on: 03/12/2022).

[176] OGC. *OGC Web Coverage Service (WCS) 2.1 Interface Standard - Core*. English. 2018. URL: http://docs.opengeospatial.org/is/17-089r1/17-089r1.html (Accessed on: 06/01/2021).

[177] Adegboyega Ojo, Edward Curry, and Fatemeh Ahmadi Zeleti. "A Tale of Open Data Innovations in Five Smart Cities". In: *2015 48th Hawaii International Conference on System Sciences*. ISSN: 1530-1605. 01/2015, pp. 2326–2335. DOI: 10.1109/HICSS.2015.280.

[178] Léon L. Olde Scholtenhuis, Timo Hartmann, and André G. Dorée. "Characterizing coordination in both loose and very tightly coupled utility reconstruction processes". In: *roceedings of International Construction Specialty Conference of the Canadian Society for Civil Engineering (ICSC) 5th edition : 2015*. Vancouver: University of British Columbia Library, 06/2015. DOI: 10.14288/1.0076361.

[179] OpenStreetMap contributors. *Planet dump retrieved from https://planet.osm.org*. Published: https://www.openstreetmap.org. 2017.

[180] TNS opinion. *PUBLIC OPINION IN THE EUROPEAN UNION*. English. Eurobarometer 83. European Comission, 07/2015. (Accessed on: 09/18/2019).

[181] Jukka Paatero. "COMPUTATIONAL STUDIES ON VARIABLE DISTRIBUTED ENERGY SYSTEMS". English. Doctoral Dissertation. Helsinki: Helsinki University of Technology, 2009. URL: https://core.ac.uk/download/pdf/301128713.pdf (Accessed on: 01/24/2021).

[182] Neal R. Peirce, W. Johnson Curtis, and Farley M. Peters. *Century of the City*. English. The Rockefeller Foundation, 12/2006. URL: https://tinyurl.com/mu9vz6yx (Accessed on: 09/12/2019).

[183] Christopher Pettit, Ashley Bakelmun, Scott N. Lieske, Stephen Glackin, Karlson 'Charlie' Hargroves, Giles Thomson, Heather Shearer, Hussein Dia, and Peter Newman. "Planning support systems for smart cities". In: *City, Culture and Society* 12 (2018), pp. 13–24. ISSN: 1877-9166. DOI: https://doi.org/10.1016/j.ccs.2017.10.002.

[184] Felipe Pezoa, Juan L. Reutter, Fernando Suarez, Martín Ugarte, and Domagoj Vrgoč. "Foundations of JSON Schema". In: *Proceedings of the 25th International Conference on World Wide Web*. WWW '16. event-place: Montréal, Québec, Canada. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, 2016, pp. 263–273. ISBN: 978-1-4503-4143-1. DOI: 10.1145/2872427.2883029.

[185] Stefan Pfenninger, Lion Hirth, Ingmar Schlecht, Eva Schmid, Frauke Wiese, Tom Brown, Chris Davis, Matthew Gidden, Heidi Heinrichs, Clara Heuberger, Simon Hilpert, Uwe Krien, Carsten Matke, Arjuna Nebel, Robbie Morrison, Berit Müller, Guido Pleßmann, Matthias Reeg, Jörn C. Richstein, Abhishek Shivakumar, Iain Staffell, Tim Tröndle, and Clemens Wingenbach. "Opening the black box of energy modelling: Strategies and lessons learned". In: *Energy Strategy Reviews* 19 (2018), pp. 63–71. ISSN: 2211-467X. DOI: https://doi.org/10.1016/j.esr.2017.12.002.

[186] I. Pispidikis and E. Dimopoulou. "CITYGML RESTFUL WEB SERVICE: AUTOMATIC RETRIEVAL OF CITYGML DATA BASED ON THEIR SEMANTICS. PRINCIPLES, GUIDELINES AND BLDG CONCEPTUAL DESIGN". en. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information*

Sciences IV-4/W6 (09/2018), pp. 49–56. ISSN: 2194-9050. DOI: 10.5194/isprs-annals-IV-4-W6-49-2018.

[187] Open Power System Data Project. *Data - wiki.openmod-initiative.org*. 2020. URL: https://tinyurl.com/39k2cnha (Accessed on: 02/22/2021).

[188] Erika Pärn. *Dr Erika Pärn on Digital Twinning [Video] | It works, Digital, University of Cambridge*. en. 2020. URL: https://www.pinterest.co.uk/pin/831758624923697389/ (Accessed on: 05/07/2021).

[189] M. Quitzau, B. Hoffmann, Jens-Phillip Petersen, A. Heller, B. Olesen, C. Rode, Vickie Aagesen, O. Mørck, Uta Lynar, L. Ebenbeck, Judith Commenges, J. Freudenberg, P. Vieweg, A. Koch, M. Peter, M. Sevenet, S. Cajot, M. Kornmann, G. Stryi-Hipp, D. Kellenberger, Christian Schmid, A. Roser, F. Banon, Kenneth Ward Church, I. Andresen, Brita Nielsen, Eszter Juhasz-Nagy, Y. Shimoda, Rita Streblow, J. Schiefelbein, Ruth Buggie, Declan Meally, H. Strasser, O. M. A. Tinkhof, \AA L. Sørensen, C. Slotterback, and J. Kimman. "Implementation of Energy Strategies in Communities (Annex 63) Volume 3: Application of Strategic Measures". In: 2018.

[190] Marvin Rees, Yvonne Aki-Sawyerr, and Giuseppe Sala. *When Cities Go Global*. 07/2019. URL: https://www.youtube.com/watch?v=bOU00-wjdCc (Accessed on: 11/03/2019).

[191] Christoph F. Reinhart and Carlos Cerezo Davila. "Urban building energy modeling – A review of a nascent field". en. In: *Building and Environment* 97 (02/2016), pp. 196–202. ISSN: 03601323. DOI: 10.1016/j.buildenv.2015.12.001.

[192] Jan Remund, Stefan Müller, Michael Schmutz, and Pascal Graf. "Meteonorm Version 8". In: *Proceedings of the EUPVSEC 2020*. Online conference, 2020. URL: https://meteonorm.com/assets/publications/5BV.3.8_pvsec_2020_mn8.pdf.

[193] Hannah Ritchie and Max Roser. "Urbanization". English. In: *Our World in Data* (2018). URL: https://ourworldindata.org/urbanization#citation (Accessed on: 02/03/2022).

[194] Julian F. Rosser, Gavin Long, Sameh Zakhary, Doreen S. Boyd, Yong Mao, and Darren Robinson. "Modelling Urban Housing Stocks for Building Energy Simulation using CityGML EnergyADE". en. In: *ISPRS International Journal of Geo-Information* 8.4 (03/2019). ISSN: 2220-9964. DOI: 10.3390/ijgi8040163. (Accessed on: 09/13/2019).

[195] L. Rowe and M. Stonebraker. "The POSTGRES Data Model". In: *Research Foundations in Object-Oriented and Semantic Database Systems*. 1987.

[196] Patrizia Russo, Rosa Lanzilotti, Maria F. Costabile, and Christopher J. Pettit. "Towards satisfying practitioners in using Planning Support Systems". In: *Computers, Environment and Urban Systems* 67 (2018), pp. 9–20. ISSN: 0198-9715. DOI: https://doi.org/10.1016/j.compenvurbsys.2017.08.009.

[197] Ted Saarikko, Ulrika H. Westergren, and Tomas Blomquist. "Digital transformation: Five recommendations for the digitally conscious firm". In: *Business Horizons* 63.6 (2020), pp. 825–839. ISSN: 0007-6813. DOI: https://doi.org/10.1016/j.bushor.2020.07.005.

[198] Jeffrey D. Sachs. *The Age of Sustainable Development | Google Preview*. en. 01/2015. ISBN: 978-0-231-17314-8. URL: https://cup.columbia.edu/the-age-of-sustainable-development/9780231173155 (Accessed on: 10/15/2018).

[199] Muhammad Saed. "Heating Energy Consumption and CO2 Production in the city of Karlsruhe, A novel clustering approach". English. Supervisor: Dr.

Jochen Wendel and Alexandru Nichersu M.Sc. MA thesis. Darmstadt: Technische Universität Darmstadt, 2015.

[200] Muhammad Saed and Jochen Wendel. "Estimating heating energy consumption and CO2 production – A novel modeling approach". en. In: (2015), p. 1.

[201] Peter Salhofer. "Evaluating the FIWARE Platform". English. In: *Proceedings of the 51st Hawaii International Conference on System Sciences | 2018*. Hawaii, USA, 2018, pp. 5797–5805. ISBN: 978-0-9981331-1-9. DOI: 10.24251/HICSS.2018.726. (Accessed on: 11/27/2021).

[202] Bahador Samadzadegan, Soroush Samareh Abolhassani, Sanam Dabirian, Saeed Ranjbar, Hadise Rasoulian, Azin Sanei, and Ursula Eicker. "Novel Energy System Design Workflow for Zero-Carbon Energy District Development". In: *Frontiers in Sustainable Cities* 3 (2021). ISSN: 2624-9634. DOI: 10.3389/frsc.2021.662822.

[203] Nuno Santos, Jaime Pereira, Nuno Ferreira, and Ricardo J. Machado. "Modeling in Agile Software Development: Decomposing Use Cases Towards Logical Architecture Design". In: *Product-Focused Software Process Improvement*. Ed. by Marco Kuhrmann, Kurt Schneider, Dietmar Pfahl, Sousuke Amasaki, Marcus Ciolkowski, Regina Hebig, Paolo Tell, Jil Klünder, and Steffen Küpper. Vol. 11271. Springer International Publishing, Cham, 2018, pp. 396–408. ISBN: 978-3-030-03673-7. DOI: https://doi.org/10.1007/978-3-030-03673-7_31.

[204] I. M. Savin and S. A. Sinenko. "Experience in digitizing the services of public authorities in the transport sector". In: *IOP Conference Series: Materials Science and Engineering* 918.1 (09/2020). Publisher: IOP Publishing, p. 012181. DOI: 10.1088/1757-899x/918/1/012181.

[205] ScaleGrid. *2019 Database Trends - SQL vs. NoSQL, Top Databases, Single vs. Multiple Database Use*. English. 09/2019. URL: https://scalegrid.io/blog/2019-database-trends-sql-vs-nosql-top-databases-single-vs-multiple-database-use/ (Accessed on: 09/23/2019).

[206] N. Schuurman. *GIS: A Short Introduction*. 2004.

[207] Monika Sester and Claus Brenner. "A vocabulary for a multiscale process description for fast transmission and continuous visualization of spatial data". In: *Computers & Geosciences* 35 (2009), pp. 2177–2184. DOI: 10.1016/j.cageo.2008.11.003.

[208] Jonny Sexton. "Implementation of the EnergyADE 0.8 DB Schema within the aEneAs Python Library - Use Case Heat Loss and Solar Radiation". English. Supervisor: Alexandru Nichersu, MSc. and Alexander Simons, MSc. MA thesis. Karlsruhe: KIT, 2016.

[209] Esther Shein. "Python for beginners". In: *Communications of the ACM* 58.3 (2015). Publisher: ACM New York, NY, USA, pp. 19–21. DOI: 10.1145/2716560.

[210] Alexander Simons. "Development of a CityGML infrastructure for the implementation of an energy demand method with different data sources". English. Supervisor: Alexandru Nichersu. MA thesis. Karlsruhe: Hochschule Karlsruhe, 2014.

[211] SmartBear. *SOAP vs REST APIs: Which Is Right For You? | SoapUI*. IT support. 12/2021. URL: https://www.soapui.org/learn/api/soap-vs-rest-api/ (Accessed on: 12/11/2021).

[212] H. Strasser, J. Kimman, A. Koch, O. Mair am Tinkhof, D. Müller, J. Schiefelbein, and C. Slotterback. "IEA EBC annex 63–implementation of energy strategies in communities". In: *Energy and Buildings* 158 (2018), pp. 123–134. ISSN: 0378-7788. DOI: https://doi.org/10.1016/j.enbuild.2017.08.051.

[213] Harald Sundmaeker, Patrick Guillemin, Peter Friess, Sylvie Woelfflé, European Commission, and Directorate-General for the Information Society and Media. *Vision and challenges for realising the Internet of things*. en. OCLC: 847355368. Luxembourg: EUR-OP, 2010. ISBN: 978-92-79-15088-3.

[214] Irina Susorova, Melissa Angulo, Payam Bahrami, and Brent Stephens. "A model of vegetated exterior facades for evaluation of wall thermal performance". In: *Building and Environment* 67 (2013), pp. 1–13. ISSN: 0360-1323. DOI: https://doi.org/10.1016/j.buildenv.2013.04.027. URL: https://www.sciencedirect.com/science/article/pii/S0360132313001388.

[215] VC Systems. *VC Database*. en-US. Software. 2021. URL: https://vc.systems/en/products/vc-database/ (Accessed on: 10/20/2021).

[216] Daniel Szepielak, Przemyslaw Tumidajewicz, and Lars Hagge. "Integrating Information Systems Using Web Oriented Integration Architecture and RESTful Web Services". In: *2010 6th World Congress on Services*. 2010, pp. 598–605. DOI: 10.1109/SERVICES.2010.115.

[217] W. Tobler. "Movie Simulating Urban Growth in the Detroit Region". In: *Economic Geography* 46 (1970), pp. 234–240. URL: https://www.jstor.org/stable/143141?origin=JSTOR-pdf&seq=1 (Accessed on: 01/24/2021).

[218] Erik H. Trainer, Arun Kalyanasundaram, Chalalai Chaihirunkarn, and James D. Herbsleb. "How to Hackathon: Socio-Technical Tradeoffs in Brief, Intensive Collocation". In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '16. event-place: San Francisco, California, USA. New York, NY, USA: Association for Computing Machinery, 2016, pp. 1118–1130. ISBN: 978-1-4503-3592-8. DOI: 10.1145/2818048.2819946.

[219] A. Tserstou, A. Jonoski, I. Popescu, T. Herman Asumpcao, G. Athanasiou, A. Kallioras, and I. Nichersu. "SCENT: Citizen Sourced Data in Support of Environmental Monitoring". In: *2017 21st International Conference on Control Systems and Computer Science (CSCS)*. ISSN: 2379-0482. 05/2017, pp. 612–616. DOI: 10.1109/CSCS.2017.93.

[220] UNCTAD. *Beyond 20/20 WDS - Table view - Total and urban population, annual*. English, French. Statistical Database. 2019. URL: https://unctadstat.unctad.org/wds/TableViewer/tableView.aspx (Accessed on: 11/01/2019).

[221] Union européenne and Commission européenne. *EU energy in figures 2018*. en. OCLC: 1079865467. 2018. ISBN: 978-92-79-88737-6.

[222] Department of Economic and Social Affairs United Nations Population Division. *The World's Cities in 2016 – Data Booklet*. United Nations, 2016. URL: https://www.un.org/en/development/desa/population/publications/pdf/urbanization/the_worlds_cities_in_2016_data_booklet.pdf (Accessed on: 09/28/2019).

[223] Trang Vo. "DEVELOPMENT AND VALIDATION OF A GREEN ROOF / GREEN FAÇADE MODEL FOR ESTIMATING THEIR IMPACT ON BUILDINGS ENERGY DEMAND". English. Supervisor: Alexandru Nichersu M.Sc. and Dr. Jochen Wendel. MA thesis. Darmstadt: Technische Universität Darmstadt, 2018.

[224] Trang Thuy Vo, Alexandru Nichersu, and Jochen Wendel. "Modeling, Monitoring, and Validating Green Roof and Green Facade Solutions with Semantic City Models Using Low Cost Sensors and Open Software Infrastructures". In: *Urban Science* 3.2 (2019). ISSN: 2413-8851. DOI: 10.3390/urbansci3020039.

[225] Panagiotis (Peter) A. Vretanos. *OGC® Web Feature Service 2.0 Interface Standard – With Corrigendum*. English. 2014. URL: http://docs.opengeospatial.org/is/09-025r2/09-025r2.html (Accessed on: 05/31/2021).

[226] W3C. *Data on the Web Best Practices*. en. 2021. URL: https://www.w3.org/TR/dwbp/ (Accessed on: 05/23/2021).

[227] Chao Wang and Yong K. Cho. "Automated gbXML-Based Building Model Creation for Thermal Building Simulation". In: *2014 2nd International Conference on 3D Vision*. Vol. 2. 2014, pp. 111–117. DOI: 10.1109/3DV.2014.109.

[228] Cheng-Kai Wang, Simon Tindemans, Clayton Miller, Giorgio Agugiaro, and Jantien Stoter. "Bayesian calibration at the urban scale: a case study on a large residential heating demand application in Amsterdam". In: *Journal of Building Performance Simulation* 13.3 (2020). Publisher: Taylor & Francis, pp. 347–361. DOI: 10.1080/19401493.2020.1729862.

[229] O. Wannenwetsch and T. A. Majchrzak. "On constructing persistent identifiers with persistent resolution targets". In: *2016 Federated Conference on Computer Science and Information Systems (FedCSIS)*. 09/2016, pp. 1031–1040.

[230] Gregory J. Ward. "The RADIANCE Lighting Simulation and Rendering System". In: *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '94. New York, NY, USA: Association for Computing Machinery, 1994, pp. 459–472. ISBN: 0-89791-667-0. DOI: 10.1145/192161.192286.

[231] Jochen Wendel, Alexander Simons, Alexandru Nichersu, and Syed Monjur Murshed. "Rapid development of semantic 3D city models for urban energy analysis based on free and open data sources and software". en. In: *Proceedings of the 3rd ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics - UrbanGIS'17*. Redondo Beach, CA, USA: ACM Press, 2017, pp. 1–7. ISBN: 978-1-4503-5495-0. DOI: 10.1145/3152178.3152193. (Accessed on: 10/08/2018).

[232] E. Widl, C. Agugiaro, and P. Puerto. "FIRST STEPS TOWARDS LINKING SEMANTIC 3D CITY MODELLING AND MULTI-DOMAIN CO-SIMULATION FOR ENERGY MODELLING AT URBAN SCALE". en. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* IV-4 (09/2018), pp. 227–234. ISSN: 2194-9050. DOI: 10.5194/isprs-annals-IV-4-227-2018. URL: https://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/IV-4/227/2018/ (Accessed on: 10/01/2018).

[233] Manfred Wieland. "Computing Solar Radiation on Buildings at an Urban Scale". English. Supervisor: Alexandru Nichersu. MA thesis. Karlsruhe: Hoch schule Karlsruhe, 2014.

[234] Manfred Wieland, A. Nichersu, S.M. Murshed, and Jochen Wendel. "Computing Solar Radiation on CityGML Building Data". In: *Proceedings of the 18th AGILE conference on Geographic Information Science.* Issue: June. 2015, pp. 2–5. URL: https://agile-online.org/conference_paper/cds/agile_2015/shortpapers/107/107_Paper_in_PDF.pdf.

[235] Mark D Wilkinson, Michel Dumontier, I Jsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, Jildau Bouwman, Anthony J Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J G Gray, Paul Groth, Carole Goble, Jeffrey S Grethe, Jaap Heringa, Peter A C 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J Lusher, Maryann E Martone, Albert Mons, Abel L Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry

Sengstag, Ted Slater, George Strawn, Morris A Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons. "The FAIR Guiding Principles for scientific data management and stewardship". eng. In: *Scientific data* 3 (03/2016). Publisher: Nature Publishing Group, pp. 160018–160018. ISSN: 2052-4463. DOI: 10.1038/sdata.2016.18.

[236] Lixian Wu, Yujie Lin, and Weilin Pang. "Distribution Network Topology Modelling and Automatic Mapping Based on CIM and GIS". In: *2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC)*. 12/2018, pp. 1–5. DOI: 10.1109/ITOEC.2018.8740469.

[237] Shinji Yamamura, Liyang Fan, and Yoshiyasu Suzuki. "Assessment of Urban Energy Performance through Integration of BIM and GIS for Smart City Planning". In: *Procedia Engineering* 180 (2017), pp. 1462–1472. ISSN: 1877-7058. DOI: https://doi.org/10.1016/j.proeng.2017.04.309.

[238] Zhihang Yao, Claus Nagel, Felix Kunde, György Hudra, Philipp Willkomm, Andreas Donaubauer, Thomas Adolphi, and Thomas H. Kolbe. "3DCityDB - a 3D geodatabase solution for the management, analysis, and visualization of semantic 3D city models based on CityGML". en. In: *Open Geospatial Data, Software and Standards* 3.1 (12/2018). ISSN: 2363-7501. DOI: 10.1186/s40965-018-0046-7.

[239] Di Zhang, Zhenyu Zhou, Shahid Mumtaz, Jonathan Rodriguez, and Takuro Sato. "One Integrated Energy Efficiency Proposal for 5G IoT Communications". In: *IEEE Internet of Things Journal* 3.6 (12/2016), pp. 1346–1354. ISSN: 2327-4662, 2372-2541. DOI: 10.1109/JIOT.2016.2599852.

[240] Qing Zhao, Hoh In, Xiaowei Wu, and Garng Huang. "Transforming legacy energy management system (EMS) modules into reusable components: a case study". In: *Proceedings 24th Annual International Computer Software and Applications Conference. COMPSAC2000*. ISSN: 0730-3157. 10/2000, pp. 105–110. DOI: 10.1109/CMPSAC.2000.884698.

[241] Junlin Zhou and Yan Fu. "Clustering High-Dimensional Data Using Growing SOM". In: *Proceedings of the Second International Conference on Advances in Neural Networks - Volume Part II*. ISNN'05. event-place: Chongqing, China. Berlin, Heidelberg: Springer-Verlag, 2005, pp. 63–68. ISBN: 3-540-25913-9.

[242] Wanji Zhu, Alexander Simons, Sven Wursthorn, and Alexandru Nichersu. "Integration of CityGML and Air Quality Spatio-Temporal Data Series via OGC SOS". en. In: *Geospatial Sensor Webs*. Münster, Germany: 52°North Initiative for Geospatial Open Source Software GmbH, 08/2016, p. 6. URL: http://ceur-ws.org/Vol-1762/Zhu.pdf (Accessed on: 09/25/2019).

[243] Dimitrios Zissis, Dimitrios Lekkas, Philip Azariadis, Paraskevas Papanikos, and Elias Xidias. "Collaborative CAD/CAE as a cloud service". In: *International Journal of Systems Science: Operations & Logistics* 4.4 (2017). Publisher: Taylor & Francis, pp. 339–355. DOI: 10.1080/23302674.2016.1186237.

[244] Marcel Šúri and Jaroslav Hofierka. "A New GIS-based Solar Radiation Model and Its Application to Photovoltaic Assessments". In: *Transactions in GIS* 8.2 (2004), pp. 175–190. DOI: https://doi.org/10.1111/j.1467-9671.2004.00174.x.