READY4SmartCities - ICT Roadmap and Data Interoperability for Energy Systems in Smart Cities

Deliverable D2.3: Ontologies and datasets for Energy Management System interoperability v2

Document Details

Delivery date:	*M23*
Lead Beneficiary:	*AEC3 Ltd.*
Dissemination Level (*):	*PU*
Version:	*2.0*
Preparation Date:	*31/08/2015*
Prepared by:	*Matthias Weise (AEC3), María Poveda-Villalón (UPM), Raúl García-Castro (UPM), Jérôme Euzenat (INRIA), Luz-Maria Priego-Roche (INRIA), Bruno Fies (CSTB), Andrea Cavallaro (DAPP), Jan Peters-Anders (AIT), Kleopatra Zoi (CERTH/ITI)*
Reviewed by:	*Jan Peters-Anders (AIT), Bruno Fies (CSTB)*
Approved by:	*Coordinator (DAPP), Technical Coordinator (UPM)*

(*) Only one choice between:

- PU = Public
- PP = Restricted to other programme participants (including the Commission Services)
- RE = Restricted to a group specified by the consortium (including the Commission Services)
- CO = Confidential, only for members of the consortium (including the Commission Services)

Project Contractual Details

Project Title:	ICT Roadmap and Data Interoperability for Energy Systems in Smart Cities
Project Acronym:	READY4SmartCities
Grant Agreement No.:	608711
Project Start Date:	2013-10-01
Project End Date:	2015-09-30
Duration:	24 months
Project Officer:	Svetoslav Mihaylov

SEVENTH FRAMEWORK PROGRAMME

Revision History

Date	Author	Partner	Content	Ver.
06/10/2014	Mari Figueroa, Matthias Weise, Andrea Cavallaro	UPM AEC3 DAPP	Submission of V1 (D2.2)	1.0
17/06/2015	Matthias Weise	AEC3	Proposal to restructure the TOC according to project agreements	1.1
15/07/2015	WP2/WP3 partners	All	Agreement to divide into new Part A and joint Part B	1.2
22/07/2015	Jérôme Euzenat, Matthias Weise	INRIA AEC3	Final agreements about TOC of Part A	1.3
07-08 2015	Strahil Birov María Poveda-Villalón Jérôme Euzenat Bruno Fies Andrea Cavallaro Jan Peters-Anders Kleopatra Zoi Tsagkari	EMP UPM INRIA CSTB DAP AIT CERTH/ITI	Part B update: sections 5 - 10	1.4
12/08/2015	Matthias Weise	AEC3	First version of chapters 2, 3 and 4	1.5
14/08/2015	Matthias Weise, Jérôme Euzenat	AEC3 INRIA	First complete draft	1.6
18/08/2015	Jan Peters-Anders	AIT	Review of Part A	1.7
02/09/2015	Bruno Fies	CSTB	Review of Part A	1.8
07/09/2015	Matthias Weise	AEC3	Integration of Part B	1.9
09/09/2015	WP2/WP3 partners	All	Final version	2.0

Statement of originality:
This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Statement of financial support:
The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. FP7-SMARTCITIES 2013-608711

Executive Summary

This document presents a final report of the work carried out as part of work package 2 of the READY4SmartCities project (R4SC), whose goal it is to identify the knowledge and data resources that support interoperability for energy management systems. The document is divided into two parts.

Part A reflects on a "change of gear" approach after the first year of the project according to which, in addition to collecting ontologies and datasets, greater effort has been directed to increase availability of more open datasets of our built environment. Those activities have already been started in the beginning of R4SC but have been strengthened after the first year due to the low number of open datasets. The report identifies main barriers being not only on technical level but are also related to the development of use cases. Requirements have been derived to mitigate those barriers and a strategy is described to increase awareness of BIM-LOD and the overall willingness to publish own datasets. A main part of the proposed strategy is to integrate existing developments and to actively involve the whole community by setting out common goals.

One of these goals is to agree on a recommended ifcOWL ontology that can be used as a reference for further developments. This proposal was well received not only by the BIM-LOD community but also by the buildingSMART organisation that is developing the IFC standard. This momentum was used by the community to work on a proposal for an official ifcOWL standard and the foundation of a working group within the buildingSMART organisation. Meanwhile, open technical issues have been solved and it is expected that such ifcOWL standard will be available in near future (beginning/mid of 2016).

Another goal is to agree on use cases that should show the benefits and the relationship to existing developments. It should identify business opportunities that can trigger further developments. It finally turned out that this discussion is far more difficult than expected. For this, the SWIMing CSA project took over the responsibility to create a W3C community group and to lead the collection and detailing of potential use cases. One of these use cases is "Indoor Navigation" that was chosen by Ready4SmartCities to be further developed as a show case. This use case was broken down into a general data publication process that is in line with the IDM/MVD methodology of buildingSMART and follows the guidelines developed in WP4. Although this show case is not directly related to the topics of energy efficiency it is believed that it can act as an important advertising vehicle for BIM-LOD.

Looking back to the beginning of the project important steps have been made with help of R4SC. A lot of work still has to be done and coordinated. However, achieved results as well as the buildingSMART working group and the W3C community group provide a good basis for further developments and take-up by the industry.

Part B introduces the methods and processes followed for collecting and analysing ontologies and datasets. Methods have been co-developed with WP3 in respect of the interoperability area energy measurement and validation. The common process for identifying and collecting relevant resources is first updated (chapters 6 and 7) from what presented in previous version of this deliverable, followed by a description of the resources collected, namely relevant ontologies, datasets and alignments and links among them (chapters 8 and 9).

For the collection of ontologies and datasets, a special online catalogue ensures that resources are collected and recorded in a standardised way. The catalogue also allows for ease of understanding and use in terms of submission of new content, visualisation of existing resources and handling of recorded items. For the collection of alignments, an alignment server identifies and documents links and alignments among the identified ontologies and datasets.

Various collection methods were used in order to identify and collect relevant ontologies, datasets and explore possible alignments. The methods include the set-up and administration of an online survey addressed to relevant experts, stakeholders in the domains identified in the previous deliverable, literature review by the study team, analysis of standardisation and institutional bodies, and screening of resource catalogues.

Ontologies were collected using a semi-automatic process, engaging contributors, who suggested which ontologies to be included in the catalogue, populators, who added new ontologies directly into the catalogue on-line, and metadata curators, who reviewed, improved and completed the metadata of ontologies already in the catalogue. As a result, 70 ontologies were included in the catalogue during the whole project duration (32 ontologies in the first year, and 38 in the second year of the project). The current ontology offer represented in the catalogue provides good coverage of the relevant domains, although some are missing. In addition, a number of new domains not identified before as part of Level 1 nor Level 2 are included in the catalogue.

Current availability of Open Linked Data(sets) related to energy in general was found to be quite limited. Nine datasets were collected in the first year, and another nine new datasets have been included for this second and final version.

Gap analysis revealed deficits in the supply of ontologies and datasets in both interoperability areas. Though the catalogue of ontologies appears quite large, some ontologies are much specialised and others very generic, leaving some relevant conceptual areas with poor coverage. As with ontologies, the current availability of open linked data falls very short of what could be envisioned. For both domains, energy management systems and energy measurement and validation, there is a significant opportunity to improve the offer of ontologies and to encourage publication of more linked open data.

The work carried out in work package 2 and 3 provides a solid basis for any stakeholder wishing to take advantage of linked data by providing the necessary tools in the form of a comprehensive catalogue with available ontologies and datasets. This technical basis combined with the comprehensive guidelines produced as part of work package 4 enables stakeholders to produce Linked Data and raises awareness of the opportunities it offers Smart Cities towards becoming interoperable.

Glossary

AEC	Architecture, Engineering and Construction
Alignment	The result of analyzing multiple vocabularies to determine terms that are common across them.
BIM	A new method to support collaborative work in the AEC and FM industry. BIM is an abbreviation for Building Information Model(ling) and defined to be *"an improved planning, design, construction, operation, and maintenance process using a standardized machine-readable information model for each facility, new or old, which contains all appropriate information created or gathered about that facility in a format useable by all throughout its lifecycle."* (definition from the NBIMS committee).
CityGML	A standard for the exchange of virtual 3D city and landscape models. It is originally published as XML Schema Definition. An OWL version of CityGML is available from academia, which similar to ifcOWL was derived from the original CityGML schema. It defines five Level of Details, which similar to Linked Open Data is using the abbreviation LOD.
Dataset	A collection of RDF data, comprising one or more RDF graphs that is published, maintained, or aggregated by a single provider. In SPARQL, an RDF Dataset represents a collection of RDF graphs over which a query may be performed.
IDM/MVD	Information Delivery Manual (IDM) and Model View Definition (MVD) is a methodology developed by the international non-profit organisation buildingSMART. It defines a set of steps to translate business requirements to IT solutions based on the IFC standard. As such it is a universal approach, but additional specifications and tools provide specific support for IFC developments.
IFC ifcOWL IFC2x3 IFC4	Industry Foundation Classes: IFC is an Open BIM standard developed by the international non-profit organisation buildingSMART. It is a very rich object-oriented data structure that is defined in the EXPRESS modelling language (ISO 10303-11). ifcOWL is a representation of IFC using the Semantic Web ontology language OWL. It is based on agreements how to translate from EXPRESS to OWL. The latest release of the IFC standard is IFC4 Addendum 1, published in summer 2015 and mainly fixing issues in IFC4. The previous major release is IFC2x3, which is currently dominating practical use.
LGDO	Linked Geo Data Ontology: Ontology for geographic information derived from the Open Street Map specification.
Linked Data (LD)	A pattern for hyperlinking machine-readable data sets to each other using Semantic Web techniques, especially via the use of RDF and URIs. Enables distributed SPARQL queries of the data sets and a browsing or discovery approach to finding information (as compared to a search strategy). Linked Data is intended for access by both humans and machines. Linked Data uses the RDF family of standards for data interchange (e.g., RDF/XML, RDFa, Turtle) and query (SPARQL).
LBD	Whereas Linked Data (LD) is without further specification of access rights restrictions or its content the deliverable is also using Linked Open Data (LOD – see glossary) and Linked Building Data (LBD).
LDAC	Linked Data in Architecture and Construction – used to refer to an initiative by a group of people to push the use of Semantic Web technology in the architecture and construction domain.
LOD	In this deliverable used as Linked Open Data (see Linked Data). Please note that in context of building data the abbreviation LOD is also used for Level of Detail (see CityGML)

OGC	Open Geospatial Consortium (www.opengeospatial.org) is charge of the definition of standards like CityGML
Ontology	A formal model that allows knowledge to be represented for a specific domain. An ontology describes the types of things that exist (classes), the relationships between them (properties) and the logical ways those classes and properties can be used together (axioms).
Open Data	Refers to content that is published on the public Web in a variety of non-proprietary formats.
OWL	Web Ontology Language (OWL) is a family of knowledge representation and vocabulary description languages for authoring ontologies, based on RDF and standardized by the W3C.
R4SC	READY4SmartCities – used in the delivarable as short form of the project name.
RDF	Resource Description Framework (RDF) is a family of international standards for data interchange on the Web produced by W3C. RDF is based on the idea of identifying things using Web identifiers or HTTP URIs, and describing resources in terms of simple properties and property values.
SAREF	Smart Appliances REFerence (SAREF) ontology: *"Is a shared model of consensus that facilitates the matching of existing assets in the smart appliances domain."* It was published in March 2015 by TNO from Netherlands (http://ontology.tno.nl/saref/).
SKOS	Simple Knowledge Organisation System (SKOS) is a vocabulary description language for RDF designed for representing traditional knowledge organization systems such as enterprise taxonomies in RDF.
SPARQL	SPARQL Protocol and RDF Query Language (SPARQL) defines a query language for RDF data, analogous to the Structured Query Language (SQL) for relational databases. It is a family of standards of the World Wide Web Consortium.
URI	A global identifier standardized by joint action of the World Wide Web Consortium and Internet Engineering Task Force. A Uniform Resource Identifier (URI) may or may not be resolvable on the Web. URIs can be used to uniquely identify virtually anything including a physical building or more abstract concepts such as colours.
VoCamp	A VoCamp is an informal event where people can spend some dedicated time creating lightweight vocabularies/ontologies for the Semantic Web/Web of Data. The emphasis of the events is not on creating the perfect ontology in a particular domain, but on creating vocabularies that are good enough for people to start using for publishing data on the Web.

Table of Contents

1 Introduction

1.1 Purpose and Methodology

Work package 2 of the Ready4SmartCities project aims at identifying the knowledge and data that can support interoperability in energy management systems by identifying and assessing relevant ontologies, vocabularies and standards, as well as relevant datasets and alignments.

During the first year of the project, partners in WPs 2 and 3 identified 42 ontologies and 9 datasets. These were allocated between the work packages following the methodology and strategy developed as part of D2.1/D3.1. D2.2 presented the first version of the set of ontologies, datasets and alignments found by the partners and analysed as relevant to the work package domains.

Work on identification of ontologies and datasets has continued in the second year, and this deliverable presents an updated version of this work (total of 70 ontologies and 18 datasets). Also, in line with project objectives, greater effort has been directed at stakeholders and users of the project results, to further assist them in making use of the collected knowledge. This change of gear has become possible as the necessary foundation to providing useful guidance to stakeholders had been laid down through project work in year one. At the end of that year we had a working catalogue of the gathered ontologies and datasets, first alignments between these, and guidelines to transforming datasets into Linked Open Data.

The change of gear in the second project year has seen greater activity in engaging relevant stakeholders and providing them with the necessary tools and assistance to work towards interoperable data for energy management systems. This work has given rise to a number of interesting results, and complements the extension of the ontology, dataset and alignment catalogue, which are shared between WP2 and WP3. We have therefore adopted the following structure for this deliverable:

1) Activities and results from the changed gear approach documented in 'Part A' of this deliverable are unique to WP2, with the BIM community as the target stakeholders. The sub-title for the deliverable reflects this: "Towards Linked Open Building Data";

2) Building on D2.2 and D3.2 this deliverable jointly provide an updated version of the collected ontologies, datasets and alignments for interoperability of energy management systems and energy measurement and validation. This update is presented as Part B and, unlike the solution adopted for D2.2/ D3.2, this part is common for both work packages. This second version now contains **all** information on ontologies, datasets and alignments produced by the consortium in work packages 2 and 3 of the project.

1.2 Document Structure

The deliverable is divided into three parts.

Part A reports on the efforts to engage the AEC industry in publishing relevant parts of BIM models as Linked Open Data. Sections 2 – 4 discuss the activities towards a standardized ifcOWL specification and the development of a data publication show case.

Part B is an update of deliverables D2.2 and D3.2 and is common for work packages 2 and 3. Sections 5 – 9 document the collection of ontologies and datasets, their documentation using catalogues and a server, as well as a comprehensive list and description of the collected 70 ontologies and 18 datasets during the project lifetime, accompanied by a gap analysis.

At the end of the deliverable a conclusion is given for both parts.

1.3 Contribution of partners

AEC3 has the main responsibility to produce this document. The following states which partners have contributed to the different sections of the deliverable:

Partner	Contributions to sections
AEC3	WP lead
UPM	Contribution to sections 5 – 9
INRIA	Contribution to sections 5, 6, 7 and 9
DAPP	Contribution to sections 8 and 9
CERTH/ITI	Contribution to section 8 and 9
AIT	Contribution to section 8
POLITO	Contribution to section 8

Part A: Towards Linked Open Building Data

2 Barriers for publishing building datasets

Research in Ready4SmartCities has shown that there is a shortage of open datasets related to Energy Management Systems interoperability. A main area of interest of R4SC are datasets about built environment, in particular all kinds of buildings as major energy consumers[1] of a city. For instance, the envelope of a building is responsible for heat losses in winter and cooling requirements in summer, and the building equipment has a major impact on overall energy efficiency and shows energy-saving potentials of technical upgrades. Today a lot of data is already available in electronic form but is currently hidden by data owners. There are different reasons for this which are identified and discussed in this chapter.

2.1 Technical Barriers

Today, most buildings are designed by the use of IT systems, which for instance are used to predict the energy consumption of a building or simulate the user comfort under various conditions. Thus, in each design, construction and maintenance stage a lot of data is already generated in order to support required design and maintenance activities.

A major challenge since the beginning of IT-supported building design has been to share and reuse the generated data, which for instance requires to understand and to integrate information of other domains. Therefor developed solution is named Building Information Modelling (BIM) and is based on a shared, standardized data structure that has been agreed on within the building domain. The latest standard is the IFC4[2] specification from buildingSMART, which was released in 2013 and updated in 2015. Conceptually, this development provides the basis not only to exchange information between different experts but also to integrate their datasets. A lot of efforts have been made since the beginning of IFC in the mid 90's, including specification, implementation and development of guidelines. Meanwhile, BIM is accepted as a major innovation in the building industry and is demanded by various governments for their public buildings. Accordingly, the integration approach followed by the building industry provides a good basis for Linked Open Data [Bizer, Heath, & Berners-Lee 2009] as proposed by R4SC. The following advantages can be mentioned:

- IFC integrates different domains within the building industry
- IFC is a formal specification
- IFC is agreed within the building industry
- IFC is open and vendor neutral
- IFC datasets can be imported and exported by many tools
- IFC modelling technology is very similar to Semantic Web ontologies

However, there are a couple of technical barriers that need to be resolved for the use of LOD.

[1] Buildings act more and more as energy producer and, maybe even more important, as energy storage.

[2] http://www.buildingsmart-tech.org/specifications/ifc-releases/summary

2.1.1 Modelling language and used tools

The roots of IFC and used technology dates back to research projects from the 80's and 90's. The architecture and main concepts of IFC are influenced by object-oriented modelling and various knowledge representation formats from that time. The specification has been developed using the EXPRESS modelling language that itself is embedded in a family of ISO standards, namely the ISO 10303 also known as STEP: STandard for the Exchange of Product data. In many aspects the EXPRESS language is very similar to the Web Ontology Language OWL[3] as being recommended in Semantic Web developments. For instance, both languages use classes or entities and know the concept of inheritance, the definition of links and consistency constraints. Accordingly various researchers have worked on solutions to transfer the IFC data structure to an OWL-based representation. The main idea of those solutions is to agree on meta-model level how to map EXPRESS-based data structures to an OWL-based ontology. Although there are many similarities there are modelling concepts that are not directly transferrable. In such cases, various strategies have been proposed leading to different OWL representations of the same IFC data structure. Proposed solutions for dealing with mapping issues typically depend on expected use cases, which sometimes result in the decision to omit specific features (e.g. WHERE rules) or require to choose from different options each coming with own advantages and disadvantages (e.g. naming of properties). This diversity currently hinders further use of Sematic Web technologies and ifcOWL.

The current situation can be summarized as follows:

Barriers:

- IFC is not based on Semantic Web technology;
- Existing IFC datasets need to be transferred to RDF-graphs based on ifcOWL;
- There is no official ifcOWL version yet.

Requirements:

a) Agreements for an official ifcOWL representation;
b) Providing of tools for connecting both technologies.

2.1.2 Domains covered by IFC

The IFC data structure has gone through 20 years of development and meanwhile covers many domains and life-cycle stages of a building. However, the schema does not cover all aspects of a building nor all related areas that might be of interest. The current schema is rather focused on building data that is relevant for design coordination and data handover (see Figure 1). For instance, GIS, weather data or highly specialized domain data like finite element meshes are out of scope. This is leading to the situation where on one hand a lot building data can be integrated into an IFC model but on the other hand will never cover all data for special design activities and the scale is limited to buildings. In order to address larger models at the district or city scale there is a need to rely on other standards like CityGML. Thus, a strategy is needed how to connect with other data sources and to support use cases that go beyond the design of single buildings. Linked (Open) Data can provide answers how to integrate other data structures. However, LOD is not yet realized as a potential solution mainly because like IFC most other data sources are not based on Semantic Web technologies.

[3] http://www.w3.org/TR/owl-ref/

The current situation can be summarized as follows:

Barriers:

- IFC is focused on building data and does not cover all aspects in all details;
- There is no agreed strategy how to integrate other data sources.

Requirements:

c) The LOD approach needs to be recognized as a potential solution for integration of data that is out of scope for IFC.

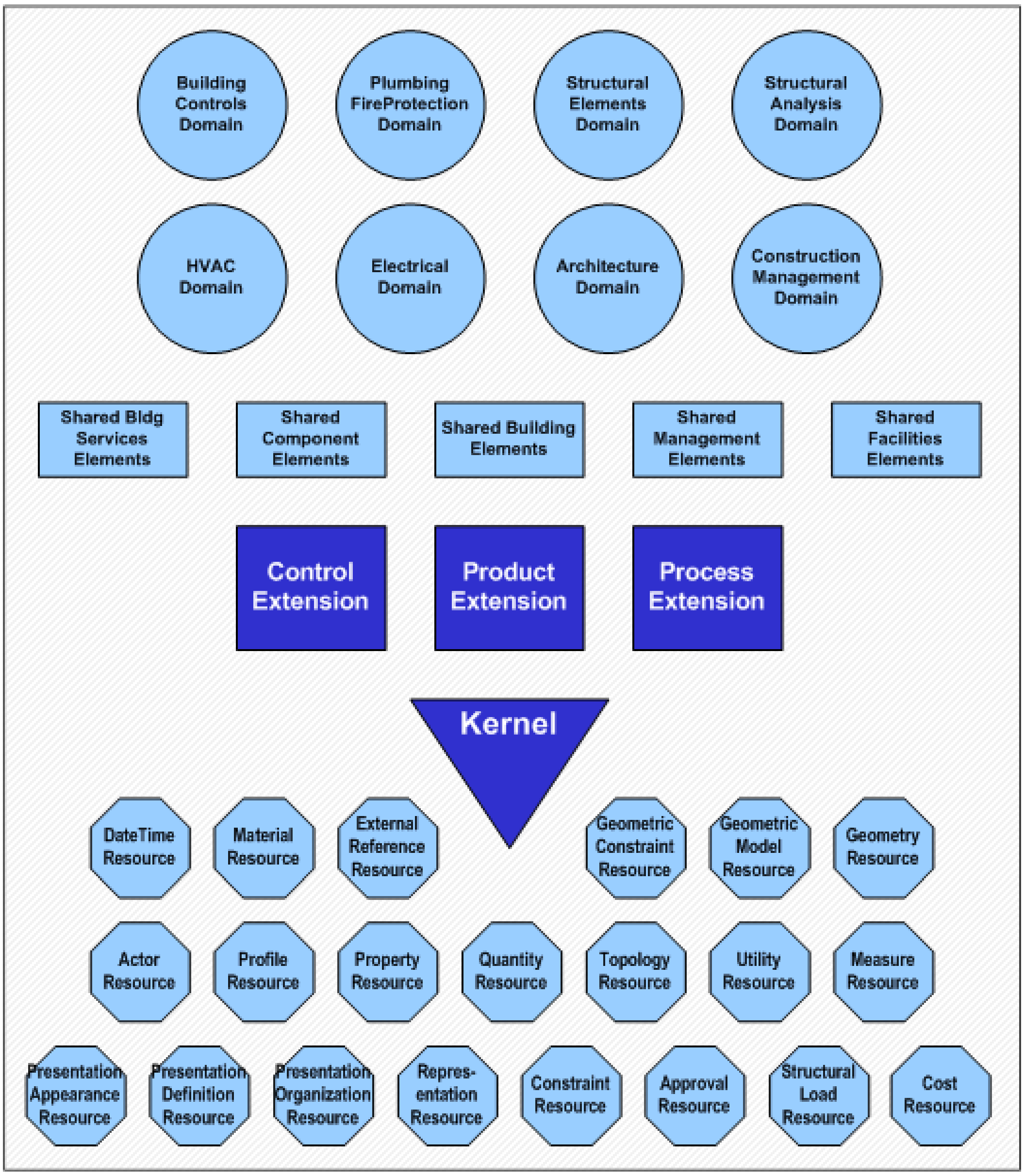


Figure 1: IFC architecture consisting of a set of sub schemata each covering specific aspects of a building

2.1.3 Implementation of IFC

IFC-based data exchange is implemented for design coordination and does not cover the whole IFC data structure. Also, interfaces are still based on IFC2x3 release and not yet on the latest IFC4. This is leading to the situation that a lot of BIM data that might be of interest in context of R4SC is not or only partially available as an IFC dataset. For instance, a lot of data that is required for energy analysis or building monitoring is not yet provided as BIM data. IFC implementation is currently focused on design coordination activities that for instance merge main physical objects (building elements, HVAC equipment) for making clash detection and visual checks. Building simulation is typically done inside the various simulation packages using their own internal data models (see Figure 2).

The current situation can be summarized as follows:

Barriers:

- The building industry is not yet ready to provide all BIM data in a neutral format.

Requirements:

d) LOD use case development needs to be aligned with capabilities of existing BIM tools.

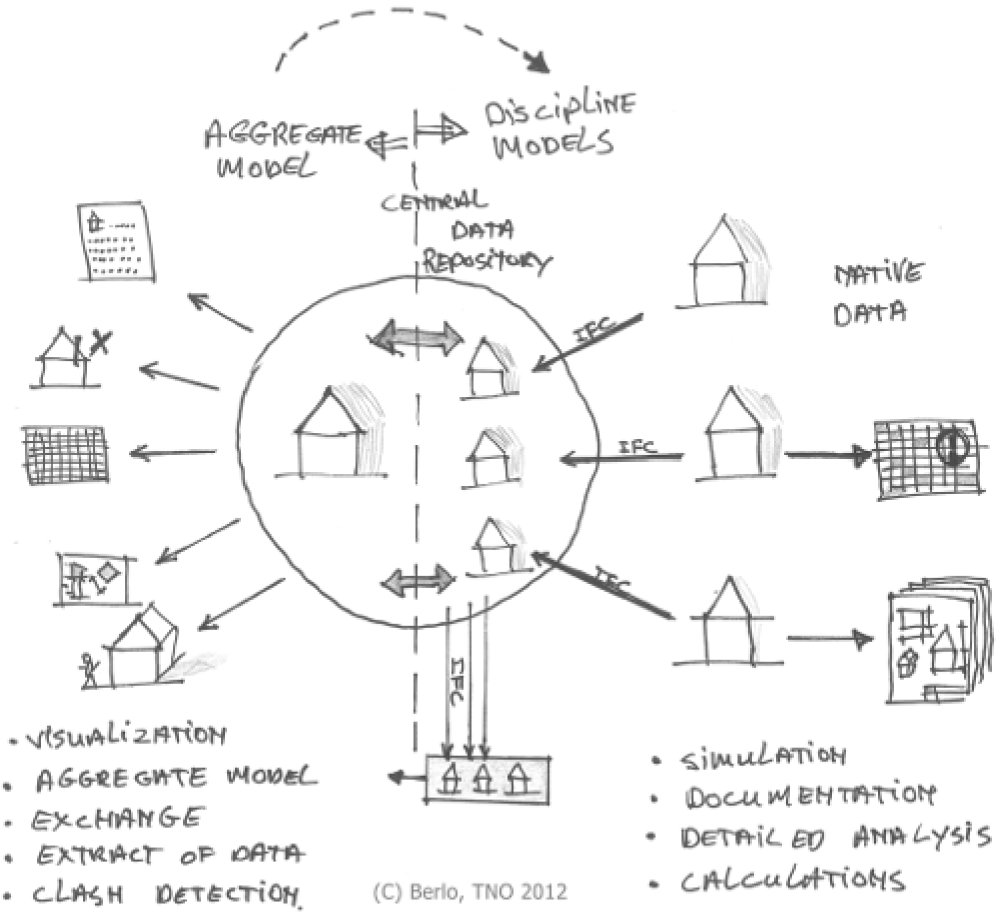


Figure 2: Reference model approach and main use cases for shared BIM data (by Leon van Berlo, TNO)

2.2 Security, privacy and other issues to avoid data publication

BIM data is shared within the design team and not published on the internet. There might be some fact sheets and renderings that are used for advertisement, but the details of a building are typically kept private. There is simply no tradition of making building data publicly available and there are only very few examples of downloadable BIM files today. There are different reasons for this:

- Hide specific business secrets from competitors (e.g. price tags of products, detailed design solutions, way of working with BIM etc.);
- Protect safety-critical data in particular for special kinds of buildings like courthouse, prisons, power plants or critical infrastructure, but also normal buildings in case of terrorist threats or risk of burglary;
- Protect private data of occupants that for instance enables to track user behaviour, availability times or a like.

The problem with BIM is that it integrates a lot of data into a single dataset, which makes it necessary to identify and remove critical data (see WP4 guidelines). The current situation can be summarized as follows:

Barriers:

- BIM data publication requires additional efforts to identify and remove critical data.

Requirements:

e) Give user easy solutions to control the data publication process.

2.3 Business Case

Linked (Open) Data is not yet a topic in the building industry as there are no clear business cases. Also, as shown in section 2.1 there is not only a noteworthy technical barrier to use Semantic Web technology but also a gap of other external data sources that are of interest for linking with BIM data. Accordingly, there is no immediate benefit in terms of financial profit for using this technology. The current situation can be summarized as follows:

Barriers:

- Noteworthy investment costs for publishing BIM data without clear strategy for a return of investment;
- Missing data sources and infrastructure to make use of LOD.

Requirements:

f) Reduce the efforts for BIM data publication (see also b – combine BIM-LOD with existing technologies);
g) Develop use cases to show benefits for data owners.

3 Steps towards Linked Open Building Data

The use of Linked Open Building Data, in particular in combination with BIM/ifcOWL developments, is still in research phase due to open questions from the industry and missing reference implementations. This chapter clarifies the state-of-the-art in using ifcOWL, the current state of the standardization effort and use case development and finally reports about events that have been supported during the last two years to push those developments.

3.1 State-of-the-art review

First proposals for translation of the IFC standard to an OWL-based ontology dates back to 2005. Beetz et al. [2005, 2009] presented two approaches to derive an OWL notation from the Industry Foundation Classes which at that time was a feasibility study without a specific use case in mind. However, the work was motivated by a set of requirements related to agent-based systems like information discovery and retrieval, distributed data storage, semantic web services and mapping to other data formats. Other projects and specific use cases followed this work (e.g. InteliGRID, DRUM[4], KnoholEM[5], DuraARK[6]) leading to adjustments for the transformation from EXPRESS to OWL. In fact, beside data publication and linking there are other main use cases like reasoning and knowledge acquisition (see Figure 4). This was effectively leading to various ifcOWL versions, each optimized for a specific use case (see Figure 5). While there have been good reasons for those configuration settings it was not an ideal situation to go towards industrial use. Accordingly, the group of experts decided to give recommendations and to reduce the number of ifcOWL versions. It was also proposed to publish ifcOWL as a buildingSMART standard.

EXPRESS construct	OWL construct
ENTITY	Class
SUB/SUPERTYPE	subClassOf
SELECT	Class and subClassOf
INVERSE	inverseOf / InverseFunction-alProperty
ENUM	DatatypeProperty […] owl:DataRange […] owl:one of […] rfd:List or enumerat-ed classes.
Cardinality constraints	owl:cardinality, owl:minCar-dinality, owl:maxCardinality
Simple Types	Simple XML Schema types
WHERE domain rules	Possibly through SWRL rules in future
collections: LIST, SET, BAG, ARRAY	Only unordered (?)

Figure 3: Decisions by Beetz et al. [2005] for mapping of IFC from EXPRESS to OWL

4 http://cse.aalto.fi/en/research/groups/distributed_systems/projects/drum/

5 http://www.knoholem.eu

6 http://duraark.eu/

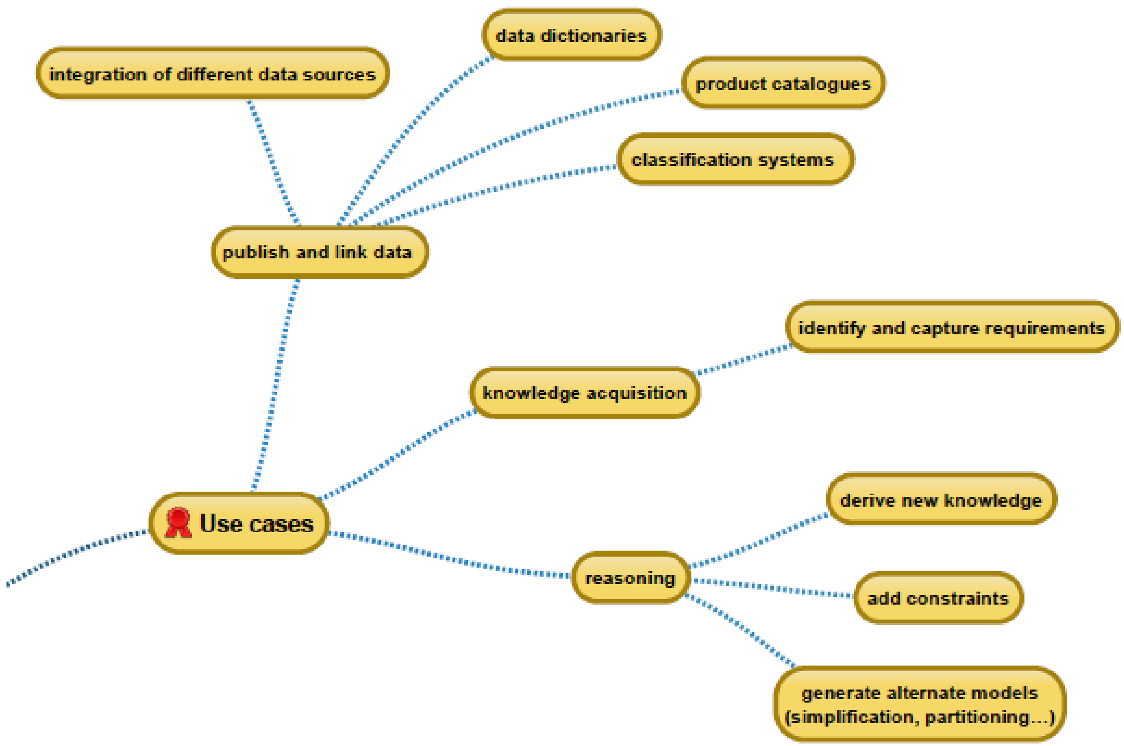


Figure 4: Main use cases for ifcOWL

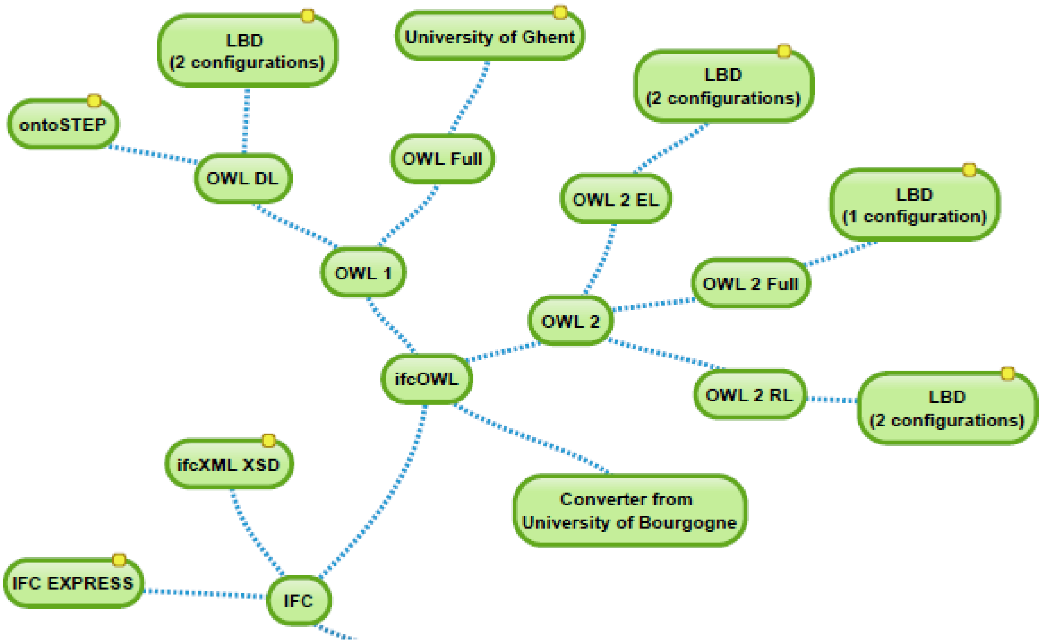


Figure 5: Overview of ifcOWL versions

3.2 Standardization of the ifcOWL ontology

Work on a recommended ifcOWL version has been started by the group of ontology experts in May 2014 as a result of LBD Workshop/VoCamp held in Espoo/Finland. Final decisions have been made in the last LBD Workshop in Eindhoven and will be presented at the buildingSMART summit in October 2015 in Singapure. In general, it was agreed to be as close as possible to the original IFC definition, but without converting rules, functions and derived attributes. The following list of issues have been discussed and decided in Eindhoven for the ifcOWL proposal[7]. It gives an impression of the number of agreements to be made (compare also with Figure 3):

1) Domains and ranges to be included or not?
2) PropertyName - Consistent use of long names or only use of long names when required (inconsistent)?
3) PropertyName - Camelcase long names or Underscore long names?
4) PropertyName - First Class Name, then Property Name?
5) PropertyName - Exclude "Ifc" from all names (classes, properties)?
6) PropertyName - name_of_IfcRoot | name_IfcRoot?
7) Use of "Grouping" SubProperty relations (Property "Name") or not?
8) How to convert LISTs and ARRAYs?
9) How to convert SETs?
10) Inverse attributes?
11) NUMBER simple datatype is considered as an INTEGER or as a REAL?
12) REAL simple datatype is considered as an xsd:double, xsd:decimal or xsd:real?
13) LOGICAL simple datatype is considered as an xsd:boolean, or as an enumeration of TRUE, FALSE, and UNKNOWN?
14) BOOLEAN simple datatype is considered as an xsd:boolean, or as an enumeration of TRUE and FALSE?
15) How to declare ENUM datatypes?
16) How to declare SELECT datatypes?
17) What namespace structure do we use?
18) Naming individuals / instances?
19) EnumName - what naming to use?
20) Which license to use?

First reference implementations for converting both the IFC schema and IFC datasets are already available, but need to be adjusted to reflect decisions from Eindhoven. The next steps will be to get accepted as a proposal by buildingSMART, which then has to be reviewed by the community and finally released as an official buildingSMART standard. Latest developments are already published on the buildingSMART website[8].

[7] Minutes available at: https://docs.google.com/document/d/1d44liBjvuzujk3_c5RjrPtPDBVb1Zv6WPubXPqvIRjg

[8] http://www.buildingsmart-tech.org/future/linked-data/linked-data

3.3 Clarification of ifcOWL use cases

A discussion about ifcOWL-based use cases has been started in parallel to the discussion about the ifcOWL standard. Clarity about possible use cases is seen as a critical step for further acceptance and take-up by the industry. In particular, the position of ifcOWL in relation to existing technologies is an important statement to agree on directions of development. The use case discussion is currently organised in a W3C community group and coordinated by the SWIMing project[9], also funded by EU as a Coordination and Support Action. While the current focus is on LOD and energy efficiency collected use cases will go beyond these application areas.

End of July 2015 the Wiki page reports about over 30 use cases grouped into different high level categories. Each use case is described with a set of meta-data in order to be able to more easily identify similarities and differences. The collected use cases are organized along three main critierias: (1) the building life-cycle stage to define “when”, (2) the involved stakeholders to define “who” and (3) the domains and types of data to define “what”. This approach fits very well to the IDM/MVD methodology proposed by buildingSMART to support use case development and software implementation. While only “domains and types of data” (3) are directly related to ontologies and datasets and thus the main area of interest in WP2/3 of R4SC, the two other criteria are related to the business case, i.e. its integration into design, construction and maintenance tasks of a building.

Collected use cases have been derived mainly from ongoing research projects. Some of these projects already make use of ontologies and Linked Data. In order to raise interest and trigger further feedback from the community a critical discussion shall show expected benefits and challenges for using Semantic Web technologies. Besides the technical developments and standardization efforts within the buildingSMART task group (see 3.2), the W3C community group is also aiming at getting in contact with other domains being outside of the typical building domain. Further information about collected use cases and the W3C community group can be found in the D4.3 of WP4.

3.4 Supported events, cooperation and main results

R4SC actively supported further development of ifcOWL, in particular its standardization through buildingSMART, and discussion about L(O)D use cases. The following events have been visited, organized or supported by R4SC and its partners, namely AEC3, CSTB, INRIA and UPM.

- **VoCamp in Brussels (Feb. 2013)**
  2nd VoCamp within the series of workshops started by CERTH/ITI end of 2012,
  The topic of the VoCamp was about Building Information Modelling (BIM) and was
  organised by the European Commission and AEC3
- **Group of onlogy experts (Feb. 2014)**
  Initiative started by R4SC to get in contact with people from research, government and industry active in the area of ontologies and building data. The request showed an overwhelming interest in joining efforts for further developments. It was the starting point for further activities and organisation of the W3C and buildingSMART groups. The list of people at around mid of 2014 is shown in Figure 6.

[9] http://www.swiming-project.eu/

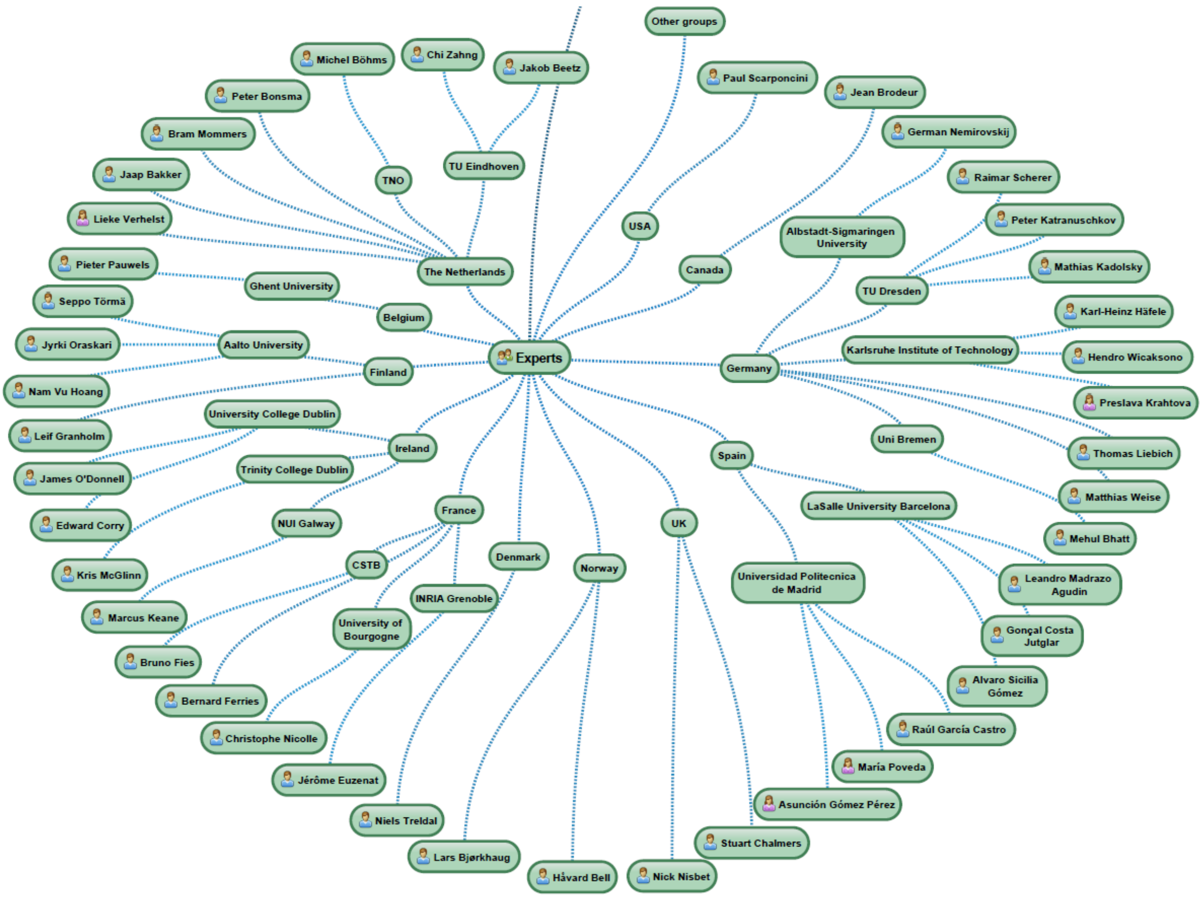


Figure 6: List of ontology experts showing interest in joining efforts

- **VoCamp in Barcelona (Feb. 2014)**
  4th VoCamp about Integrating multiple domains and scales. The workshop was organised by ARC Engineering and Architecture La Salle. The event was used to get in contact with the GIS community and the decision was made to organize a VoCamp together with the LDAC initiative.
- **LDAC Workshop/VoCamp in Espoo (May 2014)**[10]
  6th VoCamp about Linked Data in Architecture and Construction. The workshop was organised by Aalto University and Tekla. The two VoCamp champions have been funded by R4SC:
  (1) Peter Bonsma, RDF Ltd. – leading the Platform track and
  (2) Jaap Bakker, CB-NL – leading the use case track

[10] http://linkedbuildingdata.net/wp-content/uploads/2014/05/LDACworkshopreport.pdf

Figure 7: Participants of the LDAC Workshop/VoCamp in Espoo 2014

- **Workshop "Building Knowledge" in Graz (Sept. 2014)[11]**
  The workshop was part of the i-KNOW conference. R4SC presented work related to ontology alignment. The event was also used to discuss barriers for publishing BIM data.
- **buildingsSMART summit in London/Watford (March 2015)[12]**
  Technical Room "Semantic Web and bSI technical Roadmap" Working Group
  Use cases and status of technical developments have been presented and discussed with the buildingSMART community. The event was finished with the decision to found an official working group within buildingSMART.
- **LDAC Workshop in Eindhoven (July 2015)[13]**
  Final decisions towards a standardized ifcOWL representation have been made. Feedback to use cases collected by the W3C group.

Besides two series of regular web conferences the next important event is buildingSMART summit in Singapore in October 2015[14]. It is planned to present the proposal for a standardized ifcOWL release.

[11] http://duraark.eu/presentations/i-know-2014-workshop-linked-building-data/

[12] https://www.w3.org/community/lbd/wiki/images/1/1d/20150302_V2_bSI_Watford_LinkedOpenData_Agenda.pdf

[13] http://ldac-2015.bwk.tue.nl/index.html

[14] http://www.buildingsmart.org/event/standards-summit-singapore-2/

4 Results

A main motivation in the second year of WP2 was to improve availability of open BIM datasets. It was decided to follow a sustainable strategy that is focused on increasing overall awareness of LOD within the BIM community and to reduce the barriers as identified in chapter 2. It is not expected to lead to immediate results in terms of a specific number of open BIM datasets. It is rather expected to contribute to a sustainable grow of stakeholders that are willing to publish their datasets. So far, a very active core team of ontology experts is pushing further developments, they are organised within the main standardization bodies, namely buildingSMART and W3C, and have a clear agenda for the next years. This chapter describes the selection strategy, started developments as well as achieved results and shows first examples.

4.1 Strategy to publish BIM data and to link with other data sources

The vision of R4SC is based on the assumption to have open access to different kinds of datasets of the city in order to discovery energy and CO_2 savings and to improve life quality of residents. Implementation of that vision will take time and is expected to start with trials and individual initiatives. It is not expected that the full potential of LOD can be realized in short time frame and by decision of individual players within a city. Due to the number of potential stakeholders and diversity of their backgrounds it is important to get started with very simple, but meaningfull examples.

Accordingly, R4SC has been looking for a small "Hello ifcOWL World" example that acts as a show case for BIM owners. Also, the development should be driven by the community to get feedback and combine forces. Based on the requirements identified in chapter 2 the following decisions were made:

- **Select a use case with defined data requirements**
  According to requirement e) described in chapter 2.2 we do not follow the strategy to publish the whole BIM data set. It is proposed to focus on a minimal dataset that is required by a specific use case, which would also fit to the IDM/MVD methodology followed by buildingSMART. Keep it as simple and controllable as possible. This is also linked to requirement g).
- **Stick to public standards and methods**
  To raise acceptance and reuse of tools existing standards shall be used as far as possible. The example should be based on a standardized ifcOWL representation and should follow the IDM/MVD principle for use case development (see next chapter). This is derived from requirements a) and f).
- **Select use cases that go beyond the building domain and are supported by existing tools**
  There are a couple of restrictions for implementation of an realistic and meaningful use case. Following requirements c) and d) it is necessary to check if required data can be produced by existing BIM authoring tools and what kind of non-BIM data is available to demonstrate the use of links.

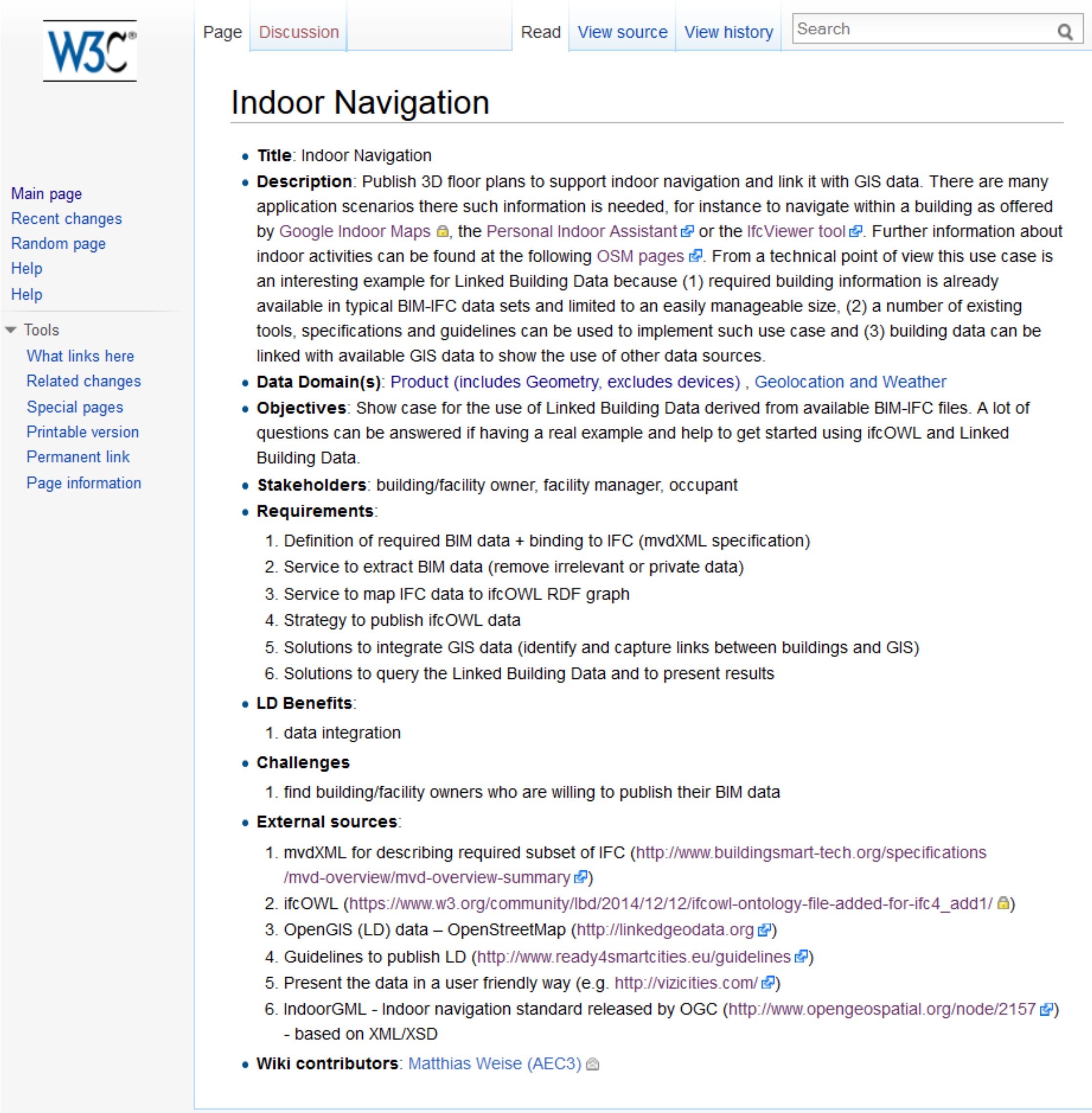
These criteria finally led to the decision to select "Indoor Navigation" as show case.

4.2 Indoor Navigation show case

The "Indoor Navigation" use case is documented at the W3C use case wiki[15] (see Figure 8) and was finally chosen, although not directly related to the topic of energy-efficiency, because of the following reasons:

[15] https://www.w3.org/community/lbd/wiki/Indoor_Navigation

- There is only very few BIM data that is required (see chapter 4.3) and most BIM authoring tools are able to produce that data.
- Data can be linked to other domain data, in this case GIS datasets, in particular Open Street Map data that is also available as RDF.
- The use case is of increasing interest and there are a couple of tools available supporting various types of indoor navigation[16]. Furthermore, since January 2015 there is an OGC standard[17] about indoor navigation available.



Page Discussion Read View source View history Search

Main page
Recent changes
Random page
Help
Help
Tools
What links here
Related changes
Special pages
Printable version
Permanent link
Page information

Indoor Navigation

- **Title**: Indoor Navigation
- **Description**: Publish 3D floor plans to support indoor navigation and link it with GIS data. There are many application scenarios there such information is needed, for instance to navigate within a building as offered by Google Indoor Maps, the Personal Indoor Assistant or the IfcViewer tool. Further information about indoor activities can be found at the following OSM pages. From a technical point of view this use case is an interesting example for Linked Building Data because (1) required building information is already available in typical BIM-IFC data sets and limited to an easily manageable size, (2) a number of existing tools, specifications and guidelines can be used to implement such use case and (3) building data can be linked with available GIS data to show the use of other data sources.
- **Data Domain(s)**: Product (includes Geometry, excludes devices) , Geolocation and Weather
- **Objectives**: Show case for the use of Linked Building Data derived from available BIM-IFC files. A lot of questions can be answered if having a real example and help to get started using ifcOWL and Linked Building Data.
- **Stakeholders**: building/facility owner, facility manager, occupant
- **Requirements**:
  1. Definition of required BIM data + binding to IFC (mvdXML specification)
  2. Service to extract BIM data (remove irrelevant or private data)
  3. Service to map IFC data to ifcOWL RDF graph
  4. Strategy to publish ifcOWL data
  5. Solutions to integrate GIS data (identify and capture links between buildings and GIS)
  6. Solutions to query the Linked Building Data and to present results
- **LD Benefits**:
  1. data integration
- **Challenges**
  1. find building/facility owners who are willing to publish their BIM data
- **External sources**:
  1. mvdXML for describing required subset of IFC (http://www.buildingsmart-tech.org/specifications/mvd-overview/mvd-overview-summary)
  2. ifcOWL (https://www.w3.org/community/lbd/2014/12/12/ifcowl-ontology-file-added-for-ifc4_add1/)
  3. OpenGIS (LD) data – OpenStreetMap (http://linkedgeodata.org)
  4. Guidelines to publish LD (http://www.ready4smartcities.eu/guidelines)
  5. Present the data in a user friendly way (e.g. http://vizicities.com/)
  6. IndoorGML - Indoor navigation standard released by OGC (http://www.opengeospatial.org/node/2157) - based on XML/XSD
- **Wiki contributors**: Matthias Weise (AEC3)

Figure 8: Description of the show case "Indoor Navigation" that was used to explain the technology

[16] http://wiki.openstreetmap.org/wiki/Indoor/Projects

[17] http://www.opengeospatial.org/standards/indoorgml

It was found that this in summary is very good basis to engage the community to take part in further use case developments.

4.3 Relevant data and ontologies

As already mentioned in chapter 4.1, the proposed use case development followed the IDM/MVD methodology from buildingSMART[18]. Besides the description of the processes and its dependencies[19] it is necessary to be clear about (1) the required data types and (2) their mapping to a technical solution in form of a data structure or ontology. The first is named Exchange Requirement and the second Model View Definition.

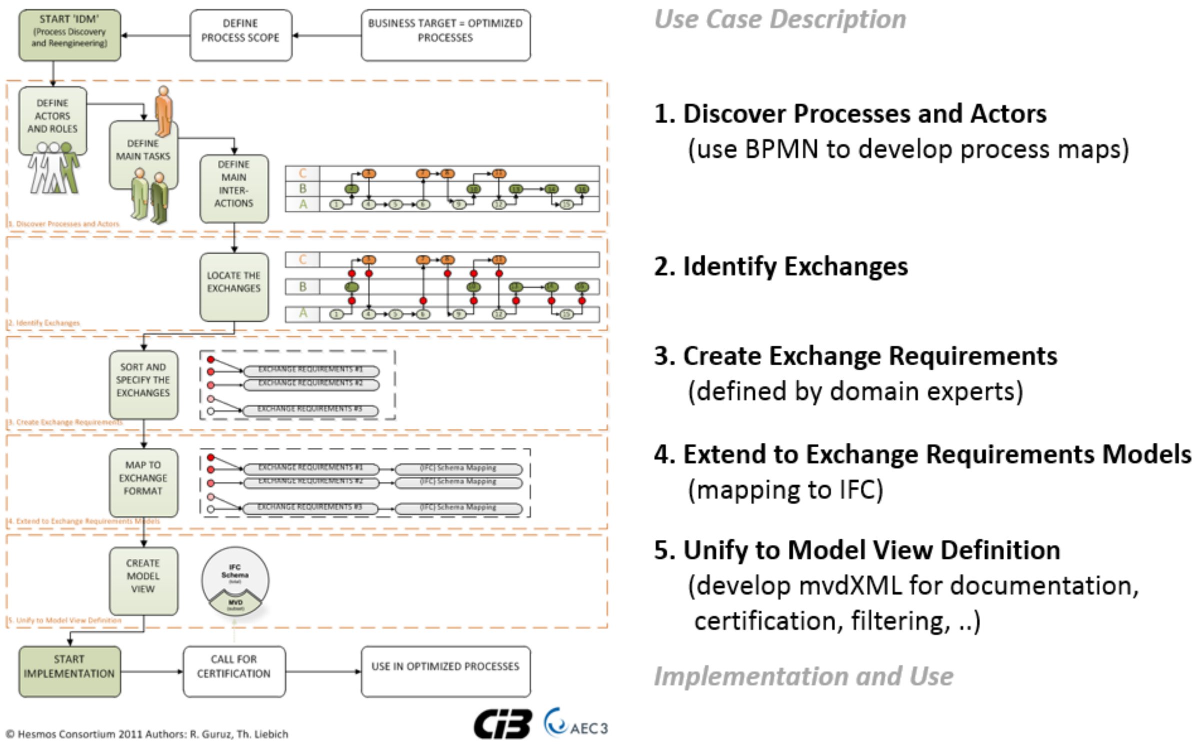


Figure 9: Overview about main step of the IDM/MVD methodology [Liebich et al. 2013]

4.3.1 Exchange Requirements

Defining exchange requirements is in fact very similar to the development of a domain vocabulary as being agreed in a VoCamp. Required input data in terms of objects, attributes and relationships are defined by domain experts using their own language. However, the result of that step is less formal and is typically captured in a spreadsheet form as shown in Figure 10 and Figure 11. For the Indoor Navigation use case 8 object types with around 50 attributes or relationships have been identified as being relevant as BIM input data.

[18] http://www.buildingsmart-tech.org/

[19] Use cases are regarded as processes or tasks that have responsibilities and dependencies. Proper modelling in form of a process map is particularly important in highly collaborative scenarios like for instance the building design.

Concept	Description	Type	Owner
Building		Object	
Address	Visitor address of the building. The address information can be used for rough positioning and can be used if geolocation of the building is not provided.	Group	Building owner
Address line	e.g. street name and street number	Text	Building owner
Country		Text	Building owner
Region		Text	Building owner
Town		Text	Building owner
Zip code		Text	Building owner
Containment	Mapping definition to IFC is tricky because the spatial structure is defined by IfcRelAggregates (for IfcSpatialStructureElements) whereas containment of building elements is defined by IfcRelContainedInSpatialStructure. Â	Group	
Containment in site		Reference	Building owner
Identification	Properties within that group shall enable to identify elements.	Group	
ID	Unique id of an element that can be used by IT systems.	Text	Building owner
Name	Id of an element that is typically used by humans to identify elements.	Text	Building owner
xx : IFC Mapping Properties	This group shall collect all definitions that are needed to specify the mapping to IFC. From a conceptual point of view they might be not relevant.	Group	
Placement		Group	
Local coordinate system	Placed relative to some other coordinate system and that is used to define the position of objects.	Data	
Entrance		Object	
Containment	Mapping definition to IFC is tricky because the spatial structure is defined by IfcRelAggregates (for IfcSpatialStructureElements) whereas containment of building elements is defined by IfcRelContainedInSpatialStructure. Â	Group	
Containment in building		Reference	
Containment in space		Reference	Building owner
Containment in storey		Reference	
Placement		Group	
Local coordinate system	Placed relative to some other coordinate system and that is used to define the position of objects.	Data	Building owner

Figure 10: Exchange Requirements for Building and Entrance

Both examples show a semi-formal definition that can either be transformed into a new ontology (in case somebody wants to start from scratch) or mapped to an existing data structure like IFC. This step is explained in the next section and is done by a modelling expert who understands the meaning of the domain requirements and knows the data structure that is capable to store that information.

At this stage, data owners can also tell whether some information can be published or shall be kept private. If for instance the address of a building shall not be published it can be marked as NOT allowed in order to be excluded from data publication. All BIM data that is not explicitly mentioned in the Exchange Requirement will be ignored by default in the Model View Definition.

4.3.2 Model View Definition

Although not only BIM data is relevant for the Indoor Navigation use case this section is focused on discussing the mapping to the IFC data structure. The overlap to CityGML and OpenStreetMap is presented in the next section showing the results of the alignment algorithms.

Starting with the Exchange Requirements coming from the domain experts a first step is to roughly identify relevant entities and attributes of the IFC data structure. This is typically captured in a separate column of the same spreadsheet (see Figure 11).

Concept	Description	Type	Owner	IFC4
Space		Object		IfcSpace
Accessibility		Group		
Connection to access system	In particular to the vertical access system.	Reference	Building owner	
Connection to space	To be defined via doors.	Reference		Fundamental Concept: Space Boundaries" (should be limited to IfcDoor only)"
Containment	Mapping definition to IFC is tricky because the spatial structure is defined by IfcRelAggregates (for IfcSpatialStructureElements) whereas containment of building elements is defined by IfcRelContainedInSpatialStructure. Â	Group		Fundamental Concepts: Composition.Aggregation""
Containment in building		Reference		IfcRelAggregates
Containment in space		Reference		IfcRelContainedInSpatialStructure
Containment in storey		Reference	Building owner	IfcRelAggregates
Geometric representation		Group		
3D brep geometry		Data	Building owner	
Identification	Properties within that group shall enable to identify elements.	Group		fundamental concept: Roots.Identiy
ID	Unique id of an element that can be used by IT systems.	Text	Building owner	IfcRoot.GlobalId
Name	Id of an element that is typically used by humans to identify elements.	Text	Building owner	IfcRoot.Name
Placement		Group		fundamental concept: Product.Placement
Local coordinate system	Placed relative to some other coordinate system and that is used to define the position of objects. NOTE: Location of space is typically defined relative to the building storey, which again is placed relative to the building. That means that the location of the all other locations is needed too.	Data	Building owner	IfcProduct.ObjectPlacement
xx : IFC Mapping Properties	This group shall collect all definitions that are needed to specify the mapping to IFC. From a conceptual point of view they might be not relevant.	Group		
Classification		Group		
DIN 277	Would enable to differentiate between space types (-> horizontal access system). But relevant only for German market, classification may not be provided.	Data		
Storey		Object		IfcBuildingStorey
Containment	Mapping definition to IFC is tricky because the spatial structure is defined by IfcRelAggregates (for IfcSpatialStructureElements) whereas containment of building elements is defined by IfcRelContainedInSpatialStructure. Â	Group		fundamental concepts: Spatial Composition, Spatial Decomposition
Containment in building		Reference	Building owner	IfcRelAggregates
Hight level		Group		
Elevation	Hight value given to some reference hight of the building.	Real	Building owner	IfcBuildingStorey.Elevation
Level name	Numbering or abbreviation describing the building level, e.g. UG, EG, OG1, ...; A, B, C, ...	Data	Building owner	Pset_BuildingStoreyCommon.Reference
Identification	Properties within that group shall enable to identify elements.	Group		fundamental concept: Roots.Identiy
ID	Unique id of an element that can be used by IT systems.	Text	Building owner	IfcRoot.GlobalId

Figure 11: Exchange Requirements for Space and Storey and its representation in IFC.

Based on this preporatory work a formal specification is developed using the IfcDoc tool from buildingSMART. It essentially allows to access to the whole IFC specification and to select required entities, attributes and relationships. Figure 12 shows the IfcDoc tool and all selected entities that for instance include the IfcDoor entity as a way to define “Connection to space” and “Connection to access system” as requested by domain experts.

Once finished the specification can be exported as:

- mvdXML file containing all definitions of the MVD
- EXPRESS schema containing relevant IFC definition only
- HTML documentation for implementation of the Indoor Navigation use case by tool vendors (same style as the IFC4 documentation)

Both, the mvdXML file as well as the EXPRESS subset schema enable to control publication of BIM data. They can be used to remove irrelevant (or critical) BIM data, and in case of mvdXML also enable to check whether all required data is properly contained in an IFC dataset or if some data like the space connection geometry is missing for indoor navigation.

After all required data has been extracted and saved back to an IFC file it can then be processed by an IFC to ifcOWL-RDF converter[20] and finally published in a triple store (see also R4SC Deliverable D4.2).

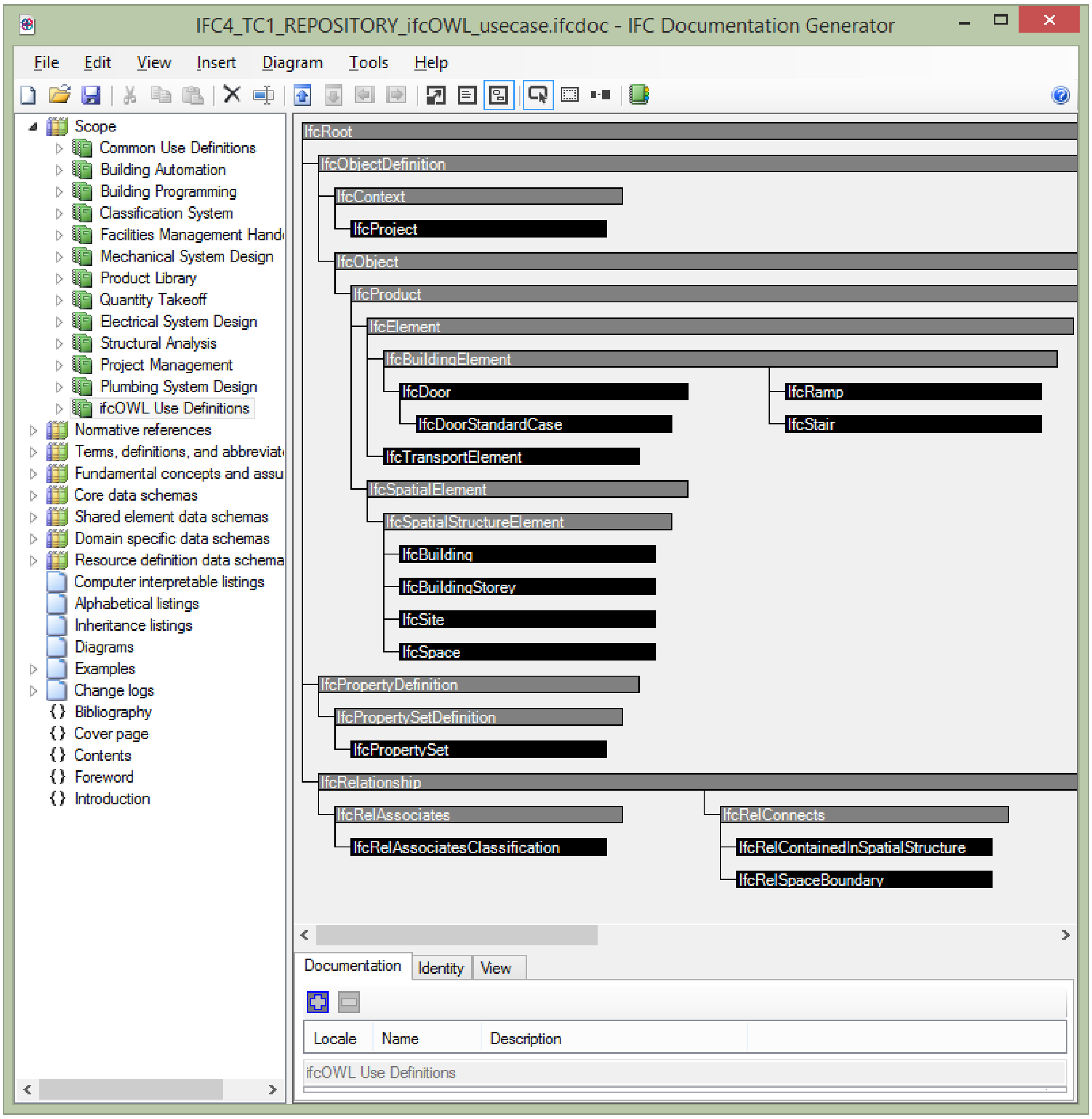


Figure 12: Formal specification of a Model View Definition using the IfcDoc tool.

[20] Conversion of IFC files has been tested with the tool developed by Pieter Pauwels, which is available as a temporary web service (http://smartlab1.elis.ugent.be:8889/IFC-repo/). However, at the time of this writing the service was not updated to latest ifcOWL agreements (see chapter 3.2) so that generated ifcOWL is not compliant with latest proposal.

4.4 Alignments

In order to gather alignments for IFC we decided to follow the regular process for generating alignments and publishing them in the Ready4SmartCities Alignment server[21]. The result of this alignment process in combination with the Model View Definition presented in chapter 4.3.2 indicates potential overlaps of the different datasets.

Below we describe the concerned ontologies, the result of the ontology matching process and discuss the specific point of matching versions of the same ontology.

4.4.1 Ontologies to be matched with IFC

In 2014 we had already matched IFC2x3 against many other ontologies identified in WP2 and WP3. It was thus part of our panel of ontologies to match.

For this use case example we used IFC4 (and not the latest IFC4 ADD1) for representing IFC4 upon discussions with the ifcOWL creator. Our goal was to match them against two eminent geographic ontologies: CityGML[22] and Linked Geo Data Ontology (LGDO) extracted from Open Street Map[23].

We have been able to provide alignments with LGDO, but not with CityGML as CityGML has a peculiar use of different namespaces within the same ontology, which confused the available matchers. Instead, we introduced alignments with SAREF[24], which is of interests in current discussions. In Table 1 we provide the respective statistics of these ontologies which were all in OWL.

	#classes	#properties	#instances	#(C+P+I)	#triples
IFC2x3	952	948	0	1901	14807
IFC4	1221	1576	1624	4421	38178
LGDO	1200	222	0	1422	24530
SAREF	110	42	73	310	1382
CityGML	184	279	0	463	2527

Table 1: Respective size of the considered ontologies.

Table 1 shows that the IFC ontologies have larger size than the others, with LGDO as a close follower. Since the IFC ontologies are about buildings, SAREF about appliances, LGDO and cityGML about geographic information,

[21] http://al4sc.inrialpes.fr

[22] http://cui.unige.ch/isi/onto/citygml2.0.owl

[23] http://linkedgeodata.org/ontology/

[24] http://ontology.tno.nl/saref

we can expect strong matching between the IFC versions, moderate matching between IFCs and all other ontologies and little matching between LGDO and SAREF.

4.4.2 Resulting alignments

The totality of the results, as well as the whole process is described in the Appendix B of this deliverable. The resulting alignments are available in the Ready4SmartCities alignment server. We discuss here the results provided by these ontologies.

	IFC2x3	IFC4	LGDO
IFC4	2644		
LGDO	102	106	
SAREF	49	111	49

Table 2: Size of raw alignments returned by the matchers.

From the sheer sizes alone (Table 2), it is possible to see that indeed, SAREF and LGDO have less in common and that IFC2x3 and IFC4 have most in common with the other combinations in the middle. Surprising results are those between SAREF and IFC2x3, especially in relation with those of IFC4. This is a clear indication that IFC4 must have been extended in the domain of appliances.

	IFC2x3	IFC4	LGDO
IFC4	2283		
LGDO	57	51	
SAREF	42	71	26

Table 3: Size of curated alignments.

Between LGDO and SAREF, there are many false positive matches: For instance, saref:Water is matched against lgdo:Water, lgdo:WaterMill, lgdo:WaterFall, lgdo:WaterTower, lgdo:WaterPoint, etc. The same applies to saref:Power. Although these correspondences have remained as plausible, they are likely to be incorrect.

LGDO and the IFC versions have expected connections with ifc:IfcBuilding being more general than lgdo:BuildingOffice, lgdo:BuildingHospital, lgdo:BuildingSchool, etc. SAREF and IFC have more matches for indoor elements: ifc:IfcDoor with saref:door, ifc:TEMPERATURESENSOR with saref:TemperatureSensor, ifc:IfcSensor with saref:Sensor, etc.

It seems possible to conclude, as expected, that these ontologies are complementary, LGDO extending IFC to the types and outsides of buildings and SAREF extending it to the inside and available appliances. However, the development of SAREF has taken IFC into account, so it is not surprising that there are matches.

Table 4 provides a different perspective by counting only those correspondences which are likely to be correct.

	IFC2x3	IFC4	LGDO
IFC4	2037		
LGDO	27	28	
SAREF	14	22	5

Table 4: Size of curated alignments with a threshold of 0.9 confidence.

Across IFC versions, there are more matches with IFC4 such as ifc:IfcPump (IFC2x3 has only ifc:IfcPumpType) with lgdo:Pump or ifc:IfcTransformer with lgdo:Transformer.

4.4.3 Matching several versions of IFC

The two versions of IFC have most of their concepts and properties in common. Hence, it is important to have very strict thresholds which require strong similarity for declaring them identical. Indeed any similar word (Ramp and Lamp) will now provide two false positive instead of one. And there are many similar words: Door, DoorType, DoorStyle, DoorTypeEnum, etc.For instance, the classes IfcCompressorType and IfcCompressorTypeEnum are both in IFC2x3 and IFC4. Hence, they are found equivalent. However, IFC4 has added new classes IfcCompressor and PredefinedType_of_IfcCompressor. Matchers find some similarity among these elements and the old ones.

This multiplies sources of errors. As shown by the raw figures, however, this does not seem a large problem: out of 2644 correspondences, only 2283 remain after curation.

Moreover, if matchers are clever, they will also find many relationships between the elements of the ontology (subClassOf and not only equivalentClass). But these are only useful for classes which have no correspondences in the other ontology, otherwise, this will be redundant (note that this is the case with other ontologies: all the subsumption correspondences with ifc:IfcBuilding are useless if all these classes are subclasses of lgdo:Building which is equivalent to ifc:IfcBuilding).

For instance, the alignment could find that:

```
ifc4:Control > ifc2x3:WorkControl
```

but this is redundant since:

```
ifc4:WorkControl rdfs:subClassOf ifc4:Control
```

(as well as `ifc2x3:WorkControl rdfs:subClassOf ifc2x3:Control`) and

```
ifc4:WorkControl = ifc2x3:WorkControl
```

Note that the found alignment suggest that:

```
ifc4:IfcController > ifc2x3:IfcFlowController
```

This is not redundant, since all the types of controller in IFC are *not* subsumed by the IfcController class. So, the alignment is certainly wrong.

Hence, it seems better in such case to enforce 1:1 matching as strongly as possible. Since we did not introduce any specifics for ontology versions in our generic workflow, this has been moderately applied during curation. It may be also possible to reduce redundancies. However, very strict threshold matching would certainly miss the spelling corrections. Similarly, 1-1 alignment would miss those cases when a class has been further specialised, or split in two classes, and so should be matched. For instance, ifc:IfcEnergyConversionDevice appears in both IFC2x3 and IFC4. However, its subclasses ifc:IfcTransformer and ifc:IfcTubeBundle only appear in IFC4. The alignment should then contain:

```
ifc:fcEnergyConversionDevice = ifc4:fcEnergyConversionDevice
ifc:fcEnergyConversionDevice > ifc:IfcTransformer
ifc:fcEnergyConversionDevice > ifc:IfcTubeBundle
```

This clearly shows, that it is worth for ontology designers to keep track of the changes made in ontologies and to publish them as an alignment instead of trying to reconstruct these changes a posteriori.

4.5 Trial implementations and example data

Various IFC data sets have been used for testing the data extraction process. The test procedure started with an initial assessment of the IFC file whether all required data is included or if something is missing. This first step was done with help of the IfcViewer tool[25] developed by the Karlsruhe Institute of Technology (KIT). It was done in two ways: (1) by checking the properties of relevant objects like for instance the address of a building, the name of spaces or its connections via door objects and (2) by using the build-in feature for indoor navigation, which essentially was used to see if the path between two randomly selected rooms could be calculated or not. In a next step all required data as defined in the developed mvdXML specification was extracted and exported back to an IFC partial file. This file was then checked again with the IfcViewer tool to see whether the same results as before could be achieved.

This test procedure was leading to the following results:

- An iteratively refined Model View Definition, which in particular was necessary to cover all geometry types used for the definition of the connection geometry.
- Only few datasets cover all required data. Sometimes even main information like spaces are missing. Nevertheless, simple indoor navigation features could be support in most cases and it has been shown that the amount of data can be dramatically reduced. A typical reduction factor of 10 or even more has been achieved with the used test files.

The next chapters describe some of the used test files and a rough assessment of the quality. All examples are based on IFC2x3.

4.5.1 Labyrinth show case from KIT

This test case was developed by the Karlsruhe Institute of Technology (KIT) and was used as reference. This artificial example was modelled by KIT to show the potential of indoor navigation as implemented in their viewer tools. Figure 13 gives an impression of the single storey building with walls arranged as a labyrinth. A special agreement is also visible: Two external spaces have been added to identify the entrance of a building.

Summary:

- Original file size and contained data: 457 kB (only architectural data)
- Reduced file size: 58 kB (factor 8)
- Quality of data: very good (data enables to calculate the space navigation path),
  The show case did not include a building address, but contained some default geolocation in the city centre of Berlin.

---

[25] http://www.iai.fzk.de/www-extern/index.php?id=1138&L=1

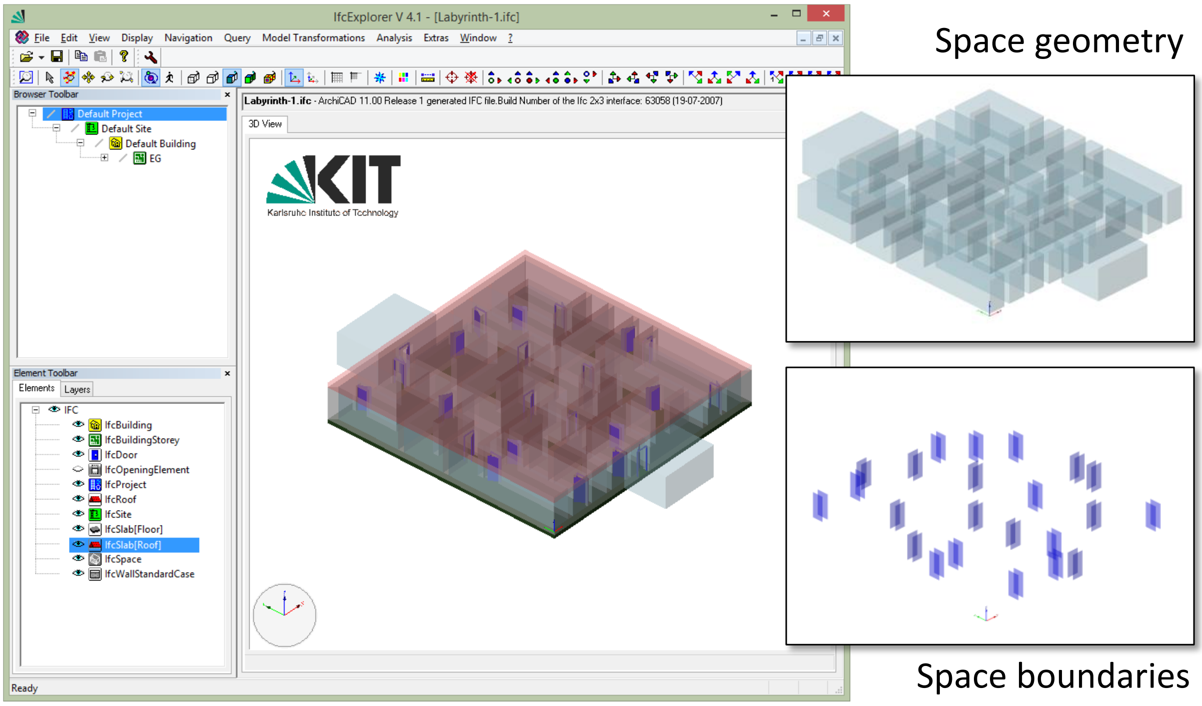


Figure 13: Show case from KIT; left screenshot with all data; right screenshot with extracted space geometry and space boundaries

4.5.2 NIBS Clinic

This example was downloaded from the websites of the National Institute of Building Science[26]. It is a real clinic example located in the south-west of the United States and is available as a set of IFC files covering different aspects of the clinic building. An extension of the MVD was necessary to support navigation between building storeys. The connection geometry is represented by a virtual space boundary which, unlike to the connection via doors, are not linked to a (virtual) element[27]. This is not only leading to a more complex calculation of navigation paths but also to a more complex MVD that increases the size of the partial model. The necessary extension of the MVD was finally leading to three times more data than originally created. In addition to this, stairs are added as elements with full geometry to more easily allow to locate the vertical access system. Figure 14 shows a screenshot of both the whole architectural model as provided on the website and the created partial model.

Summary:

- Original file size and contained data: 17,7 MB (only architectural data)
- Reduced file size: 1,9 MB (factor 9,3)

[26] http://www.nibs.org/?page=bsa_commonbimfiles

[27] Connection between two spaces is typically modelled via a connection element, which is mostly a door or a virtual element. If the connection element is missing, then (in case of a virtual connection) the connection geometry needs to be checked to find the path between two adjacent spaces.

- Quality of data: fair (partially usable for space navigation path),
  Building location covered by geolocation, no building address.

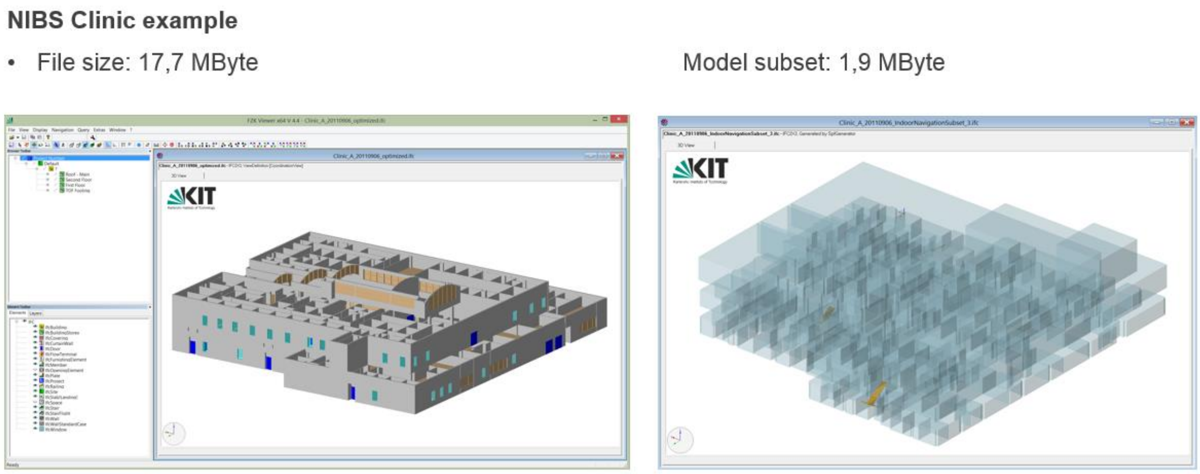


Figure 14: Clinic example provided by NIBS

4.5.3 STREAMER hospitals

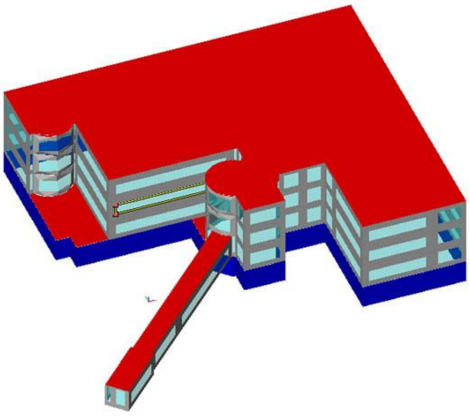
STREAMER is an integrated project funded by the European Commission that is dealing with energy efficient design and retrofitting of hospital districts. Four hospitals have been selected for testing the developed methods, which first of all means to build-up a simple 3D BIM model in order to document the as-is state and the new space layout. Collected data is mainly in an early design stage, but should also be able to fulfil the data requirements for the Indoor Navigation show case. Tests have been conducted with intermediate results for two of the four projects.

Rijnstate Hospital – Arnhem/Netherlands

This demonstration case includes an expansion of an existing building with around 5.000 m². It is modelled on three scale levels to answer different research questions: on (1) building, (2) functional area and (3) room level. For Indoor Navigation the most detailed model on room level has been used (see Figure 15). Tested data, which is still in an intermediate state, enable very basic space navigation due missing door information.

Summary:

- Original file size and contained data: 4,3 MB (only architectural data)
- Reduced file size: not tested due to technical issues
- Quality of data: missing space connection, but spaces and its location available
  Building location covered by geolocation, no building address (wrong address line).



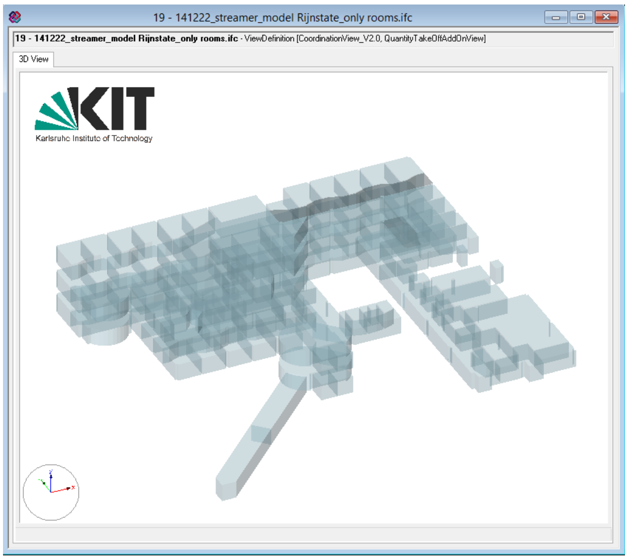


Figure 15: Left: 3D-Model of the Rjinstate hospital (from deliverable D7.2 of the STREAMER project) Right: Available space geometry without door connections

NHS hospital – Rotherham/UK

Two areas from the existing district of over 80.000 m² have been selected for this case study: the Outpatients Department on Level C and the Ophthalmology Ward B6 on Level B. A main challenge is to create a 3D model from drawings and to integrate all relevant information for making an assessment of the current situation and to develop a proposal for an iterative upgrade of those areas. An early version of the created 3D model is shown in Figure 16, which however is not yet on necessary level of detail to support indoor navigation.

Although it is expected that required data will be added soon this example nicely shows that the quality and level of detail of available BIM data is extremely different, if available at all.

Summary:

- Quality of data: indoor navigation not possible due to missing spaces (defines the structure and functional areas of a building only).

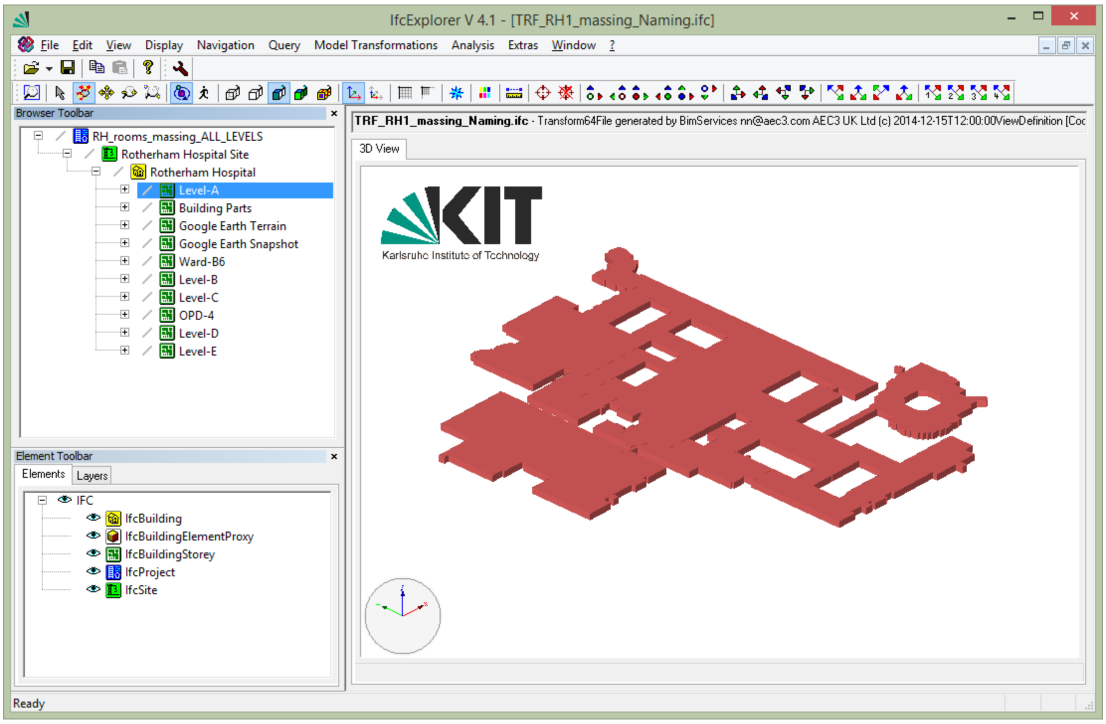


Figure 16: Early version of the 3D IFC model of the Rotherham example.

4.6 Next steps

The main focus has been to extract required IFC data applying the IDM/MVD methodology from buildingSMART. A couple of steps remain to publish, interlink and use the BIM data. The principles are shown in the guidelines developed in work package 4. They have been partially tested with existing tools like the IFC to OWL conversion services from Ghent university and the Fuseki server[28] for hosting the IFC-RDF data. Yet missing is an official ifcOWL standard that is expected in near future. This will provide a reliable basis for development of further services, like for instance a user interface for Indoor Navigation.

Indoor Navigation seems to be of rapidly increasing interest, because a lot of developments are happening in that area. Accordingly, a business case for using parts of BIM datasets is already out there and provision of high quality BIM data becomes a challenge. In general, this is a big chance for the combination of BIM and Linked Open Data to be recognized by the building industry as a valuable technical solution for data publication.

[28] http://jena.apache.org/documentation/serving_data/

Part B: Ontologies and Datasets for Energy Management Systems Interoperability v2

5 Collection of ontologies and datasets

During the second year of the project, the partners involved in the ontologies and dataset collection have followed the approach set in the first year. Therefore, we refer the reader to D2.2 for further details about the project partner involvement. D2.2 also contains detailed information about the different ways used by partners to collect ontologies and datasets, while in the following only updates are included.

- Stakeholder involvement
- Review literature for ontology seeking
- Analysis of Standardization and Institutional Bodies
- Lookup Resource Catalogues

During this second year the ontology catalogue has been presented in the "Building Knowledge Workshop" (Graz, 16th September 2014) and the "5th Workshop on EeB Data Models" (Vienna, 18th September 2014). In both events, a second version of the catalogue was presented including webpages for each ontology or dataset containing details for it and evaluation features for the ontologies. This feature together with other improvement is presented in next section.

5.1 Stakeholder involvement

An **online survey** was set up and launched in March 2014 to enable capturing contributions by the stakeholder community. The idea of the survey is to provide an easy way for stakeholders to take part in the project activities, while also offering the possibility for more experienced stakeholders to provide detailed information. This has been realised by creating two versions of the survey. The first asks stakeholders to only provide the location (URL) of the resource they are aware of, and the follow up research of the resource is done by the project partners. A second survey provides an interface with all information necessary to record an ontology or dataset. If filled by a stakeholder, this information is saved in the database and only needs to be checked by the curator of this database (for the ontology catalogue, this is UPM, Empirica is the curator for the gathered datasets). The survey links will remain active throughout the project lifetime in order to provide a way for new ontologies and datasets to be included. The following links are used for this purpose:

- http://survey.ready4smartcities.eu/index.php/638667/ - short ontology survey
- https://docs.google.com/forms/d/1kTrNUKRnAlN5bBnOwTzQjWwQLinKFQcW4EqXDOYbFsQ/viewform - long ontology survey
- http://survey.ready4smartcities.eu/index.php/162877/ - short dataset survey
- https://docs.google.com/forms/d/1EUISLPLpVHmBaUy2gI76LjE_UPkgPaSW9J1nDruKS0U/viewform - long dataset survey

The target audience for the online survey consisted primarily of stakeholders having access or connected somehow to energy-related data. Such stakeholders were reached through various channels as listed below:

- Mailing list of relevant partners/projects – each partner from the READY4SmartCities consortium shared a number of their partners from other projects based on their background and their relevance to the survey. The mailing list created counted more than 1000 people and was used to introduce the R4SC project and to invite interested people to fill in the survey.

- eeSemantics wiki – CERTH partner is responsible for the maintenance of the eeSemantics wiki, forum and document library on Semantic Interoperability of Energy Efficiency ICT Tools for eeBuildings and beyond and therefore has access to the whole member list of relevant stakeholders (counting more than 500 members). An introduction to the R4SC project and concept was sent, followed by an invitation to participate in the survey, by both a post in the Forum and an email sent to the mailing list.
- READY4SmartCities Portal – the survey was made available and promoted on the R4SC website http://www.ready4smartcities.eu/ and was posted on the website's newsletter.
- Social Networks – the questionnaire invitation was published through the R4SC project's social networks, namely LinkedIn and Twitter, early established in the project.
- VoCamp Participants – during the VoCamps in Germany and Finland, participants with high relevance to energy-related data were approached and were requested to dedicate some time to answer the survey.

5.2 Review literature for ontology seeking

Some of the ontologies included in the READY4SmartCities catalogue[29] have been gathered through the revision of related literature. It is important to mention that the search has been focused on ontologies or vocabularies already implemented in an ontology language, such as RDF and OWL. Thus, when the ontology was only a non-implemented model, such ontology was not taken into account.
The general ontology collection process was:

- UPM read each corresponding document and search for references to ontologies
- When a reference to a relevant ontology is found in the text, two different situations can occur:
    - Such a reference directly leads to a link in which the ontology (implemented in an ontology language) is available. In this case, UPM downloaded the ontology and reviewed the ontology code. After that, UPM acted as catalogue populator by means of providing ontology metadata through the online form.
    - Such a reference is just a textual reference (normally the ontology name). In this case, UPM performed a broad search in the Internet looking for documents about such ontology. When documents were found, UPM started again the general process. On the contrary, UPM had to contact people involved in the ontology development and/or related with such an ontology. UPM directly contacted paper authors, deliverable contributors and/or project coordinators in order to ask for (a) other relevant papers and/or documents in which the ontology is described, (b) information about the ontology files (e.g., if exists, the site in which the ontology is available for downloading), and (c) any other relevant data. However, UPM discovered cases in which it were not possible to contact people (document authors, project coordinators, etc.) involved in the ontology development or related to the ontology building.

      As a result of the contacts conducted, the possible responses obtained were:

        - Confirmation that the ontology is not available on-line, but the ontology file was sent via email
        - Confirmation that there is no ontology implemented
        - Confirmation that the ontology is not public
        - Information about the current status of the ontology development (e.g., the ontology implementation is in progress, our plans includes the development of an ontology).
        - No reply was obtained at the moment of writing this document

The revision of related literature included the following sources:

[29] http://smartcity.linkeddata.es/

- *eeSemantics wiki*[30]. UPM has reviewed pages in the wiki looking for ontologies related to the energy efficiency domain. In particular, pages on the 'Examples and Implementations' and 'eeBuilding Data Models' sections were inspected. In some cases, it was also needed to search for related papers and/or documents. As a result of reviewing this source, five ontologies were included in the catalogue.
- *eeBuilding Data Models workshop proceedings*. Proceedings of this series of workshops were reviewed in order to find related ontologies. The ontologies found in such proceedings were already included in the catalogue while checking other sources.
- *ETSI Smart Appliances workshop report*. The document, D-S1 Interim Study Report, presents a list of existing semantic assets and use case assets, describes their semantic coverage, and proposes an initial semantic mapping. In some cases, it was also needed to search for related papers and/or documents. As a result of the revision of this report one ontology has been included in the catalogue.
- *European project production*. Documents produced within 70 energy-related projects (such as STREAMER, SESAME-S, S4EEB, HYDRA, and SEEMPUBS) have been reviewed. As an outcome of this literature checking, five ontologies were included in the catalogue by UPM acting as a catalogue populator.
- *Other related research literature*. Papers in the area of energy efficiency have been reviewed. UPM included in the catalogue eight ontologies (e.g., DogOnt, ontologies developed in the context of ThinkHome project) found during the inspection of this source.

Finally, it is also important to mention that UPM has checked READY4SmartCitites Deliverable D4.1 in order to include in the catalogue those ontologies mentioned in the described guidelines. In addition, UPM considered useful to have ontologies in the geographical domain, thus literature in such an area was reviewed. The effect of this revision was the inclusion of two ontologies (OGC GeoSPARQL and WGS84 Geo Positioning).

5.3 Review literature for datasets

The datasets included in the READY4SmartCities catalogue have been gathered mainly through desk research, which, however, relates also to surveying related literature sources. It is important to mention that the search has been focused on datasets that are linked and open, i.e. the data should be in RDF. This meant that other datasets which weren't linked or open were not added to the catalogue, they were, however, taken into account specifically for the gap analysis (see section 9.1).

Relevant sources for the datasets came from the expertise of the involved project partners, the survey entries, and suggestions from experts and stakeholders contacted by the consortium as part of WP1 activities. Some of the portals that were pointed as possible sources of information include:

- **Reegle**[31]: the gateway has already established itself as a popular information portal in the fields of renewable energy and energy efficiency. It offers all of its data under W3C standards, i.e. it is open and Linked Data in a non-proprietary format (RDF).
- **OpenEI**: a collaborative knowledge-sharing platform with free and open access to energy- related data, models, tools, and information. OpenEI features over 55,000 content pages, more than 600 downloadable data sets, regional gateways on a variety of energy-related topics, and numerous online tools.
- **Datahub**: this powerful data management platform covering a wide range of topics. It offers data collections, some of which are linked and open.

[30] https://webgate.ec.europa.eu/fpfis/wikis/display/eeSemantics/Home

[31] http://www.reegle.info/

The dataset collection process is similar to the one used to collect ontologies. An identified dataset that meets the requirements of Linked Open Data is added to the catalogue by the dataset curator EMP (only metadata) through the corresponding online form.

5.4 Analysis of Standardization and Institutional Bodies

In general, standardization and institutional bodies are a valuable source of information when it comes to identify agreements for information exchange and reuse of data. Seamless exchange of digital data has been an issue from the very beginning of computer based work and a lot of efforts have already been made to reach consensus between different parties about how to organize and structure shared data. The Open Linked Data Approach based on general web standards like URI, XML, RDF, OWL and SPARQL is a relatively new approach compared to other technologies like SQL, IDEF or STEP-EXPRESS. The main use case of (Open) Linked Data is to publish and interlink pieces of information and thus differs from current exchange and integration approaches. Meanwhile, after several years of research, standardization bodies took notice of this new technology and its potential benefits. While there are still ongoing discussions about use cases and how to position OLD to existing developments, it became clear that both approaches can benefit from each other. On one side there are rich vocabularies, model schemata and business logic developed in many years of standardization work and on the other side there is a new technology to support the web of data with all promised advantages. While our search for ontologies and open datasets published by standardization bodies was not really successful we realized that there are ongoing discussions and preparation work for further standardisation. A short summary of the current situation as well as activities of R4SC towards support actions is given below.

W3C

W3C is seen as the most relevant standardization body for OWL-based ontologies. The partner UPM is active in working groups related to the standardization of different technologies in the W3C. Different ontologies and vocabularies developed in the W3C and widely used were included in the catalogue for representing generic concepts (e.g., time, organizations) and some specific ones (e.g., sensor networks, statistical data). More domain specific W3C standards are currently developed or discussed for instance with support from OGC (Spatial Data on the Web Working Group)[32] or AEC researches (Linked Building Data Community Group)[33].

ETSI

From summer 2013, the European Commission has the intention to launch a standardization exercise at ETSI to propose a high-level model (an ontology) for smart appliances, as an ETSI standard. The first step consists in a pre-normative study that will be done by the Dutch TNO. This project is called “Study on Semantic Assets for Smart Appliances Interoperability” and consists in defining/ identifying a common vocabulary for appliances product information, commands, signals and in a second step agrees on an abstract architecture compatible with the current machine-to-machine (M2M) standards. The outcomes of this study is highly relevant for our project and already ontologies coming from 17 relevant initiatives or project have been translated into Turtle language and are available for download (https://sites.google.com/site/smartappliancesproject/ontologies).

UPM and other project partners participated in the DG CONNECT & ETSI Workshop on Smart M2M Appliances, held in Brussels on 27-28 May 2014. In that workshop, a study on available semantics assets for the interoperability of smart appliances was presented. The document, D-S1 Interim Study Report, presents a list of existing semantic assets and use case assets, describes their semantic coverage, and proposes an initial semantic mapping. We

[32] http://www.w3.org/2014/05/geo-charter

[33] http://www.w3.org/community/lbd/

took into account the ontologies described in that document and, in some cases, we also needed to search for related papers and/or documents

AENOR

UPM is member of the AENOR (the Spanish standardization body) Technical Committee for Smart Cities (CTN 178). For this version of the catalogue a current working draft of a standard on open data for smart cities was analysed in order to search for relevant ontologies.

buildingSMART

buildingSMART is an international non-profit organization that develops open standards for the AEC and FM industry. Since nearly 20 years buildingSMART is pushing the BIM technology. Meanwhile its open IFC standard is supported by all major CAD software tools. AEC3 is very active in this organization and started to facilitate discussions about an ifcOWL standard[34] as a baseline for further developments. The Joint workshop on Linked Data in Architecture and Construction (2nd LDAC Workshop & 6th eeSemantics VoCamp, Espoo/Finland, 26-27 May 2014), co-organised and supported by the Ready4SmartCities project, brought together ontology and AEC experts and was used to discuss two main topics: (1) use case scenarios for linked building data and (2) requirements for a unified ifcOWL representation. Also, it was decided to give feedback to the buildingSMART organization and to facilitate a buildingSMART working group that puts this topic on its agenda.

ISO

ISO is a well-known international standardization body for a broad spectrum of engineering applications. The partner AEC3 is involved in standardization work in the building and construction sector, in particular in publishing the IFC model as an ISO standard (ISO 16739). OWL ontologies are not yet a topic, but there are similarities to XML schema-based definitions. Within the STEP familiy of standards (ISO 10303) the EXPRESS language as used for the IFC specification is defined. For support of XML schema a mapping approach is used that includes a standard mapping configuration that can also be adapted to specific purposes. This approach fits to proposals that have been made by several researchers to transfer the EXPRESS-based IFC model to an OWL representation. These proposals could be a baseline for a general mapping approach that then would allow to map other EXPRESS-based standards to a W3C conformant representation.

Other Standardisation and Institutional Bodies

There are a couple of efforts towards the aim of Ready 4 Smart Cities, e.g. the Energy Performance Buildings Directive from CEN or the draft about a Facility Smart Grid Information Model from ASHRAE. Also, there are a couple of data exchange standards that are relevant in context of smart cities use cases. However, they typically do not make use of the Open Linked Data approach or underlying technologies so that we decided to ignore such efforts for our catalogue or further discussions.

[34] As buildingSMART already publishes a mature, object-oriented data model the strategy from researchers has been to work on mapping proposals from the EXPRESS language to a proper OWL representation of IFC. Depending on use case scenarios and used ontology toolsets there are different flavours for such mapping definitions. Thus, while all available ifcOWL representations are derived from the original IFC specification there is not yet a common agreement within this community which of those should be preferred or the “standard” representation.

5.5 Lookup Resource Catalogues

There are several ontology search engines that UPM has analysed for identifying ontologies that are relevant to READY4SmartCities: Watson[35], Swoogle[36], and Linked Open Vocabularies (LOV)[37].

The main resource used during the ontology catalogue has been LOV as it includes information about creators, maintainers and publishers that are not always included in the ontology encoding nor the documentation associated, if any. As LOV does not cover all the ontologies gathered during this collection process this approach does not ensure to find such metadata for all possible cases.

Another catalogue that UPM analysed was the Collaborative platform Joinup[38]. This platform offers several services that aim to help e-Government professionals share their experience regarding interoperability solutions with each other. Although the vocabularies are not directly related to the energy efficiency or the smart cities domain, UPM considered useful to review ontologies and vocabularies recommended in such a platform. The effect of this inspection was the inclusion of the Registered Organization Vocabulary in the ontology catalogue.

[35] http://watson.kmi.open.ac.uk/

[36] http://swoogle.umbc.edu/

[37] http://lov.okfn.org/dataset/lov/

[38] https://joinup.ec.europa.eu/

6 Recording of ontologies and datasets

The following section present improvement in the ontology catalogue and dataset catalogue respectively in comparison to the version presented in the previous version of this deliverable. Beside such improvement, there are also common features affecting both the ontology and dataset catalogues, namely:

- A SPARQL endpoint containing data in RDF for both catalogues has been made available at http://smartcity.linkeddata.es/sparql.
- A dcat (data catalogue vocabulary)[39] description containing metadata information about both catalogue has been produced and made available at http://smartcity.linkeddata.es/metadata/dcatSmartcities.ttl
- Filtering by domain feature has been added to the index pages for both catalogues.
- Description pages for ontologies and datasets

6.1 Ontology catalogue

This section shows the updates on the ontology catalogue implementation. For the overview and catalogue generation we refer the reader to the previous version of this deliverable. The catalogue of ontologies about smart cities, energy and other related fields can be accessed through a web application available at http://smartcity.linkeddata.es/ontologies/.

As in the previous version and as it is shown in Figure 17, the index catalogue allows visualizing metadata about the listed ontologies. For each ontology, the metadata are shown in the columns: "Syntax", "Domain", and "Natural Language". The values shown in each cell of the table contain different information both represented by text and by colour; for ontology metadata, colours have the following meaning: "plain information" for blue and "unknown" for grey. Furthermore, in addition to the colour, each cell contains detailed information when available.

Apart from ontology metadata, the catalogue presents in the first three columns the quality indicators for the ontologies defined in [Garcia-Castro et al, 2014]: "Online Availability', "Open License", and "Ontology Language". For the quality indicators, colours have the following meaning: "success" for green, "warning" for orange, "danger" for red, and "unknown" for grey.

As in the first version of the catalogue, the values of the "Open License" and "Ontology Language" indicators are taken from the ontology metadata and the evaluation results are stated using colour. For example, in the column "Open License" we can see that the ontologies "Units of Measure (OM)" and "The W3C Organization Ontology" are both published under an open license as the colour of the cell is green, while detailed information about the licenses is also provided. More precisely, these licenses are "CC-BY 3.0" (Creative Commons Attribution 3.0 Unported) and "W3C" (W3C Software Notice and License) respectively.

The "Online Availability" indicator represents whether the ontology is available in the Web in RDF and in HTML format. The evaluation of this indicator is performed on execution time when the catalogue is generated, that is, it is updated every time the catalogue is rebuilt.

[39] http://www.w3.org/TR/vocab-dcat/

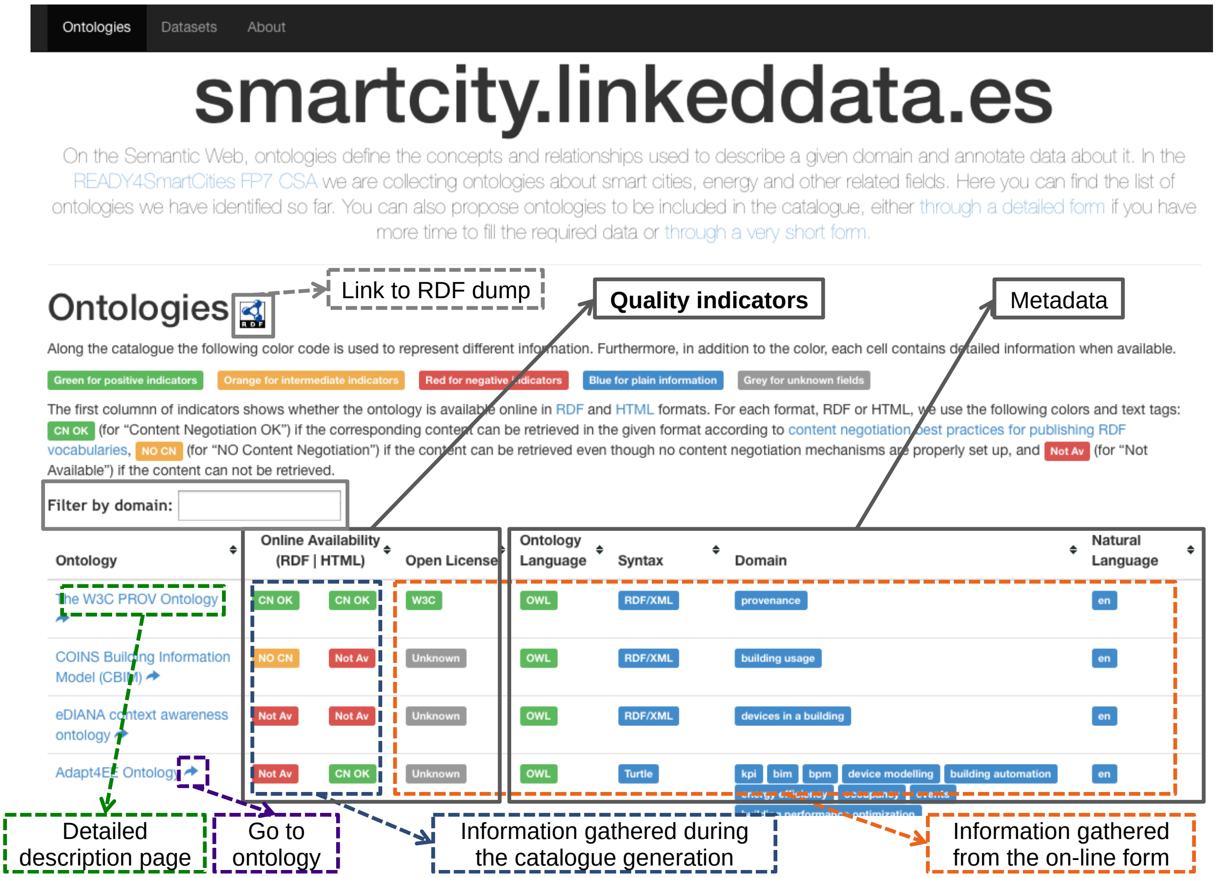


Figure 17. Screenshot of the ontology catalogue home page

In this second year some improvements have been implemented in the catalogue. As Figure 17 shows, the ontology column has been extended so that the link in the ontology title goes to a webpage describing each ontology, while the link arrow near to the name redirects to the ontology itself.

Figure 18 shows an example of a webpage describing a particular ontology. These pages show detailed information about the ontology gathered in the submission form like title, URI, description, languages, ontology languages and formats, issued and modified date, version and license. When the ontology is available and accessible via its URI it is analysed by OOPS! (OntOlogy Pitfall Scanner![40], [Poveda-Villalón et al., 2012]) and the evaluation results are provided in the same webpage as shown in Figure 18. Such evaluation results consist on a list of detected pitfalls (situations that represent an error in ontologies or might lead to errors). For each detected pitfall it is shows its title, how many times it appears, how important is the pitfall (minor, important or critical), a description of what the pitfall consists on and the list of elements affected.

[40] http://oops.linkeddata.es

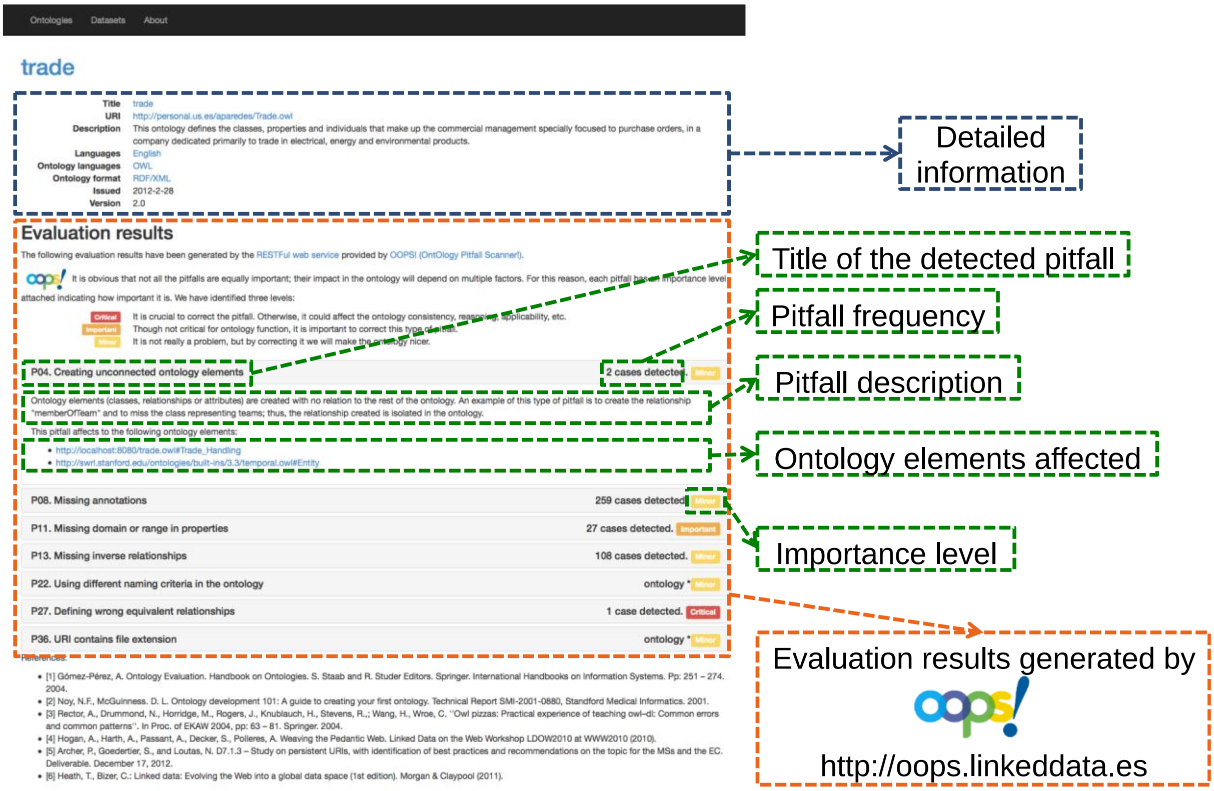


Figure 18. Screenshot of ontology description page example

6.2 Dataset catalogue

This section shows the updates on the dataset catalogue implementation. For the overview and catalogue generation we refer the reader to the previous version of this deliverable. The catalogue of datasets about smart cities, energy and other related fields can be accessed through a web application available at http://smartcity.linkeddata.es/datasets/.

As in the previous version and as it is shown in Figure 19, the catalogue allows visualizing metadata about the listed datasets. For each dataset, the metadata are shown in the columns. More precisely the columns "Digital form", "Publicly available", "Free of charge", "Available online", "Machine readable", "Available in bulk", "Open License" and "Up to date", represent the considered quality indicators as defined in [Garcia-Castro et al, 2014] while the columns "Domain" and "Natural language" provide general information about the dataset. The values shown in each cell of the table contain different information both represented by text and by colour; for ontology metadata, colours have the following meaning: "plain information" for blue and "unknown" for grey. Furthermore, in addition to the colour, each cell contains detailed information when available.

In this second year some improvements have been implemented in the catalogue. As Figure 19 shows, the dataset column has been extended so that the link in the dataset title goes to a webpage describing each dataset, while the link arrow near to the name redirects to the dataset itself.

Figure 18 shows an example of a webpage describing a particular dataset. These pages show detailed information about the dataset gathered in the submission form like its title, description, the domains addressed in the dataset, versioning information, Creation date and last update, contact person, publisher, license, format, language, update frequency, whether it is available online, publicly available, free of charge, available in a machine readable format and available via bulk. Finally, at the bottom of the page, the ontologies used by the dataset are listed. In case the ontology is already included in the catalogue, a link to its specific page is provided following the ontology name with the label "see in the catalogue".

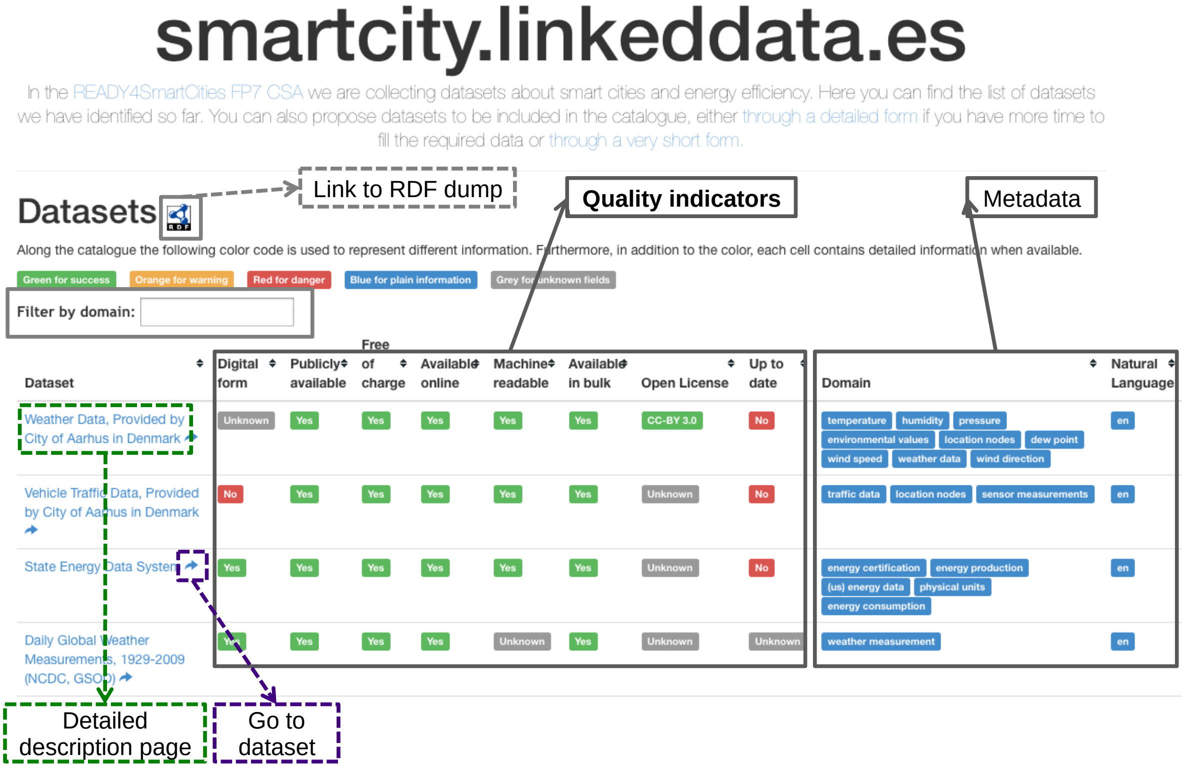
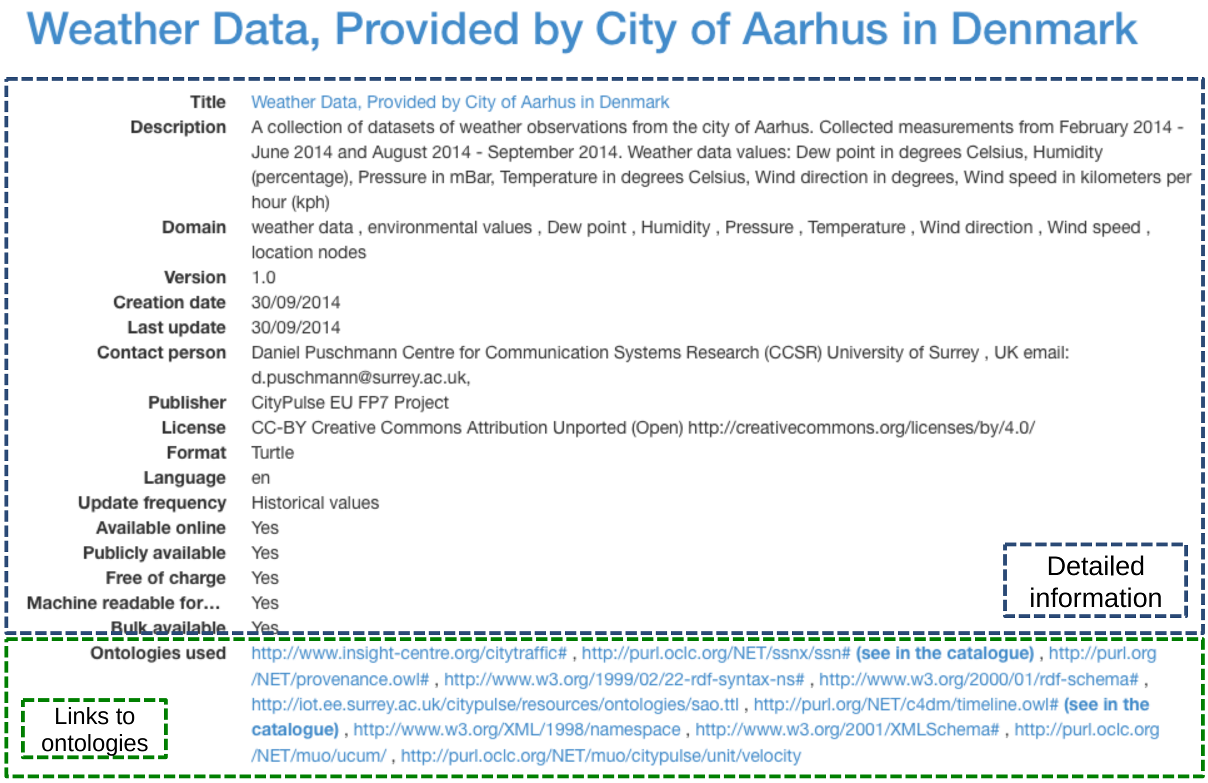


Figure 19. Screenshot of the dataset catalogue home page



Weather Data, Provided by City of Aarhus in Denmark

Title	Weather Data, Provided by City of Aarhus in Denmark
Description	A collection of datasets of weather observations from the city of Aarhus. Collected measurements from February 2014 - June 2014 and August 2014 - September 2014. Weather data values: Dew point in degrees Celsius, Humidity (percentage), Pressure in mBar, Temperature in degrees Celsius, Wind direction in degrees, Wind speed in kilometers per hour (kph)
Domain	weather data , environmental values , Dew point , Humidity , Pressure , Temperature , Wind direction , Wind speed , location nodes
Version	1.0
Creation date	30/09/2014
Last update	30/09/2014
Contact person	Daniel Puschmann Centre for Communication Systems Research (CCSR) University of Surrey , UK email: d.puschmann@surrey.ac.uk,
Publisher	CityPulse EU FP7 Project
License	CC-BY Creative Commons Attribution Unported (Open) http://creativecommons.org/licenses/by/4.0/
Format	Turtle
Language	en
Update frequency	Historical values
Available online	Yes
Publicly available	Yes
Free of charge	Yes
Machine readable for...	Yes
Bulk available	Yes
Ontologies used	http://www.insight-centre.org/citytraffic# , http://purl.oclc.org/NET/ssnx/ssn# **(see in the catalogue)** , http://purl.org/NET/provenance.owl# , http://www.w3.org/1999/02/22-rdf-syntax-ns# , http://www.w3.org/2000/01/rdf-schema# , http://iot.ee.surrey.ac.uk/citypulse/resources/ontologies/sao.ttl , http://purl.org/NET/c4dm/timeline.owl# **(see in the catalogue)** , http://www.w3.org/XML/1998/namespace , http://www.w3.org/2001/XMLSchema# , http://purl.oclc.org/NET/muo/ucum/ , http://purl.oclc.org/NET/muo/citypulse/unit/velocity

Figure 20 Screenshot of dataset description page example

6.3 Alignments catalogue

The alignment catalogue is implemented as an alignment server sharing alignments on the web. Below, we describe briefly the architecture of the alignment server and the methodology used for generating the alignments it contains

6.3.1 Overview of the Alignment server

The Alignment server can supply alignments for people to inspect and for systems to reuse. More than a simple catalogue, it offers the opportunity to generate, organise and manipulate alignments online.

The goal of the Alignment server is that different actors can share available alignments and methods for finding alignments. Such a server enables to match ontologies, store the resulting alignment, store manually provided alignments, extract merger, transformer, mediators from those alignments.

The Alignment server is built around the Alignment API. It thus provides access to all the features of this API. The server architecture is made of three layers:

- **A storage system** providing persistent storage and retrieval of alignments. It implements only basic storage and runtime memory caching functions. The storage is made through a DBMS interface and can be replaced by any database management system as soon as it is supported by jdbc.
- **A protocol manager** which handles the server protocol. It accepts the queries from plug-in interfaces and uses the server resources for answering them. It uses the storage system for caching results.
- **Protocol plugs-in** which accept incoming queries in a particular communication system and invoke the protocol manager in order to satisfy them. These plugs-in are ideally stateless and only translator for the external queries.

This infrastructure is able to store and retrieve alignments as well as providing them on the fly. We call it an infrastructure because it will be shared by the applications using ontologies on the semantic web. However, it may be seen as a directory or a service by web services, as an agent by agents, as a library in ambient computing applications, etc.

Services that are provided by the Alignment server are:

- storing alignments, whether they are provided by automatic means or by hand;
- storing annotations in order for the clients to evaluate alignments and to decide to use one of them or to start from it (this starts with the information about the matching algorithms, the justifications for correspondences that can be used in agent argumentation, as well as properties of the alignment);
- producing alignments on the fly through various algorithms that can be extended and parametrised;
- manipulating alignments by inverting them, applying thresholds;
- generating knowledge processors such as mediators, transformations, translators, rules as well as to process these processors if necessary;
- finding similar ontologies and contacting other such services in order to ask them for operations that the current service cannot provide by itself.

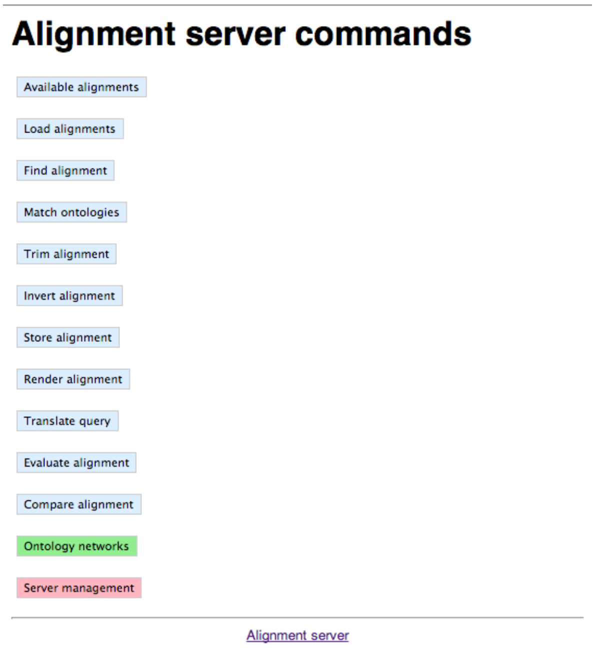


Figure 21. Menu of the services provided through the Alignment server

The menu of these services through the HTML plug-in is seen on Figure 21. For Ready4SmartCities, we introduced in the server the notion of ontology network which group together a set of ontologies and a set of alignments for better visibility.

The server is accessible from the ontology catalogue (but currently not the other way around because the ontologies refer only to their actual URI).

This section serves as a documentation for the connection between the ontology catalogue and the Alignment server. The main point would be that it is possible to link these. This has to be performed through web services call invocation. We describe here the REST interface, however a SOAP interface is also available.

There are two main ways which can be used to connect the Ontology catalogue to the Alignment server.

The ontology catalogue provides for each ontology access to the alignments mentioning it in the Alignment server. This is achieved by generating a URL such as:

http://al4sc.inrialpes.fr/html/listalignments?uri1=http%3A%2F%2Fwww.geonames.org%2Fontology&uri2=all

This redirects to the list of all alignments involving the geoname ontology as shown in the following figure:

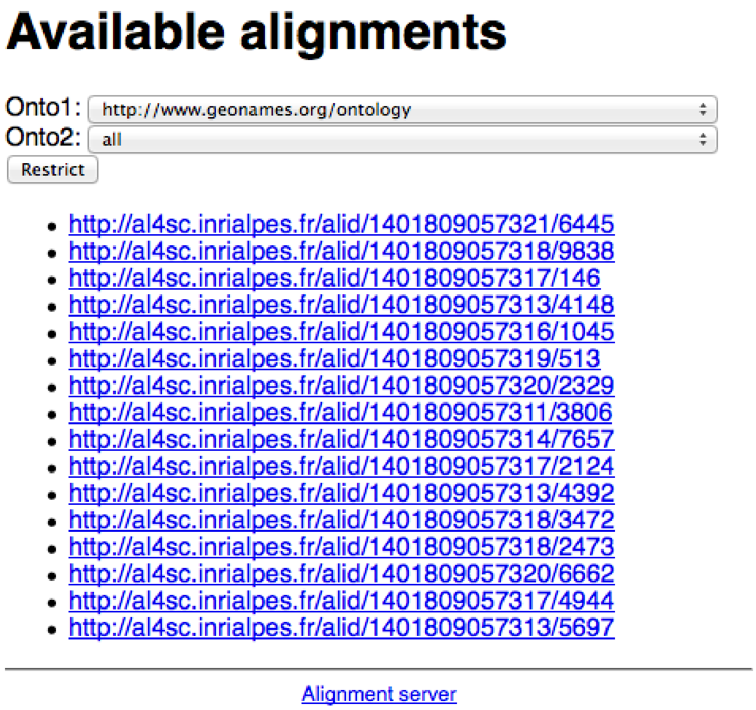


Figure 22. List of alignments involving the geoname ontology

In counterpart, each alignment description features two annotations (cat1 and cat2) which refer to the URLs of each ontology in the catalogue.

6.3.2 Methodology of alignment generation

The generation of the network of alignments for the Alignment server has been spread on the two years of the project. In 2014, a network with a core of 10 ontologies has been generated. In 2015, a network involving 42 ontologies has been generated filling largely the gap of missing alignments.

Figure 23 describes the adopted methodology spanning the two years.

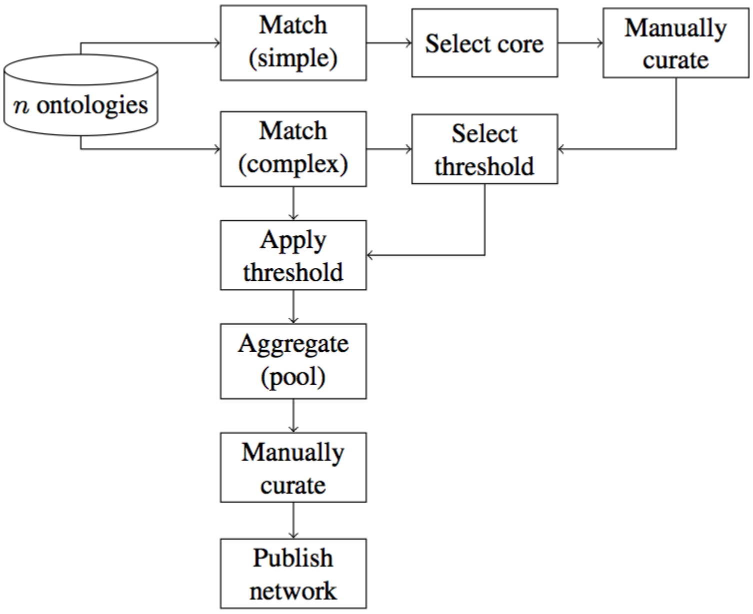


Figure 23. Workflow of the alignment generation and curation process

The methodology has taken the following steps:

- match all (simple): The ontology are matched on the basis of string equivalence.
- select core: A subset of ontologies is selected by taking the most connected ontologies.
- match core (average): The core ontologies are matched with basic string-based matchers (SMOA and EDNA).
- select threshold: A threshold on the alignments so generated is chosen so they generate only 33% additional correspondences in addition to the simple matchers.
- apply threshold: The selected threshold is applied to the alignments and they are merged.
- manually curate: The resulting alignments are manually curated (this was performed in 2014).
- match (complex): The ontologies are matched with a larger panel of matching system, including Aroma and LogMap.
- aggregate: For each pair of ontologies, all the alignments between this pair are aggregated in a single alignment containing all their correspondences with a confidence corresponding to the proportion of matchers which have found it.
- manually curate: The whole network is manually curated by using systematic confidence levels
- publish network: The network is published on the Alignment server. This results in alignments generated by the process of Figure 24.

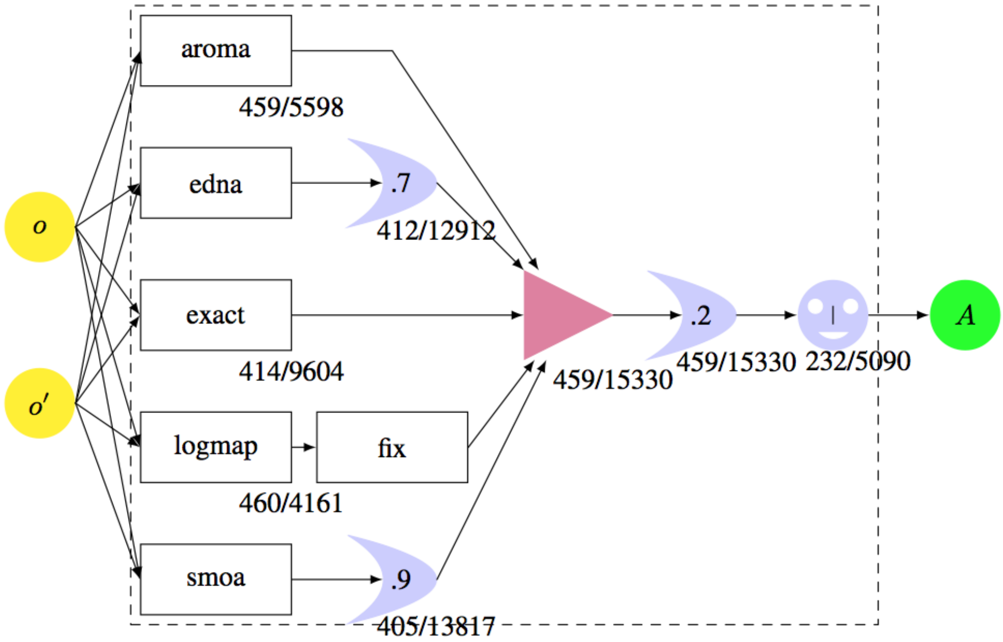


Figure 24. Process which actually led to the ontologies between alignments in the style of [Euzenat and Shvaiko, 2013].

The curation process has been rationalised by standardising the confidence measures associated with the correspondence. We reproduce below the table, provided to the curator, giving the semantics of confidence measures:

confidence	explaination
1.	Most certainly correct correspondence
.8	Certainly correct correspondence, but there may be some subtle difference
.7	Likely correct correspondence
.6	To check, with some chance to be correct
.5	To check
.4	To check, but not surprised if incorrect
.2	Unlikely correspondence, but who knows
.0	Most certainly incorrect correspondence (such correspondences are discarded) .

6.4 Overview of ontologies and datasets gathered during the first and second project years

6.4.1 Ontologies, vocabularies and standards

General overview of the Ontology Catalogue

- At the moment of writing this deliverable, the Ready4SmartCities Ontology Catalogue contained **70 ontologies**.
- UPM analysed these ontologies in order to provide a general overview of the ontology languages and format used, the natural languages in which ontologies are expressed, and the licenses attached to these ontologies.
- INRIA performed a content analysis covering other relevant aspects

The most common ontology language in the Ready4SmartCities Catalogue is **OWL**, followed by RDF-S. 65 ontologies are implemented in OWL, while only 3 ontologies are also coded in RDF-S, finally 1 ontology is coded only in RDF-S and 1 ontology is represented in SKOS. The distribution of ontology languages in the catalogue is shown in Figure 25. It is worth mentioning that five ontologies are in more than one ontology language. These ontologies are Timeline Ontology, Data Cube and Stream Annotation Ontology.

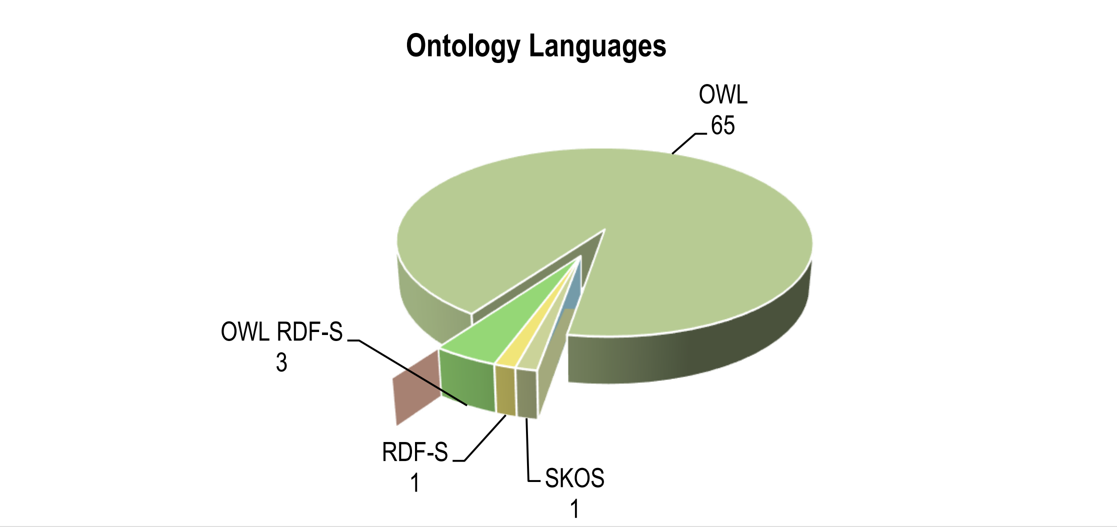


Figure 25. Ontology languages distribution

Regarding the ontology syntaxes, **RDF/XML** is the most usual one followed by Turtle. 51 ontologies are written using the RDF/XML syntax among other formats, while 20 are using the Turtle syntax within their serializations. As in the case of ontology languages, there are 8 ontologies in the catalogue provided with more than one format. These ontologies are Km4city, Units of Measure (OM), Measurement Ontology, The W3C Organization Ontology, IFC2X3 **-** University of Ghent, Places Ontology, Registered Organization Vocabulary, Stream Annotation Ontology - SAO. It is important to mention that for 3 ontologies the ontology syntax is not known. The distribution of ontology formats in the catalogue is shown in Figure 26.

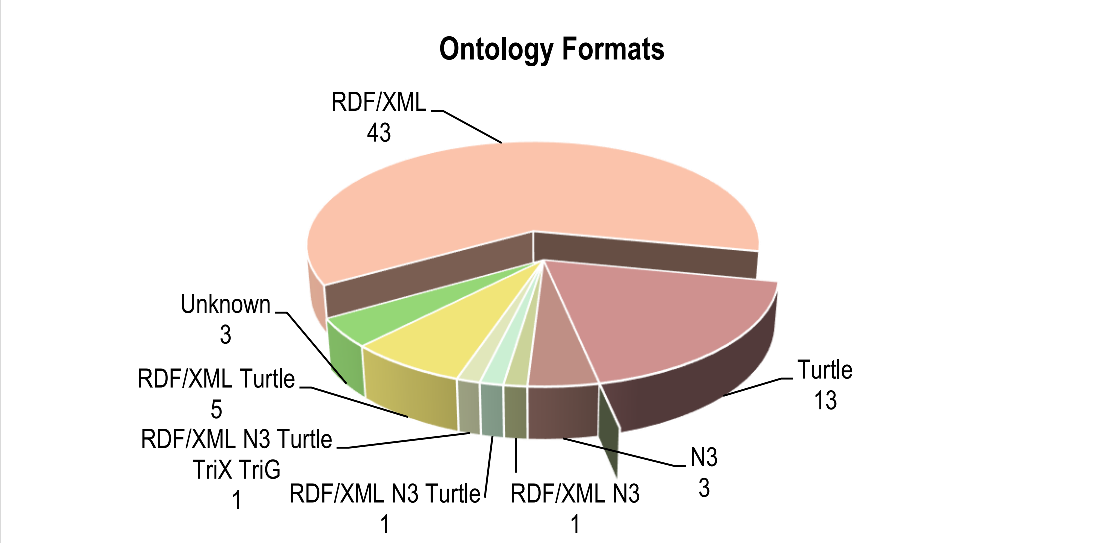


Figure 26. Ontology formats distribution

With respect to the natural language used for naming ontology elements, the most common one is **English** (67 ontologies are written at least in such a language). There are 3 ontologies which natural language is 'Unknwon'[41]. Apart from English, that might be consider the base language, there appear other languages for the mulitilingual ontologies, namely: Bulgarian-bg, Czech-cs, German-de (2 ontologies), Spanish-es (2 ontologies), French-fr (2 ontologies), Hungarian-hu, Italian-it (3 ontologies), Dutch-nl, Norwegian-no, Polish-pl, Romanian-ro, Russian-ru, Slovak-sk and Swedish-sv.There are seven ontologies in the catalogue that are written in more than one natural language. These ontologies are Geonames, Units of Measure (OM), The W3C Organization Ontology, DUL (DOLCE+DnS Ultralite), URBAMET Ontology, Eurobau Utility Ontology, and FreeClassOWL Ontology. The distribution of natural languages used in the catalogue is shown in Figure 27.

[41] This situation occurs because the ontology documentation does not provide information about the natural language used. In addition, the code for those ontologies was not available, so it was not possible to discovery the language used for naming ontology elements.

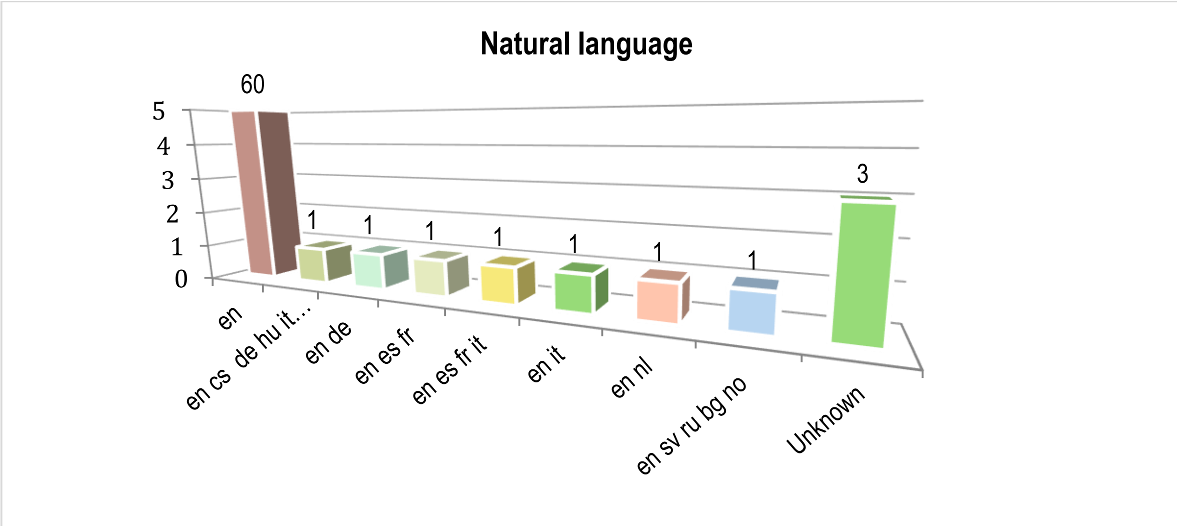


Figure 27. Distribution of natural languages in ontologies

Most of the ontologies (45 out of 70) in the catalogue have no information about licenses (ontology license is **Unknown**). In those cases in which authors provide license information, the most usual licenses are the CC-BY 3.0 Creative Commons Attribution Unported (8 ontologies have this type of license) and W3C software license (another 6 ontologies have this kind of license). The distribution of ontology licenses in the catalogue is shown in Figure 28. Such figure also shows that most of the licenses when available are open, more precisely among the 25 specified licenses, there is 1 ad-hoc license, 3 no open licenses set as “all rights reserved”, and 21 open licenses of different types.

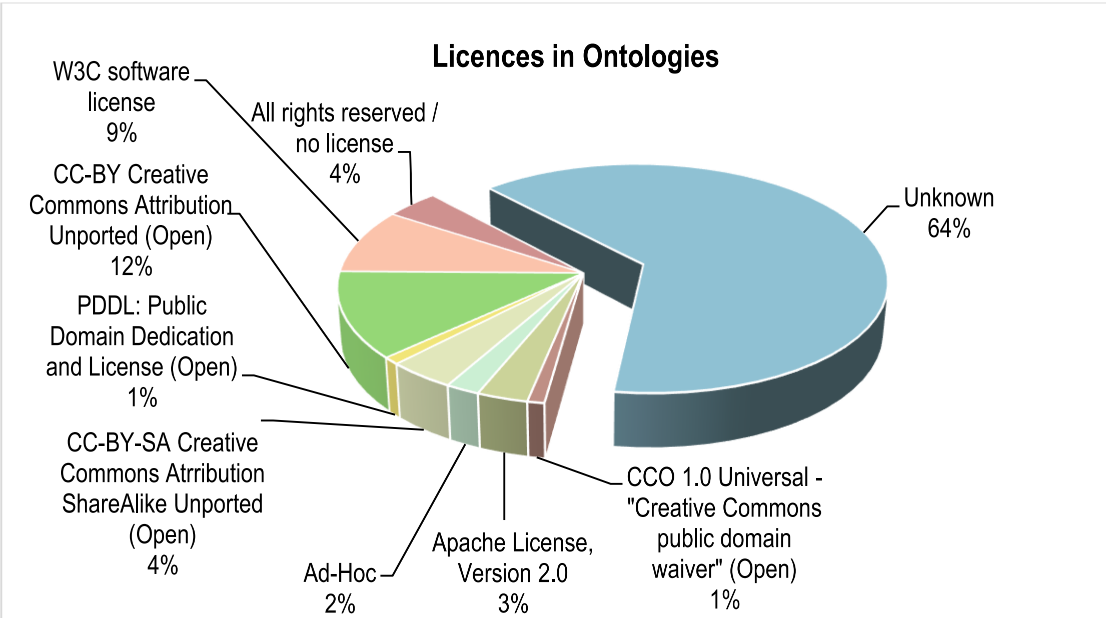


Figure 28. Ontology licenses distribution

UPM also analyzed the 70 ontologies in the catalogue with respect to the following quality indicators: online availability of ontologies and open license attached to the ontologies.

Regarding the online availability of ontologies, UPM performed two analyses: the first one refers to the availability of ontology code (RDF) and the second one refers to the availability of ontology documentation (HTML). In both cases[42] the study refers to:

- whether the corresponding content (RDF or HTML) can be retrieved in the given format according to content negotiation best practices for publishing RDF vocabularies ("Content Negotiation")
- whether the content can be retrieved even though no content negotiation mechanisms are properly set up ("No Content Negotiation")
- whether the content can not be retrieved ("Not Available")
- other situations[43] ("Unknown")

In the first case, **32 out of 70 ontologies can be retrieved in RDF**. However, 22 out of these 32 are retrieved although content negotiation mechanisms have not been properly set up. In addition, 4 ontologies cannot be retrieved in RDF and 6 probably are not available or are published in a wrong way. The distribution of RDF availability in the catalogue is shown in Figure 29.

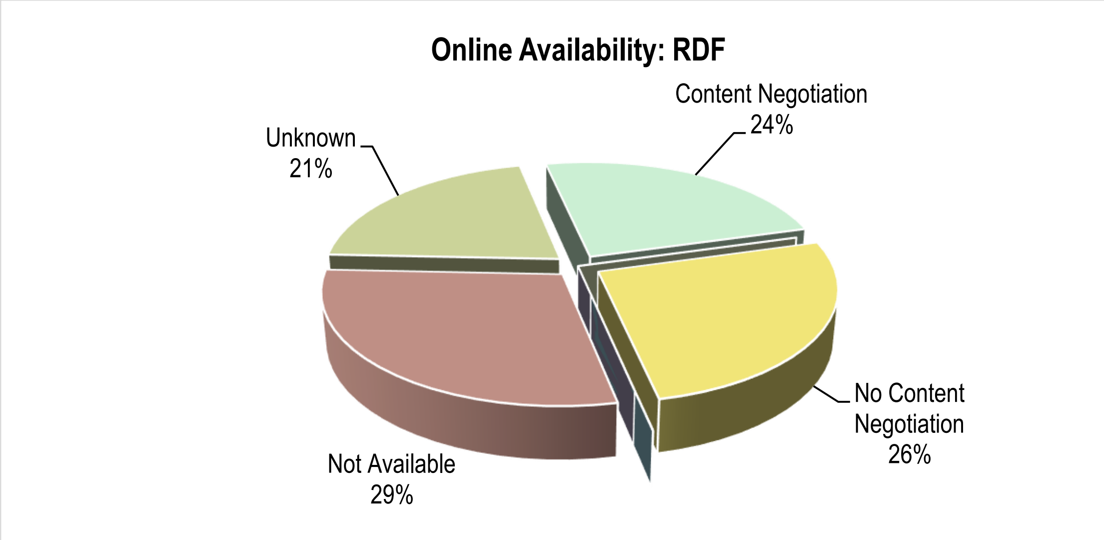


Figure 29. Distribution of RDF availability

In the second case, **26 out of 70 ontologies can be retrieved in HTML**; 15 of them have property content negotiation mechanism implemented. 34 ontologies cannot be retrieved in HTML and 10 probably are not available or are published in a wrong way. The distribution of HTML availability in the catalogue is shown in Figure 30.

[42] In order to check content negotiation mechanisms for RDF and HTML formats, the linked data validator Vapour (http://validator.linkeddata.org/vapour) is used while the RDF content of the available ontologies are loaded in a JENA (http://jena.apache.org/) model.

[43] This means that Vapour provides an exception.

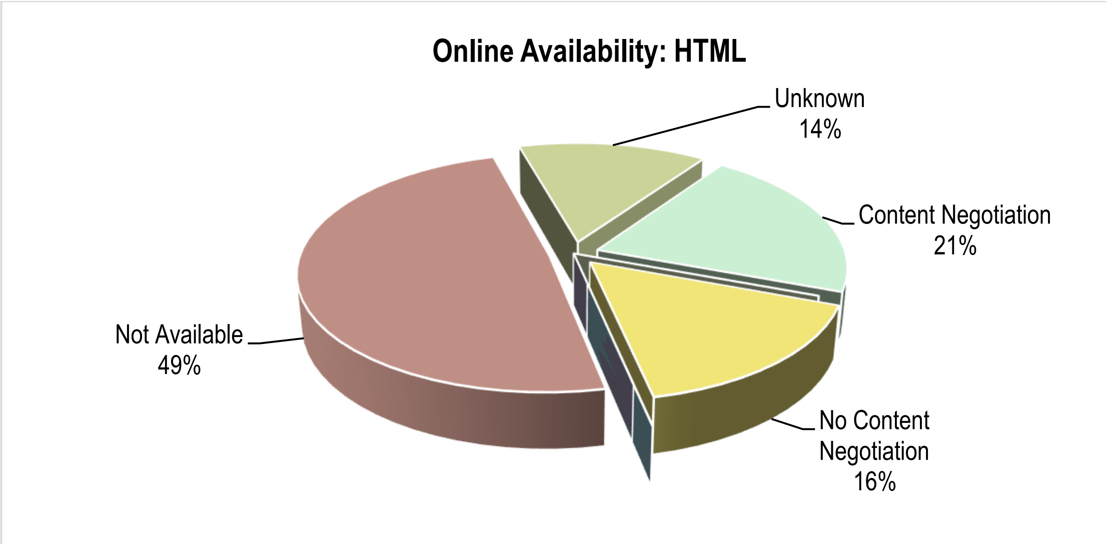


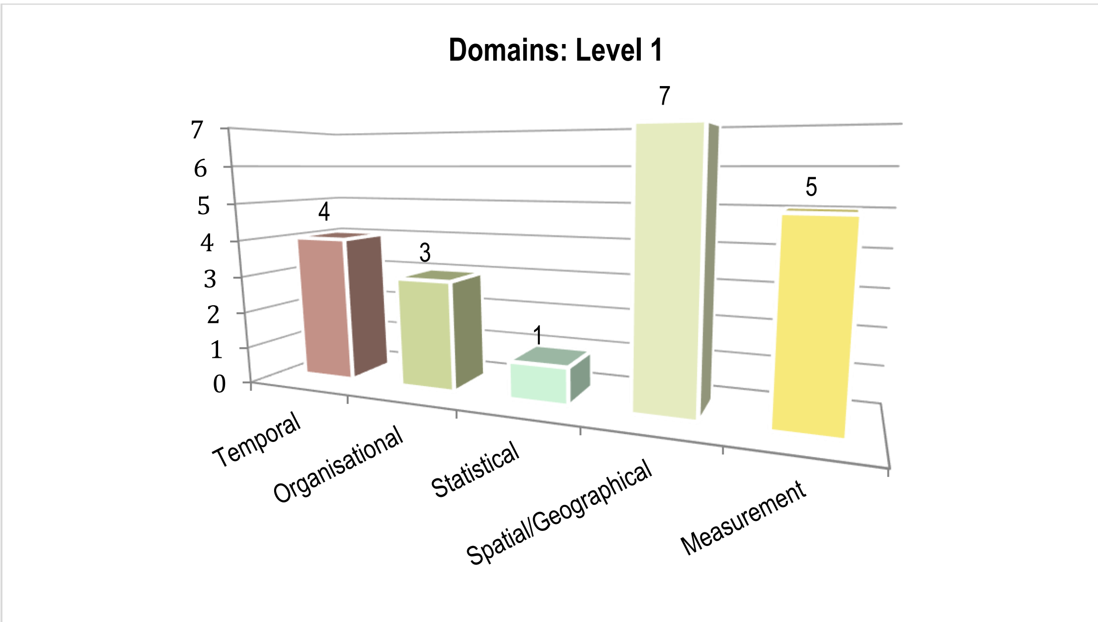
Figure 30. Distribution of HTML availability

Domain coverage analysis

Regarding the specific domains identified in Deliverable D3.1, at first the set of ontologies in the catalogue covers

- the ***five domains identified for Level 1***, that is, Temporal, Organisational, Statistical, Spatial/Geographical, and Measurement
- ***3 out 7 domains identified for Level 2***. These domains are Energy, Weather, and Building. Thus, Climate Zone, Environmental, Occupancy, and User Behaviour do not seem to be covered.

Total figures of ontologies related with Level 1 domains and with Level 2 domains are shown respectively in Figure 31 and Figure 32.



13

Figure 31. Number of ontologies in Level 1 domains

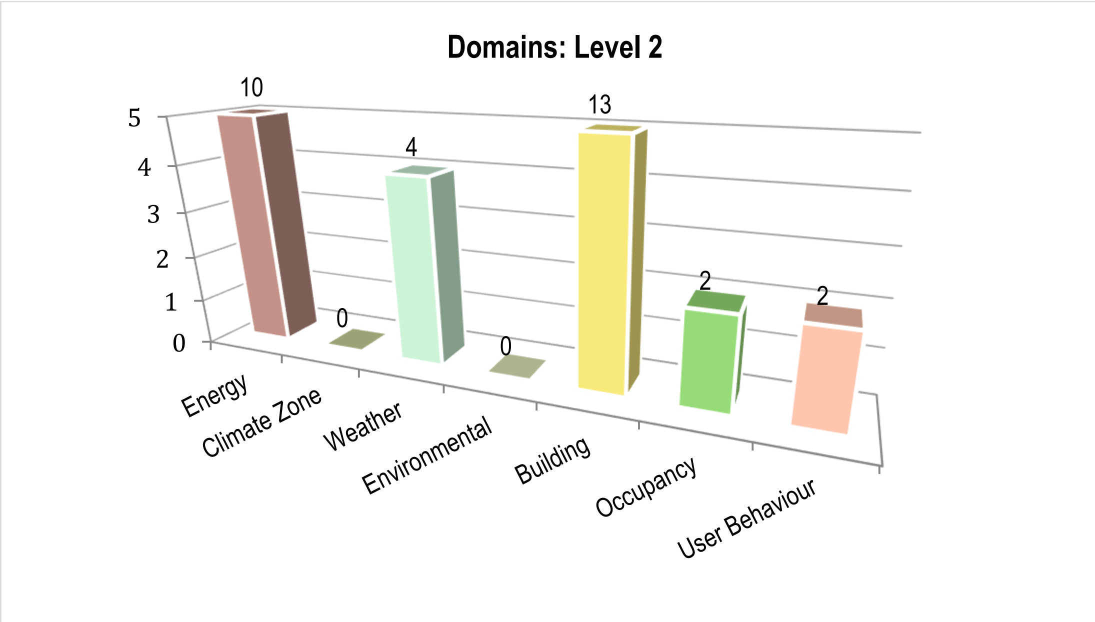


Figure 32. Number of ontologies in Level 2 domains

UPM also analyzed the list of domains attached to the ontologies by catalogue populators. As a result of this analysis, 16 new domains were identified. The full list of domains found and the number of ontologies in which they appear are shown in Table 4.

.

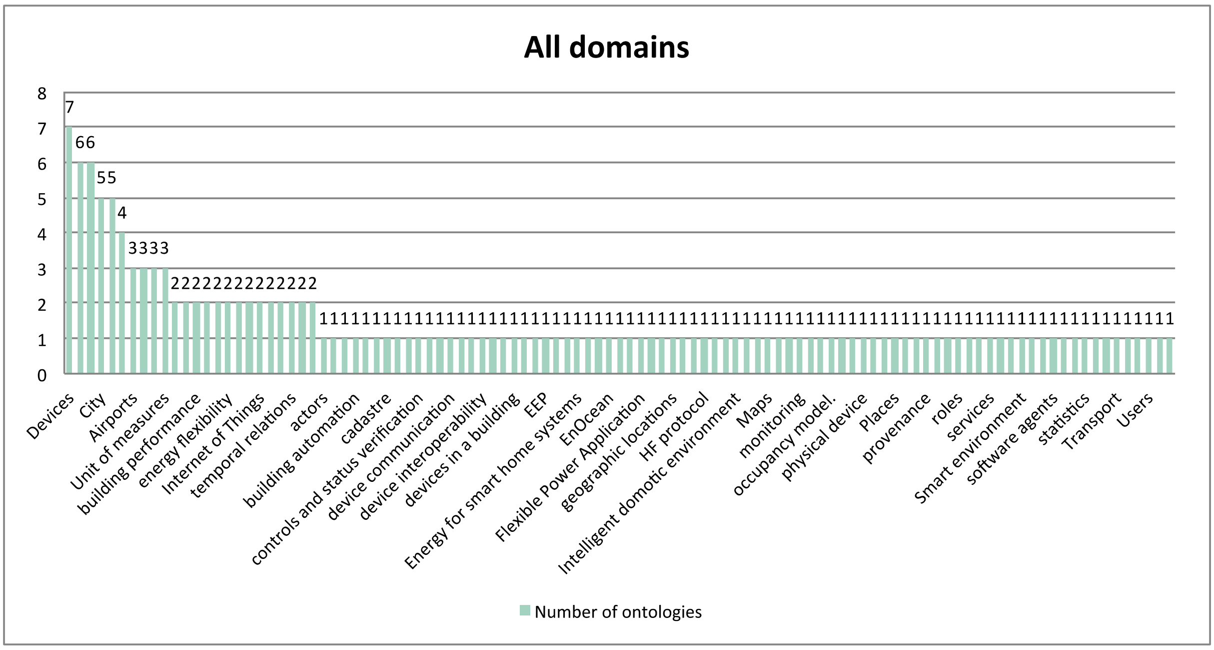


Figure 33. Number of ontologies in new domains

6.4.2 Datasets

At the moment of writing this deliverable, the Ready4SmartCities Dataset Catalogue contained twenty datasets. In the following, a summary of the main characteristics of the datasets is presented.

The datasets cover the domains *building design and measurement*, *building operation*, *outcome metrics*, and *weather and climate data, energy, housing market, location, traffic, parking* and *pollution*,

For six datasets no license has been given (unknown); the datasets with a license include *CC-BY-SA Creative Commons Attribution-ShareAlike Unported (Open)*, CC-BY Crative Commons Atttribution unported (Open), UK Open Govermment Licence (*OGL)* and *PDD* as shown in Figure 34

The format of the datasets is usually N triples and RDF as Figure 35 depicts. Out of the twenty datasets in the catalogue, seven have been recorded as originating from a European project. Two of the datasets are not available in bulk.

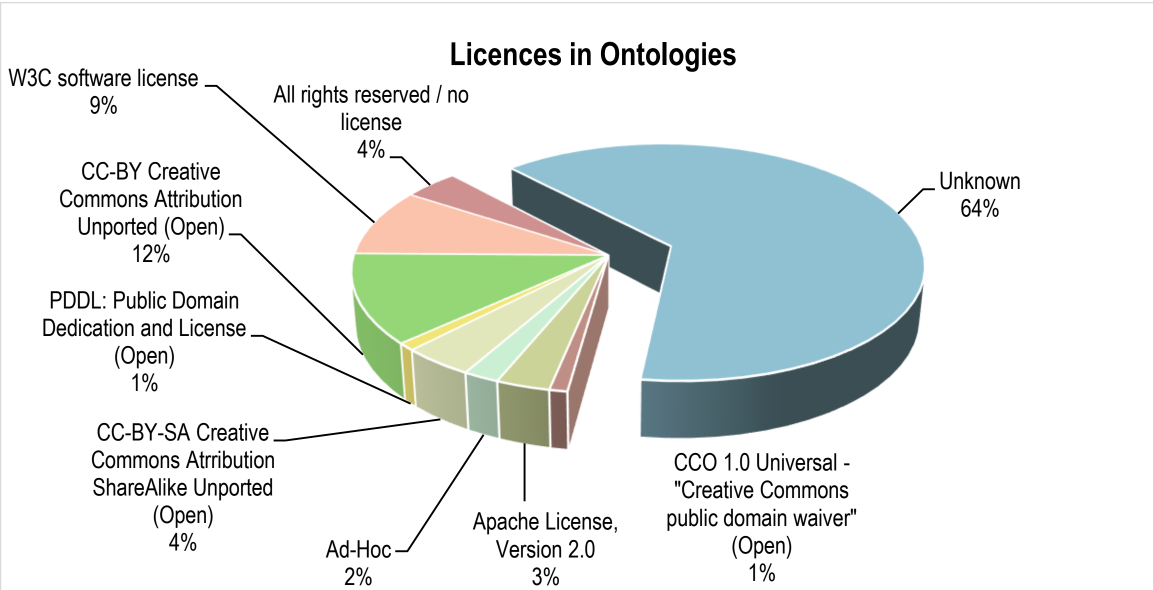


Figure 34. Dataset licenses distribution

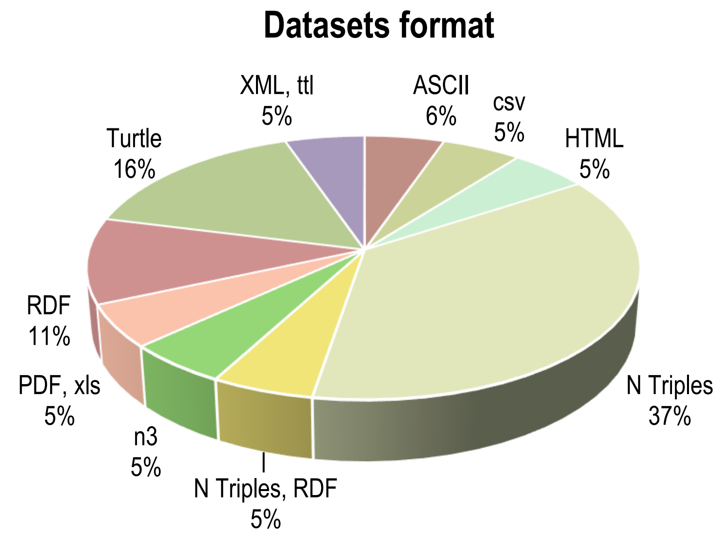


Figure 35. Datasets formats provided

6.4.3 Ontology alignments

As expected as the beginning of the project, there are not many alignments available. Some stakeholders told us that such alignments were part of their proprietary ontologies. Isolating and sharing alignments, however, has the benefit that it can be adopted and improved by others.

So, we take the active step of trying to obtain alignments from the ontologies themselves as described in Section 6.3. We review here the result of this process as available in the Alignment server.

The following table provides the list and statistics about the ontologies which are matched in the Alignment server. These ontologies come from the R4SC ontology catalog. We simply selected those which were available at the moment we started. But we discovered that some of these ontologies were importing or referring to other relevant ontologies, so we included these as well. Finally, during the process three ontologies were considered as worthwhile additions, so they have been included as well. Hence, the following table is organised in four parts:

- the 10 core ontologies identified in 2014,
- the available ontologies end of 2014,
- the additional ontologies, and
- the referred ontologies.

	Nickname	#Class	#Properties	#Instances	#C+P+I (glob)	#Triples	Imports
2014 core	bfo	39	0	0	48	429	
	bonsai	54	72	1	216	845	CoDAMoS, services
	cose	138	58	34	235	690	
	dog	763	79	0	848	10625	
	dolce	37	70	0	107	752	
	dul	75	109	0	188	1807	
	energyresource	523	110	16	723	9331	time
	ifc2x3	952	948	0	1901	14807	
	semanco	978	851	21	10280	55961	sumo
	um	933	56	1700	2689	28000	
2014 catalog	actor	36	20	6	132	837	time
	building	31	5	0	36	227	
	cmo	2	4	0	8	21	qudt
	cube	15	21	1	37	266	
	geonames	7	29	699	735	6846	
	measurement	6	8	2	16	95	
	muo	10	17	0	27	127	
	org	9	35	1	45	748	
	places	47	4	0	51	566	
	process	156	38	6	268	2985	time
	rov	1	6	0	7	86	
	ssn	41	39	0	278	2370	dul
	time	11	41	14	67	297	
	timeline	22	43	1	211	1262	time, time-entry
	trade	141	106	34	302	1645	temporal
	weather	106	33	11	223	3124	time
	wgs84	0	5	0	5	33	
new	ifc4	1221	1576	1624	4421	38178	
	lgdo	1200	222	0	1422	24530	
	saref	110	42	73	310	1382	wgs, time
Extra	codamos	40	37	0	77	217	
	dbpedia	*					
	foaf	15	68	13	96	631	
	geography	2	18	0	20	80	
	geovocab	1	16	0	17	138	
	goodrelations	37	112	47	196	1834	
	qudt	199	72	0	271	1992	
	schema	295	184	0	479	3495	
	scovo	3	5	0	8	84	
	sumo	4525	802	85644	90971	587998	
	temporal	7	5	7	19	82	
	time-entry	17	48	7	72	312	

From these ontologies the following table provides the total number of alignments, non empty alignments and correspondences provided by each method on the whole network or the core ontologies.

	whole network			core network (raw)		
	#al	#ne-al	#corresp	#al	#ne-al	#corresp
aroma	429	204	4048	45	41	1267
edna .7	411	233	8021	45	40	1105
exact	410	161	6307	45	33	606
logmap	427	263	3377	46	38	634
smoa .9	403	239	5862	45	40	1268
Total	2080	1100	27615	226	192	4480
Aggregated	457	348	10342	45	45	2497

Once aggregated, these alignments have generated 10342 correspondences distributed in 348 alignments. The 10342 correspondences of these alignments where curated by hand, as described in Section 6.3. The final result of this process is a network of alignments containing 5786 correspondences in 317 alignments.

	#al	#corresp
Initially	348	10342
Global	317	5786
IFC4-ifc2x3	1	2283
without IFC4-ifc2x3	316	3503

As explained in part A of this deliverable, one alignment between the two versions of IFC, contains 2283 correspondences. The list of the 21 largest alignments in the network given below shows a quick decrease of the size of alignments.

#corresp	alignment
2283	IFC4-ifc2.rdf
473	dog-energyresource.rdf
179	IFC4-um.rdf
125	ifc2-um.rdf
76	saref-energyresource.rdf
73	saref-dog.rdf
71	IFC4-saref.rdf
70	IFC4-semanco.rdf
70	dul-ssn.rdf
67	IFC4-dog.rdf
66	IFC4-energyresource.rdf
57	lgdo-ifc2
55	IFC4-cose.rdf
54	energyresource-semanco.rdf
51	lgdo-IFC4
46	IFC4-dul.rdf
42	saref-ifc2.rdf
42	semanco-um.rdf
41	cose-dog.rdf
36	ifc2-semanco.rdf

This is further confirmed by Figure 25 (in which the IFC4-ifc2x3 alignment is not taken into account) which shows the long tail shape of the distribution of alignments along their size.

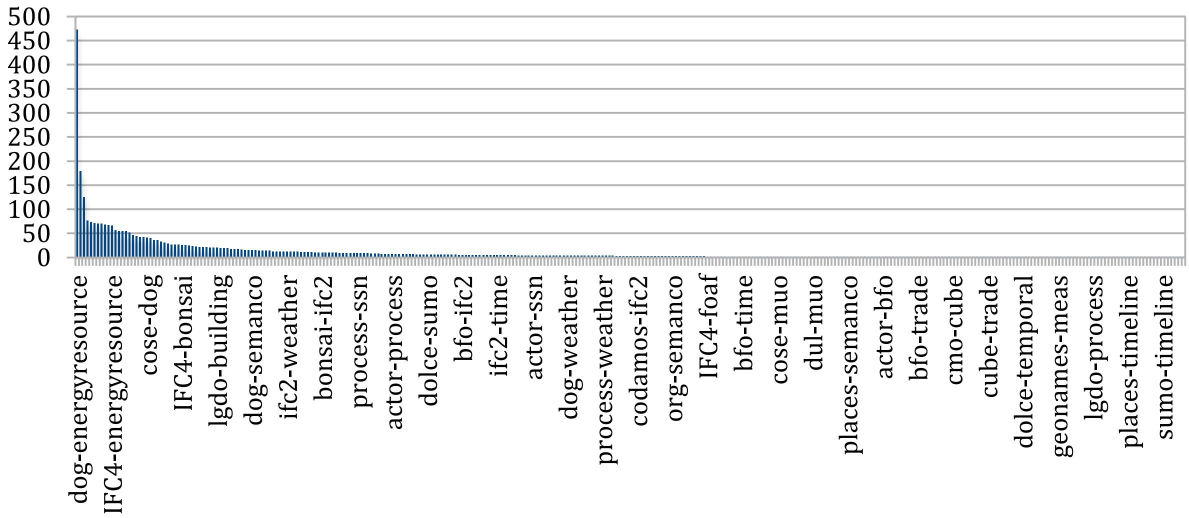


Figure 36. The long tail distribution of the alignments according to their size

In order to have an idea of the quality of the resulting alignments, we compared the raw alignments after aggregation and the final curated alignments with those obtained last year on the core ontologies. The results are reported in the following table. The measures used for comparing them (unlabelled) are precision, recall and F-measure, usually used in information retrieval and for evaluating alignments. We also report the respective size of R, the alignment obtained in 2014 and A the first and final alignments.

As expected, the first step does preserve all the 888 correspondences obtained last year and produces 1609 new correspondences. Hence, the recall attains 100% while the precision is low at 36%. After curation, the balance is reinstated with a 67% precision which gives a 75% F-measure. After curation, 144 correspondences are missing and 732 new correspondences have been added to the alignments. On the missing correspondences, 98 come from the alignment between dog and energy resources and this calls for more inspection.

These figures, however, should not be taken too seriously as the alignments of 2014 are not a paramount reference. They simply shows that the two processes have provided results which are largely commensurate.

It is necessary to have more scrutiny by experts of the domain.

Ontology	2014	first (2015)						second (2015)					
	$\|R\|$				$\|A \cap R\|$	$\|A\|$	$\|R \setminus A\|$				$\|A \cap R\|$	$\|A\|$	$\|R \setminus A\|$
bfo-bonsai	-	-	-	-	0	11	0	-	-	-	0	0	0
bfo-cose	-	-	-	-	0	9	0	-	-	-	0	2	0
bfo-dog	-	-	-	-	0	18	0	-	-	-	0	2	0
bfo-dolce	4	.21	.35	1.0	4	19	0	.66	.8	1.0	4	6	0
bfo-dul	9	.45	.62	1.0	9	20	0	.66	.76	.88	8	12	1
bfo-energyresource	-	-	-	-	0	12	0	-	-	-	0	2	0
bfo-ifc2	4	.66	.8	1.0	4	6	0	.6	.67	.75	3	5	1
bfo-semanco	-	-	-	-	0	23	0	-	-	-	0	1	0
bfo-um	-	-	-	-	0	11	0	-	-	-	0	2	0
bonsai-cose	6	.29	.44	1.0	6	21	0	.62	.71	.83	5	8	1
bonsai-dog	10	.21	.34	1.0	10	48	0	.48	.64	1.0	10	21	0
bonsai-dolce	2	.18	.31	1.0	2	11	0	.5	.67	1.0	2	4	0
bonsai-dul	8	.4	.57	1.0	8	20	0	.78	.82	.87	7	9	1
bonsai-energyresource	7	.13	.23	1.0	7	53	0	.29	.45	1.0	7	24	0
bonsai-ifc2	6	.21	.35	1.0	6	28	0	.6	.75	1.0	6	10	0
bonsai-semanco	2	.05	.09	1.0	2	40	0	.11	.18	.5	1	9	1
bonsai-um	11	.42	.59	1.0	11	26	0	1.0	1.0	1.0	11	11	0
cose-dog	20	.27	.42	1.0	20	75	0	.49	.66	1.0	20	41	0
cose-dolce	1	.11	.20	1.0	1	9	0	.33	.5	1.0	1	3	0
cose-dul	8	.5	.67	1.0	8	16	0	.88	.94	1.0	8	9	0
cose-energyresource	15	.27	.43	1.0	15	55	0	.45	.61	.93	14	31	1
cose-ifc2	13	.28	.43	1.0	13	47	0	.58	.69	.85	11	19	2
cose-semanco	8	.18	.31	1.0	8	44	0	.47	.61	.87	7	15	1
cose-um	12	.33	.5	1.0	12	36	0	.44	.61	1.0	12	27	0
dog-dolce	1	.04	.08	1.0	1	24	0	.2	.33	1.0	1	5	0
dog-dul	1	.04	.08	1.0	1	25	0	0.0	0.0	0.0	0	1	1
dog-energyresource	524	.81	.89	1.0	524	649	0	.9	.86	.82	426	473	98
dog-ifc2	8	.18	.3	1.0	8	45	0	.47	.64	1.0	8	17	0
dog-semanco	10	.09	.17	1.0	10	107	0	.66	.8	1.0	10	15	0
dog-um	1	.01	.02	1.0	1	105	0	.5	.67	1.0	1	2	0
dolce-dul	20	.57	.73	1.0	20	35	0	.88	.81	.75	15	17	5
dolce-energyresource	2	.09	.17	1.0	2	21	0	1.0	.67	.5	1	1	1
dolce-ifc2	2	.29	.44	1.0	2	7	0	.2	.29	.5	1	5	1
dolce-semanco	1	.04	.07	1.0	1	26	0	0.0	0.0	0.0	0	1	1
dolce-um	-	-	-	-	0	8	0	-	-	-	0	1	0
dul-energyresource	1	.04	.07	1.0	1	27	0	.2	.33	1.0	1	5	0
dul-ifc2	17	.29	.45	1.0	17	59	0	.64	.76	.94	16	25	1
dul-semanco	3	.11	.2	1.0	3	27	0	.6	.75	1.0	3	5	0
dul-um	2	.09	.16	1.0	2	23	0	.67	.8	1.0	2	3	0
energyresource-ifc2	8	.17	.29	1.0	8	48	0	.5	.64	.87	7	14	1
energyresource-semanco	16	.13	.23	1.0	16	122	0	.3	.46	1.0	16	54	0
energyresource-um	1	.01	.03	1.0	1	73	0	.25	.4	1.0	1	4	0
ifc2-semanco	21	.23	.37	1.0	21	93	0	.55	.19	.95	20	36	1
ifc2-um	90	.4	.57	1.0	90	227	0	.52	.6	.72	65	125	25
semanco-um	13	.15	.26	1.0	13	88	0	.31	.47	1.0	13	42	0
Overall	888	.36	.52	1.0	888	2497	0	.67	.75	.84	748	1120	144

For a more qualitative insight, we reproduce below the table produced in previous deliverable for the ifc2x3-bonsai alignment:

IFC2x3	Bonsai	SMOA .9	EDNA .7	Observation
parameter	parameter	1.0	1.0	
IfcBuilding	Building	1.0	1.0	
frequency	frequency	1.0	1.0	
IfcPoint	Point	1.0	1.0	
values	value	.97	.83	?
mode	Model	.97	.8	#
IfcActuatorType	Actuator	.94		hasType
inputPhase	hasInput	.94		?
IfcCondition	AirCondition	.93	.75	>
ParameterList0	parameter	.93		?
IfcParameterValue	parameter	.93		hasValue
IfcServiceLife	Service	.93		?
IfcSensorType	Sensor	.92		hasType
IfcBuildingStorey	Building	.92		< ?
IfcPressureMeasure	Pressure	.91		
pointParameter	parameter	.91		< ?
IfcBuildingElement	Building	.91		isPartOf
rateDateTime	dateTime	.9		< ?
IfcActuatorTypeEnum	Actuator	.9		

Clearly, the four first correspondences seem to be correct, then half of the supplementary correspondences. EDNA thresholded at .7 finds fewer correspondences (13) which are, in general, less meaningful.

This can be compared with the final result for the same two ontologies in the full network of ontologies.

IFC2x3	Bonsai	Relation	Confidence	Observation
parameter	parameter	=	1.0	
frequency	frequency	=	1.0	
IfcPoint	Point	=	0.8	They are penalised because of the Ifc prefix
IfcBuilding	Building	=	0.8	
parameter	hasParameter	=	0.8	
values	value	=	0.8	

IfcPressureMeasure	Pressure	=	0.7	
IfcCondition	AirCondition	>	0.5	
inputPhase	hasInput	<	0.2	
panelOperation	hasOperation	<	0.2	

The four first correspondences are still there and have been consolidated. The policy penalises the matches with Ifc prefix because it is impossible to know if they are here for a good reason or not (maybe IfcBuilding is a particular type of building). Only two correspondences are new and many hasardous correspondences have been discarded in the final alignment.

7 The Interoperability Areas: Energy Management Systems and Energy Measurement and Validation

The domains covered in work packages two and three come from two main application areas which have common aspects that not only allow to follow the same methodology within both work packages but also to share a lot of resources in terms of ontologies, datasets and alignments. There is no clear borderline as one may expect, which finally led to the decision to have a single point of information for the catalogues. Nevertheless, there are important differences between the two application areas that are described below. However, using linked data we expect that both application areas will more and more converge in the future, which will lead to more robust and flexible solutions for both application areas. In order to pinpoint the common areas, the two tables below should provide the scope of work in both packages. The first table tries to characterise and compare both application areas, whereas the second table shows typical domains covered by work package 3.

WP2 is reviewing the linked data situation for Energy Management Systems (EMS). In general, EMS has a very broad scope and includes a lot of domains and stakeholders that depend on each other and must interact in order to be able to control and monitor energy production and consumption of electro-mechanical facilities. For several reasons it was decided in WP2 to first focus on the construction sector, which not only is a major energy consumer with high potentials for energy savings and peak energy balancing but it is also an energy producer and even a way for energy storage. There are a lot of use cases for smart cities that directly or indirectly relate to buildings, e.g. prediction of energy demands (based on the heating, cooling and lighting demands of buildings that is also linked to user behaviours) or traffic management (for e.g. travelling between office and residential areas). Also, the construction industry is an interesting environment for testing and promoting the linked data approach as there are many different stakeholders that must collaborate and share information.

WP3 addresses the need to validate the results of energy-efficiency actions by analysing their measured impact. Measuring consumption in smart cities provides the source of data to be validated (including measurement methods, predictive models and algorithms), but other factors also play a role in the analysis, such as weather and climate data, building characteristics, user behaviour, etc. Measurement and validation requires complete terminology for experimentation and piloting including experimental group, control group, statistical significance, outcome metrics (key performance indicators, KPIs), modelling parameters (e.g. occupancy, comfort levels, meteorology, etc.).

The ontologies and datasets described in the next sections have been selected because they address one or more of the topics work packages 2 or 3 focus on. Concerning alignments, their generation in a nearly blind way already allows for clustering ontologies and identifying clusters of ontologies related to these topics.

Table 5. Application areas of the domains in work packages 2 and 3

	Energy Management System (WP2)	Energy Measurement and Validation (WP3)
Main application area	Controlling a “single” electro-mechanical system either for energy production or energy consumption, automation of systems (machine-to-machine communication)	Measure and validate energy consumption and/or production to provide key figures for strategic and operative decisions, decision support and awareness services
Characteristics of used data		
degree of standardization	Medium	Low
degree of structured data	Very high	Medium
degree of complexity	High	Medium
degree of openness	Very low (outside of the “system” environment) Medium (within the “system”, if different players must work together)	Medium to High
fault tolerance	Low to very low	Medium
security requirements	Very high	Low to medium
amount of data	Medium to high	Very high
real-time requirements	Medium to very high	Low to medium

In total 70 ontologies and 18 datasets have been identified and catalogued. An overview can be seen in Table 6 and Table 7. For more results, see the gap analysis and the list of ontologies and datasets.

Table 6. Overview of ontologies identified in the project categorised in domains

	Metrics and indicators	Methods of measurement	Predictive models / Energy analysis	User behaviour	Building design and refurbishment	Monitoring	Controlling	Optimizing performance	Building operation	GIS	Systems: BACS, BEMS	Groups (experimental, control)	Statistics	Outcome metrics (KPIs)	Modelling parameters	Piloting	Organisation	Energy data	Weather and climate data	Environmental data	Upper Ontologies	Measurement	Time	Devices/Sensors	Provenance
Architecture and Building Physics Information					■																				
The W3C Organization Ontology																■	■								
IFC2X3 - University of Ghent					■																				
IFC2X3 - NIST OntoSTEP Converter					■																				
The W3C Time Ontology	■	■																					■		
BFO (Basic Formal Ontology)																					■				
Weather and Exterior Influence Information							■								■				■	■			■		
Units of Measure (OM)	■																					■	■		
Measurement Ontology		■																				■			
Users and Preference Information				■																					

Energy and Resource Information
MUO - Measurement Units Ontology
Casas Ontology for Smart Environments (COSE)
DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering)
DUL (DOLCE+DnS Ultralite)
Timeline Ontology
SESAME-S Smart Building Ontology
Simulation Information Model (SIM) Ontology
Performance Information Model (PIM) Ontology
The W3C SemanticThe W3C Sensor Network Ontology
Building Information Model (BIM) Ontology
Global City Indicator Foundation Ontology
User Behavior and Building Process Information
Cadastre and Land Administration Thesaurus (CaLAThe)
CASCADE airport ontology
Nikola Tesla Airport (NTA) Ontology

trade							■																		
Geonames					■																				
Data Cube													■												
The W3C PROV Ontology																									■
DogOnt																								■	
SUMO (Suggested Upper Merged Ontology)																					■				
BOnSAI																								■	
OGC GeoSPARQL										■															
WGS84 Geo Positioning		■								■															
Open Street Map (OSM) ontology	■									■															
Places Ontology										■															
eDIANA context awareness ontology																								■	
Urban Energy Ontology					■					■															
Concept Modelling Ontology (CMO)																					■				
Registered Organization Vocabulary																	■								

The Event Ontology																						■			
km4city	■											■					■	■	■				■		
Internet of Things (IoT) Ontology						■									■								■		
OpenIoT Ontology																							■		
SPITFIRE Ontology																							■		
Eurobau Utility Ontology					■				■		■														
FreeClassOWL Ontology					■																				
CERISE CIM Profile for Smart Grids																					■				
COINS Building Information Model (CBIM)										■															
CASCADE Fiumicino Airport ontology								■	■														■		
CASCADE Malpensa Airport ontology								■	■														■		
Energy in Buildings Ontology						■			■																
INERTIA Ontology																							■		
INSPIRE Data Specification on Transport Networks																					■				
CityGML Ontology										■															

URBAMET Ontology
SAREF: The Smart Appliances REFerence ontology
DECT ULE ontology
Echonet ontology
EnOcean ontology
FAN FPAI ontology
FIEMSER ontology
FIPA Device Ontology
Hydra Basic Device Information ontology
SmartCoDE ontology
Mirabel ontology
Stream Annotation Ontology - SAO
Adapt4EE Ontology
ROUTE - Ontology of Urban Transportation Entities

Table 7. Overview of datasets identified in the project categorised in domains

	Metrics and indicators	Methods of measurement	Predictive models / Energy analysis	User behaviour	Building design and refurbishment	Monitoring	Controlling	Optimizing performance	Building operation	GIS	Systems: BACS, BEMS	Groups (experimental, control)	Statistics	Outcome metrics (KPIs)	Modelling parameters	Piloting	Organisation	Energy data	Weather and climate data	Environmental data	Upper Ontologies	Measurement	Time	Devices/Sensors	Provenance
Eurobau database					X				X																
Daily Global Weather Measurements, 1929-2009 (NCDC, GSOD)																			X						
Repener building energy	X																	X							
Enipedia Energy Industry Data	X									X								X							
Linked Clean Energy Data																		X	X						
State Energy Data System	X																	X							
Energy efficiency assessments and improvements														X											
Residential Energy Consumption Survey						X			X																

Housing Market Indicators
INERTIA Ontology dataset instance
Number of dwellings by tenure and district in the UK
Impact indicator: energy efficiency of new build housing in the UK
Vehicle Traffic Data, Provided by City of Aarhus in Denmark
Parking Data Stream, Provided by City of Aarhus in Denmark
Pollution Data, Provided by City of Aarhus in Denmark
Weather Data, Provided by City of Aarhus in Denmark
Energy time-series mapping from University of Southampton
Linked geodata dataset

8 Collected ontologies

8.1 Gap analysis

The version of the Ready4SmartCities Ontology Catalogue current at the time of writing contained **70 ontologies.** According to relevant domains, the current set of ontologies in the Ready4SmartCities catalogue covers the 5 domains identified for Level 1 (Temporal, Organisational, Statistical, Spatial/Geographical, and Measurement) and 4 out 7 domains identified for Level 2 (Energy, Weather, Building, and User Behaviour). Thus, there are three domains for which there are no ontologies in the catalogue, namely, Climate Zone, Environmental, and Occupancy. It is worth mentioning that 16 additional domains are also covered.

Regarding the ontology language, 74% of the ontologies in the catalogue are implemented in OWL, one of the most common languages for developing ontologies. Only three ontologies are implemented using more than one ontology language; these are Timeline Ontology, Stream Annotation Ontology - SAO and Data Cube, which are implemented in OWL and RDF-S. In order to benefit the interoperability and the usability of ontologies in different contexts, it could be beneficial to have more ontologies both in OWL and in RDF-S.

With respect to the syntaxes or formats for ontologies, 71'43% of them are provided in RDF/XML and 28.5% of these ontologies are in Turtle (it is worth noting that some ontologies are given in RDF/SML and also turtle). There are eight ontologies provided in more than one format

67 ontologies in the catalogue are written in English, which is the most common natural language in research tasks. Currently, there are seven ontologies specially written in more than one natural language; namely Geonames, Units of Measure (OM), The W3C Organization Ontology, DUL (DOLCE+DnS Ultralite), URBAMET Ontology, Eurobau Utility Ontology, and FreeClassOWL Ontology. The natural languages used in such ontolgoires are Bulgarian, Czech, German, Spanish, French, Hungarian, Italian, Dutch, Norwegian-, Polish, Romanian, Russian, Slovak and Swedish. Since multilingualism is a key issue, the catalogue should include more ontologies written in different languages.

A not really good point in the catalogue is that only open licenses are attached to those ontologies with license information. Regarding ontologies 64% of them provide no information about licensing.

With respect to the online availability of the ontologies in the catalogue, half of the ontologies can be retrieved in RDF. However, 26% of the ontologies do not have content negotiation mechanisms properly set up for this format and 29% cannot be retrieved in RDF. This situation should be corrected. Regarding HTML availability, 37% of the ontologies can be retrieved in such a format. However, 49% of the ontologies cannot be retrieved in HTML, which normally provides ontology documentation. Thus, in order to benefit the understanding and reuse of the ontologies, this situation should be also improved.

In addition, it is worth mentioning that in some cases the negotiation mechanisms seem to be good established, however the retrieved content does not correspond with the expected ones. This occurs when the ontology URI follows the pattern "www.owl-ontologies.com/" or contains only names (e.g., "CityEnergyInvestmentStudy"). This situation should also be corrected.

8.2 List of ontologies

Architecture and Building Physics Information

Name	Architecture and Building Physics Information
Author and License	Institute of Computer Aided Automation, Vienna University, Austria unknown license
URL	https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/BuildingOntologySharedVocabulary.owl
Description	An ontology representing building information (e.g. structure, material, architecture) for Smart Home Systems. Classes, axioms and customized datatypes have been retrieved from gbXML (www.gbxml.org). (for further information see: https://www.auto.tuwien.ac.at/projectsites/thinkhome/building-information.html). The mapping from gbXML is done through an XSLT script, which is also available on the website.
Scope (Domain)	Buildings, Energy Analysis
Use cases (Motivation, Relevance)	There are many use cases for smart cities where building data is of relevance. gbXML data is typically used for energy analysis, which is done in the design phase of a building.
Data sets	gbXML datasets can be generated and imported by many CAD and energy analysis tools. However, these tools export a XML file according to the gbXML schema definition and thus has to be mapped to an RDF representation according to this ontology. Sample gbXML files are available at www.gbxml.org. However, building data is typically not published as it is mainly shared within the design team only or are handed over to contractual partners.
Open issues/ Challenges	There is an agreed schema developed by the gbXML initiative. This is the baseline definition from which this ontology was derived based on an XSLT script.
Tool support	Population of the ontology through mapping approaches from traditional CAD tools.

The W3C Organization Ontology

Name	The W3C Organization Ontology
Author and License	Dave Reynolds, Epimorphics Ltd. W3C license
URL	www.w3.org/ns/org#

Description	Vocabulary for describing organizational structures, specializable to a broad variety of types of organization.
Scope (Domain)	Organization, Piloting
Use cases (Motivation, Relevance)	The motivation for creating the ontology was seen in the need to publish information relating to government organizational structure as part of the data.gov.uk initiative. The approach chosen was to develop a small, generic, reusable core ontology for organizational information and then let developers extend and specialize it to particular domains. In the energy domain, the ontology can be used and extended to describe organisations and sites that partake in energy-related projects, e.g. piloting innovative solutions that save energy, developing and testing new technologies like smart metres, etc.
Data sets	Based on the listed implementation of the ontology, it has been used in domains such as healthcare and public organisations (universities, libraries, museums), but not in the energy domain. No datasets could be found thus far that use the ontology.
Open issues/ Challenges	
Tool support	

IFC Ontology

Name	IFC2X3 - University of Ghent
Author and License	Davy Van Deursen, Pieter Pauwels (mapping configuration from IFC2x3 Express specification from buildingSMART), unknown license
URL	http://multimedialab.elis.ugent.be/organon/ontologies/IFC2X3#
Description	OWL representation of the buildingSMART data model. The IFC data model is written in an EXPRESS schema (IFC2x3). This ontology is the result of an automated transformation of this EXPRESS schema into an OWL ontology. In this transformation process, every EXPRESS element that has a direct equivalent in OWL is mapped onto this equivalent. More specifically, for each ENTITY element in EXPRESS a corresponding OWL class is generated, EXPRESS attributes are converted into the appropriate OWL properties, etc.
Scope (Domain)	Buildings, AEC industry, BIM
Use cases (Motivation,	The IFC data model supports data sharing of BIM data. It supports coordination of design activities and hand-over of design and maintenance data. There are many use cases for smart cities where building data is of relevance, either to be referenced (in particular for

Relevance)	EMV use cases) or actively used by building simulation and maintenance (EMS use cases).
Data sets	IFC datasets can be generated by all major CAD tools. However, these tools export IFC data in the original SPF format only and thus has to be mapped to an RDF representation according to this ontology. Public IFC files are available from pilot and research projects mainly. However, building data is typically not published as it is mainly shared within the design team only or are handed over to contractual partners.
Open issues/ Challenges	There is an agreed standard developed by the buildingSMART organisation. This is the baseline definition from which an ifcOWL representation can be derived and enriched. So far, there are several mapping approaches, all of them dealing with slightly different requirements and boundaries. All mapping approaches will lose some sort of information as OWL is not able to deal with all constraints specified in the original IFC EXPRESS definition. Also, none of available mapping approaches is enriching the original definition.
Tool support	

NIST OntoSTEP Converter plugin for Protégé

Name	IFC2x3 NIST OntoSTEP Converter
Author and License	Rachuri Sudarsan, Raphael Barbau, Sylvere Krima; (developer of this tool) (mapping configuration from IFC2x3 Express specification from buildingSMART -> OWL-DL representation), unknown license
URL	http://www.nist.gov/OntoSTEP/ifc2x3 (download of the tool)
Description	See IFC Ontology
Scope (Domain)	Buildings, AEC industry, BIM
Use cases (Motivation, Relevance)	See IFC Ontology
Data sets	See IFC Ontology
Open issues/ Challenges	See IFC Ontology
Tool support	Plugin for Protégé that enables to convert EXPRESS schemata and SPF datasets.

The W3C Time Ontology

Name	The W3C Time Ontology
Author and License	Jerry R. Hobbs, Feng Pan W3C license
URL	http://www.w3.org/2006/time
Description	This ontology of temporal concepts provides a vocabulary for expressing facts about topological relations among instants and intervals, together with information about durations and about date time information.
Scope (Domain)	Metrics and indicators, Methods of measurement (scales, units, classifications), Time
Use cases (Motivation, Relevance)	The specification of temporal information is necessarily required for bringing the Semantic Web into reality. In ubiquitous and pervasive computing, a time ontology is crucial for modelling and reasoning about the time dimension of the context. When it comes to measuring energy consumption, the temporal aspect is clearly of relevance (e.g. When/How often is energy usage measured? – date, time, interval).
Data sets	
Open issues/ Challenges	The OWL Time ontology is in the state of a "first public working draft" (FPWD), which has been created by the Semantic Web Best Practices and Deployment Working Group (SWBPD). The SWBPD has finished in 2006 and so work on the Time ontology has been discontinued.
Tool support	

BFO (Basic Formal Ontology)

Name	BFO (Basic Formal Ontology)
Author and License	Pierre Grenon. License: CC-BY Creative Commons Attribution Unported (Open) http://creativecommons.org/licenses/by/3.0/
URL	http://www.ifomis.org/bfo/1.1
Description	BFO is an upper level ontology that is designed for use in supporting information retrieval, analysis and integration in scientific and other domains. However, it does not contain physical, chemical, biological or other terms which would properly fall within the coverage domains of the special sciences.

Scope (Domain)	Top level ontology
Use cases (Motivation, Relevance)	Upper level ontologies could be used for data integration across datasets
Data sets	Upper level ontologies could be used in a high number of datasets as they represent top concepts
Open issues/ Challenges	Unknown
Tool support	Unknown

Weather and Exterior Influence Information

Name	Weather and Exterior Influence Information
Author and License	Automation Systems Group, Institute of Computer Aided Automation, Vienna University of Technology unknown license
URL	https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/WeatherOntology.owl
Description	This smart home ontology for weather phenomena and exterior conditions was issued in 2011 as part of the ThinkHome project, which aimed to create an adaptive regulation for maximising energy efficiency in buildings. Shortly HOMEWEATHER, the ontology imports and extends W3C's Time ontology.
Scope (Domain)	Weather and climatic data, environmental data (e.g. pollution), Time, Modelling parameters, Controlling
Use cases (Motivation, Relevance)	The ontology covers a wide range of weather and climate data, such as atmospheric pressure, humidity, precipitation, temperature, wind, etc. In a smart home context, these data can be used to infer the proper action and perform tasks most energy-efficiently.
Data sets	
Open issues/ Challenges	
Tool support	

Units of Measure (OM)

Name	Units of Measure (OM)
Author and License	Hajo Rijgersberg, Mark van Assem, Don Willems, Mari Wigham, Jeen Broekstra, Jan Top CC-BY 3.0 license
URL	http://www.wurvoc.org/vocabularies/om-1.8/
Description	The Ontology of units of Measure and related concepts (OM) models concepts and relations important to scientific research. It has a strong focus on units and quantities, measurements, and dimensions.
Scope (Domain)	Measurement, Time, Metrics and indicators
Use cases (Motivation, Relevance)	Some classes relevant to the energy domain include electricity and magnetism (e.g. electric charge, electric conductivity, current, etc.) and space and time (e.g. area, height, length, period, time, etc.).
Data sets	
Open issues/ Challenges	
Tool support	

Measurement Ontology

Name	Measurement Ontology
Author and License	Ian Jacobi, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology unknown license
URL	http://www.telegraphis.net/ontology/measurement/measurement#
Description	The Measurement Ontology is an ontology in which measurements may be rendered. A measurement is a statistic that measures a quantity that may or may not have units. Relevant classes include measurement, quantity, unit, etc.
Scope (Domain)	Measurement, Methods of measurement (e.g. scales, units, classifications)
Use cases (Motivation,	SmartHome Weather references it

Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

Users and Preference Information

Name	Users and Preference Information
Author and License	Institute of Computer Aided Automation, Vienna University, Austria unknown license
URL	https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/ActorOntology.owl
Description	An ontology describing user information and preferences for Smart Home Systems. User profiling knowledge includes information about human characteristics (e.g age and gender) and preferences (e.g. visual and thermal habits) allowing the formulation of different habit patterns. This ontology came as a result of ThinkHome project, which utilizes artificial intelligence to improve control of home automation functions provided by dedicated automation systems. (for further information see https://www.auto.tuwien.ac.at/projectsites/thinkhome/user-information.html)
Scope (Domain)	User Preferences, User Profiling, User Scheduling, Energy Management
Use cases (Motivation, Relevance)	There are many use cases for smart cities where smart home occupancy data is of relevance. In particular, these data offer valuable information about : • thermal and visual preferences • configured schedules for energy profiling Advanced control automations related to this data can significantly improve energy-efficiency and energy-saving, yet preserving used comfort and preferences.
Data sets	As reported in ThinkHome project, all data collected will be publicly available through a dedicated web-site. There is no other evidence that this ontology has already been used by other projects/applications, in order to seek for more available data-sets.
Open issues/ Challenges	

Tool support	

Energy and Resource Information

Name	Energy and Resource Information
Author and License	TU Vienna
URL	https://www.auto.tuwien.ac.at/projectsites/thinkhome/facilities-and-energy-information.html https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/energy/changelog/EnergyResource_Revision_1.03.txt
Description	An ontology representing energy information for Smart Home Systems.
Scope (Domain)	Home Automation
Use cases (Motivation, Relevance)	
Data sets	https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/EnergyResourceOntology.owl https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/EnergyResourceOntologyExample.owl
Open issues/ Challenges	
Tool support	

Measurement Units Ontology

Name	MUO Measurement Units Ontology
Author and License	Luis Polo, Diego Berrueta, Fundación CTIC License not specified
URL	http://mymobileweb.morfeo-project.org/specs/name (Not available)
Description	Ontology representing measurements units, in terms of base, complex, derived units.
Scope (Domain)	All measured entities

Use cases (Motivation, Relevance)	It is relevant due to the necessity to compare same type entities specified in different measure units, such as energy expressed in cal rather than J or Wh.
Data sets	
Open issues/ Challenges	
Tool support	

Casas Ontology for Smart Environments (COSE)

Name	Casas Ontology for Smart Environments (COSE)
Author and License	School of Electrical Engineering and Computer Science, Washington State University, Box 642752, Pullman, WA, 99164-275
URL	-
Description	The number of smart appliances and devices in the home and office has grown dramatically in recent years. Unfortunately, these devices rarely interact with each other or the environment. In order to move from environments filled with smart devices to smart environments, there must be a framework for devices to communicate with each other and with the environment. This enables reasoners and automated decision makers to understand the environment and the data collected from it. Semantic web technologies provide this framework in a well-documented and flexible package. In this paper we present the Casas Ontology for Smart Environments (COSE) and accompanying data from a test smart environment and discuss the current and future challenges associated with a Smart Environment on the Semantic Web.
Scope (Domain)	Smart Environments, Ambient Assisted Living
Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

DOLCE

Name	DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering)
Author and License	Claudio Masolo License unknown.
URL	http://www.loa-cnr.it/ontologies/DOLCE-Lite.owl#
Description	DOLCE is the first module of the WonderWeb Foundational Ontologies Library (WFOL). DOLCE has a clear cognitive bias, in the sense that it aims at capturing the ontological categories underlying natural language and human common-sense. its authors believe that such bias is very important for the Semantic Web. DOLCE is an ontology of particulars, in the sense that its domain of discourse is restricted to them. A basic choice we make in DOLCE is the so-called multiplicative approach: different entities can be co-located in the same space-time (e.g. the vase and the amount of clay).
Scope (Domain)	Top level ontology
Use cases (Motivation, Relevance)	Upper level ontologies could be used for data integration across datasets
Data sets	Upper level ontologies could be used in a high number of datasets as they represent top concepts
Open issues/ Challenges	
Tool support	

DUL Ontology

Name	DUL (DOLCE+DnS Ultralite)
Author and License	Aldo Gangemi. License unknown.
URL	http://www.ontologydesignpatterns.org/ont/dul/DUL.owl
Description	It is a simplification and an improvement of some parts of DOLCE Lite-Plus library (cf. http://dolce.semanticweb.org), and Descriptions and Situations ontology (cf. http://www.ontologydesignpatterns.org/wiki/Ontology:DnS)Its purpose is to provide a set of

	upper level concepts that can be the basis for easier interoperability among many middle and lower level ontologies.
Scope (Domain)	Top level ontology
Use cases (Motivation, Relevance)	Upper level ontologies could be used for data integration across datasets
Data sets	Upper level ontologies could be used in a high number of datasets as they represent top concepts
Open issues/ Challenges	Unknown
Tool support	Unknown

The Timeline Ontology

Name	The Timeline Ontology
Author and License	Yves Raimond, Samer Abdallah. Centre for Digital Music in Queen Mary, University of London. Licensed under a Creative Commons Attribution License.
URL	http://motools.sf.net/timeline/timeline.n3
Description	This ontology defines the TimeLine concept, representing a coherent backbone for addressing temporal information. Each temporal object (signal, video, performance, work, etc.) can be associated to such a timeline. Then, a number of Interval and Instant can be defined on this timeline.
Scope (Domain)	Time managing. It useful for anything related to time or time depending.
Use cases (Motivation, Relevance)	The principal applications interests are any non-static process that need to gather information using a precise and synchronous time reference.
Data sets	
Open issues/ Challenges	The primary scope of this ontology (music and videos) could make the Timeline Ontology and its related tools more difficult to use in the Smart Cities contest.
Tool support	A tool created to manipulate data in this ontology: http://sourceforge.net/projects/motools/

SESAME-S Smart Building Ontology

Name	SESAME-S Smart Building Ontology
Author and License	Research Centre for Telecommunication (FTW, http://www.ftw.at/), Austria unknown license
URL	http://datahub.io/dataset/smartbuilding-sesames https://commondatastorage.googleapis.com/ckannet-storage/2012-08-20T165445/SmartBuildingv3.owl
Description	SESAME-S = Semantic Smart Metering Services for Energy Efficient Houses This ontology is a typical example of a purpose-built ontology. It was developed within the SESAME project, which is already finished. The ontology is not maintained anymore and no further documentation is available. It contains about 20 class and 30 property definitions, thus being a rather small ontology in terms of size and scope. It is focused on the data that has been managed in the two real-world examples, e.g. measurements of temperature, humidity, light and presence of persons.
Scope (Domain)	Smart Sensors, Devices
Use cases (Motivation, Relevance)	The ontology was developed to show the "next generation of energy efficient buildings". It is part of a prototype development to proof the concept of semantic smart metering and providing services for energy efficiency. One of the goals was to raise awareness for taking care of reducing energy consumption within a building by providing measured data to people who are using the facilities. More information about the project is available on their website (http://sesame-s.ftw.at/) and in a number of research publiciations, e.g. in: Girtelschmid, S., Steinbauer, M., Kumar, V., Fensel, A., Kotsis, G. "On the Application of Big Data in Future Large Scale Intelligent Smart City Installations", International Journal of Pervasive Computing and Communications, Emerald Group Publishing, Vol. 10 Iss: 2 (2014).
Data sets	Data from two real-world examples have been managed with this ontology. One example from Austria, the Kirchdorf school example, and another one from Russia, the Chernogolovka factory example. The datasets are not public available as there are strong concern regarding security and privacy issues (actual energy consumption, usage patterns of the building).
Open issues/ Challenges	There is no maintenance and no further documentation of this ontology. It is used by the authors as a baseline for follow-up projects.
Tool support	Prototypes/tools developed in SESAME-S project.

Simulation Information Model (SIM) Ontology

Name	Simulation Information Model (SIM) Ontology
Author and License	unknown license
URL	http://www.modelservers.org/public/ontologies
Description	Developed and used in the IntUBE project (Intelligent Use of Building's Energy Information), which was carried out from 2007 to 2010.
Scope (Domain)	Building usage, Building performance
Use cases (Motivation, Relevance)	Simulated data generated by energy simulation tools (including their input parameters)
Data sets	Examples from the IntUBE project available (see URL).
Open issues/ Challenges	The status of the ontology is unclear. Website (and domain) is not available
Tool support	Data managed in the "Energy-information integration platform"

Performance Information Model (PIM) Ontology

Name	Performance Information Model (PIM) Ontology
Author and License	unknown license
URL	http://www.modelservers.org/public/ontologies
Description	IntUBE project (Intelligent Use of Building's Energy Information) – finished 2011 (project website no more available)
Scope (Domain)	Building usage, Building performance
Use cases (Motivation, Relevance)	dynamic data obtained from monitoring systems, including climate, building use and energy performance

Data sets	
Open issues/ Challenges	see Simulation Information Model (SIM) Ontology
Tool support	

The W3C Sensor Network Ontology

Name	The W3C Sensor Network Ontology
Author and License	W3C Semantic Sensor Network Incubator Group W3C Software Notice and License
URL	http://purl.oclc.org/NET/ssnx/ssn
Description	This ontology describes sensors and observations, and related concepts. It does not describe domain concepts, time, locations, etc. these are intended to be included from other ontologies via OWL imports. (For further information see : http://www.w3.org/2005/Incubator/ssn/wiki/Report_Work_on_the_SSN_ontology)
Scope (Domain)	Sensors, Sensors Measuring, Monitoring, Devices
Use cases (Motivation, Relevance)	Measuring and Monitoring support the basis of the intelligent operation. Valuable information about • Sensors as a device () • Measuring operations and measuring capability. • Device http://purl.oclc.org/NET/ssnx/ssn#Device Five working examples are already included in the reference wiki page, illustrating the application of different parts of this ontology, such as: University deployment, Smart product, Wind sensor, Agriculture Meteorology and Linked Sensor Data.
Data sets	The W3C Semantic Sensor Network Incubator Group maintains hosts a wiki reference page since 2005, providing the respective ontologies for public uses and allowing interaction with public via open data and communication methods via a W3C list (public-xg-ssn@w3.org) It is expected that Data-sets based on this ontology may already exist from other projects.
Open issues/ Challenges	
Tool support	

Building Information Model (BIM) Ontology

Name	Building Information Model (BIM) Ontology
Author and License	unknown license
URL	http://www.modelservers.org/public/ontologies
Description	IntUBE project (Intelligent Use of Building's Energy Information) – finished 2011 (project website no more available)
Scope (Domain)	Building
Use cases (Motivation, Relevance)	Static data about the building in general, such as building location, process stage, spaces, envelopes and building services
Data sets	
Open issues/ Challenges	see Simulation Information Model (SIM) Ontology
Tool support	

Global City Indicator Foundation Ontology

Name	Global City Indicator Foundation Ontology
Author and License	"Global City Indicators©" is a term created by the Global City Indicators Facility in 2010 at the University of Toronto. All rights apply. GCI refers to the indicators created by the GCIF to establish a global standard of over 100 city indicators with a standardized definition and methodology, tested with over 250 cities globally since 2010. The GCIs are now in a draft international standard currently being voted upon by member countries with a view to publishing the GCIs in 2013
URL	
Description	Cities are moving towards policy-making based on data. But as Hoornweg et al.[44] state: "Today there are thousands of different sets of city (or urban) indicators and hundreds of

[44] Hoornweg, D., Nunez, F., Freire, M., Palugyai, N., Herrera, E.W., and Villaveces,M., (2007), "City Indicators: Now to Nanjing", World Bank Policy Research Working Paper 4114.

	agencies compiling and reviewing them. Most cities already have some degree of performance measurement in place. However, these indicators are usually not standardized, consistent or comparable (over time or across cities), nor do they have sufficient endorsement to be used as ongoing benchmarks." In response to this challenge, the Global City Indicator (GCI) Facility was created by the World Bank to define a set of city indicators that can be consistently applied globally.
Scope (Domain)	city performance measurement
Use cases (Motivation, Relevance)	www.cityindicators.org
Data sets	
Open issues/ Challenges	
Tool support	

User Behavior and Building Process Information

Name	User Behavior and Building Process Information
Author and License	TU Vienna, unkown license
URL	https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/ProcessOntology.owl
Description	An ontology representing processes in Smart Home Systems.
Scope (Domain)	Occupancy, building domain
Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

Cadastre and Land Administration Thesaurus (CaLAThe)

Name	Cadastre and Land Administration Thesaurus (CaLAThe)
Author and License	Professor Erik Stubkjær, Department of Planning, Aalborg University, Denmark, Dr. Volkan Cagdas, Department of Surveying Engineering, Yildiz Technical University, Turkey. licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License
URL	http://www.cadastralvocabulary.org (available on request)
Description	This ontology provides a controlled vocabulary, which is derived mainly from the ISO/DIS 19152 Land Administration Domain Model and related to existing thesauri, primarily the GEMET thesaurus, the AGROVOC thesaurus, and the STW Thesaurus for Economics.
Scope (Domain)	Buildings, cadastre, geography
Use cases (Motivation, Relevance)	In smart cities application, it could be useful where certain buildings data are needed; e.g. geographical positioning, internal divisions (apartments), spatial representation.
Data sets	
Open issues/ Challenges	CaLAThe is encoded as a Simple Knowledge Organization System (SKOS), according to specifications developed by the World Wide Web Consortium (W3C).
Tool support	

CASCADE airport ontology

Name	CASCADE airport ontology
Author and License	Institute Mihajlo Pupin, Sanja Vranes, Nikola Tomasevic, Marko Batic, CASCADE ICT for Energy Efficient Airports Unknown license
URL	https://webgate.ec.europa.eu/fpfis/wikis/display/eeSemantics/CASCADE+Modelling+Ontology https://webgate.ec.europa.eu/fpfis/wikis/download/attachments/44483343/CASCADE%20Core%20Airport%20Ontology%20%28class%29.owl?version=1&modificationDate=1399554858401&api=v2
Description	The CASCADE Core airport ontology provides a generic model of the airport facility as a set of concepts and corresponding relationships among them. The purpose of the Core airport ontology is to provide the modelling guidelines and to describe the technical

	characteristics/relations of related systems installed at the site, their topological profile, as well as to facilitate the interpretation of signals.
Scope (Domain)	Airports, automated buildings
Use cases (Motivation, Relevance)	Even if this ontology is oriented to create a model of airport facility, it can be used also in generic buildings modelling, particularly public buildings or complexes, due to the commonality with airport sub-functions.
Data sets	
Open issues/ Challenges	The CASCADE deontology is characterized by a partial superposition with other ontologies taken into account (regarding geography or buildings). It would be expectable to reach a
Tool support	

Nikola Tesla Airport (NTA) Ontology

Name	Nikola Tesla Airport (NTA) Ontology
Author and License	Possibly: University of Belgrade, Institute Mihajlo Pupin
URL	
Description	The ontology facilitates the interpretation and semantic enrichment of SCADA signals using the underlying spatial and topological model of the airport infrastructure as well as vendor data regarding the equipment characteristics, protocols and standards used. http://www.e-drustvo.org/icist/2012/html/pdf/495.pdf
Scope (Domain)	airport managament, emergency management, facility management
Use cases (Motivation, Relevance)	"Nikola Tesla" airport Belgrade "For improving and providing more intelligent , holistic, airport facility management systems that rely on contemporary management platforms such as Supervisory Control and Data Acquisition (SCADA) systems, classification and description of various information/data within the airport infrastructure"[45]
Data sets	

[45] http://www.e-drustvo.org/icist/2012/html/pdf/495.pdf

Open issues/ Challenges	
Tool support	

Trade

Name	Trade
Author and License	Antonio Paredes-Moreno. No license information.
URI	http://personal.us.es/aparedes/Trade.owl
Description	This ontology defines the classes, properties and individuals that make up the commercial management specially focused to purchase orders, in a company dedicated primarily to trade in electrical, energy and environmental products.
Scope (Domain)	Energy trade
Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

Geonames Ontology

Name	GeoNames Ontology
Author and License	Bernard Vatant, GeoNames. Creative Commons CC BY 3.0
URL	http://www.geonames.org/ontology/ontology_v3.1.rdf

Description	The GeoNames Ontology makes it possible to add geospatial semantic information to the World Wide Web. All over 8.3 million geonames toponyms now have a unique URL with a corresponding RDF web service. Other services describe the relation between toponyms.
Scope (Domain)	Geography
Use cases (Motivation, Relevance)	Relevant to guarantee unique reference to toponyms and easy information access through the GeoNames database (http://sws.geonames.org), especially geographic position.
Data sets	At http://www.geonames.org/advanced-search.html all of the rdf produced by GeoNames are available.
Open issues/ Challenges	
Tool support	

Data Cube

Name	Vocabulary for multi-dimensional (e.g. statistical) data publishing
Author and License	Contributors: Arofan Gregory, Dave Reynolds, Ian Dickinson, Jeni Tennison, Richard Cyganiak W3C license
URL	http://www.w3.org/TR/vocab-data-cube/
Description	This vocabulary allows multi-dimensional data, such as statistics, to be published in RDF. It is based on the core information model from SDMX (Statistical Data and Metadata Exchange).
Scope (Domain)	Statistics
Use cases (Motivation, Relevance)	This vocabulary was originally developed and published outside of W3C, but has been extended and further developed within the Government Linked Data Working Group. It is aimed at people wishing to publish statistical or other multi-dimension data in RDF. The cube model is very general and so the Data Cube vocabulary can be used for various data sets such as survey data, spreadsheets and OLAP data cubes. Energy-related datasets can therefore also be used.

Data sets	Datasets are at the core of the vocabulary structure. The vocabulary defines them any collection of statistical data that corresponds to a defined structure. Different views of the data can be achieved through slicing.
Open issues/ Challenges	
Tool support	

The PROV Ontology

Name	PROV-O: The PROV Ontology
Author and License	Timothy Lebo, Satya Sahoo, Deborah McGuinness. Copyright © 2013 W3C® (MIT, ERCIM, Keio, Beihang), All Rights Reserved
URL	http://www.w3.org/ns/prov-o
Description	The PROV Ontology (PROV-O) expresses the PROV Data Model [PROV-DM] using the OWL2 Web Ontology Language (OWL2) [OWL2-OVERVIEW]. It provides a set of classes, properties, and restrictions that can be used to represent and interchange provenance information generated in different systems and under different contexts. It can also be specialized to create new classes and properties to model provenance information for different applications and domains.
Scope (Domain)	General, provenance
Use cases (Motivation, Relevance)	In smart cities case, it could be useful to classify pieces of information in terms of trust and reliability, due to the high level of integration of information by different sources
Data sets	
Open issues/ Challenges	
Tool support	

DogOnt

Name	DOGONT - Ontology Modeling for Intelligent Domotic Environments
Author and License	Dario Bonino
URL	http://www.cad.polito.it/pap/exact/iswc08.html
Description	The DogOnt ontology supports device/network independent description of houses, including both controllable and architectural elements
Scope (Domain)	Architecture
Use cases (Motivation, Relevance)	
Data sets	http://elite.polito.it/ontologies/dogont.owl
Open issues/ Challenges	
Tool support	

SUMO (Suggested Upper Merged Ontology)

Name	SUMO (Suggested Upper Merged Ontology)
Author and License	Adam Pease. License unknown.
URL	http://www.ontologyportal.org/
Description	The Standard Upper Ontology is the result of a joint effort to create a large, general-purpose, formal ontology. It is promoted by the IEEE Standard Upper Ontology working group, and its development began in May 2000. The participants were representatives of government, academia, and industry from several countries. The effort was officially approved as an IEEE standard project in December 2000.
Scope (Domain)	Top level ontology
Use cases (Motivation,	Upper level ontologies could be used for data integration across datasets

Relevance)	
Data sets	Upper level ontologies could be used in a high number of datasets as they represent top concepts
Open issues/ Challenges	Unknown
Tool support	Unknown

BOnSAI

Name	Bonsai - Smart Building Ontology for Ambient Intelligence
Author and License	Thanos G. Stavropoulos Dimitris Vrakas Danai Vlachava Nick Bassiliades No license information.
URI	http://lpis.csd.auth.gr/ontologies/bonsai/BOnSAI.owl
Description	The ontology extends and benefits from existing ontologies in the field, but also adds classes needed to sufficiently model every aspect of a service-oriented smart building system. Namely, it includes concepts modeling all functionality (i.e. services, operations, inputs, outputs, logic, parameters and environmental conditions), QoS (resources, QoS parameters), hardware (smart devices, sensors and actuators, appliances, servers) users and context (user profiles, moods, location, rooms etc.). (Literally taken from https://www.researchgate.net/publication/254006761_BOnSAI_a_smart_building_ontology_for_ambient_intelligence)
Scope (Domain)	Smart buildings
Use cases (Motivation, Relevance)	The ontology is designed for the Smart IHU ambient setting whose goal is to provide automation and energy savings at the International Hellenic University (IHU) premises. This environment is equipped with sensors and actuators (so-called smart devices) in large scale, which interact with the rest of the system using the web service interface (Literally taken from https://www.researchgate.net/publication/254006761_BOnSAI_a_smart_building_ontology_for_ambient_intelligence).
Data sets	

Open issues/ Challenges	
Tool support	Smart IHU Smart Building environment

OGC GeoSPARQL

Name	OGC GeoSPARQL
Author and License	Open Geospatial Consortium No license information.
URI	http://www.opengis.net/ont/geosparql
Description	An RDF/OWL vocabulary for representing spatial information. This vocabulary is based on the effort of OGC to provide 'standard' terms in RDF for describing geographic data on the Web.
Scope (Domain)	Spatial information, Geographic information
Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

WGS84 Geo Positioning

Name	WGS84 Geo Positioning
Author and License	Dan Brickley, Tim Berners-Lee, Unknown
URI	http://www.w3.org/2003/01/geo/wgs84_pos
Description	A vocabulary for representing latitude, longitude and altitude information in the WGS84 geodetic reference datum.

Scope (Domain)	Geographic information
Use cases (Motivation, Relevance)	A *basic* RDF vocabulary that provides the Semantic Web community with a namespace for representing **lat**(itude), **long**(itude) and other information about spatially-located things, using WGS84 as a reference datum.
Data sets	
Open issues/ Challenges	
Tool support	

Open Street Map (OSM) ontology

Name	Open Street Map (OSM) ontology
Author and License	Unknown
URI	http://mapserv.kt.agh.edu.pl/ontologies/osm.owl
Description	The ontology defines classes of objects appearing on maps: roads, railways, water ways, amenities, emergency infrastructure, public transport, shops, tourist attractions, etc. This large ontology contains about 660 classes, which were identified based on the published set of OSM tags and their values.
Scope (Domain)	Physical features on the ground, Maps
Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	Non accessible web page.
Tool support	

Places Ontology

Name	Place ontology
Author and License	Michael Smethurst Rob Styles Tom Scott Licence: CC0 Universal (http://creativecommons.org/publicdomain/zero/1.0/)
URI	http://purl.org/ontology/places
Description	The Places Ontology is a simple lightweight ontology for describing places of geographic interest.
Scope (Domain)	Places of geographic interest.
Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

eDIANA context awareness ontology

Name	eDIANA context awareness ontology
Author and License	Unknown
URI	http://www.owl-ontologies.com/ContextAwareness_eDIANA.owl (N.B: wrong URI in OWL file! URL:: https://sites.google.com/site/smartappliancesproject/ontologies/ediana.owl)
Description	The main objective of this ontology is to define the universe of concepts and their relations in the domain of eDIANA Platform Architecture, related to device awareness. The eDIANA Platform Architecture provides a wide and heterogeneous list of devices in hierarchical levels: MacroCell and Cell.
Scope (Domain)	Devices

Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

Urban Energy Ontology

Name	Urban Energy Ontology
Author and License	Apache License Version 2.0 (www.apache.org/licenses/)
URI	http://www.semanco-tools.eu/urban-enery-ontology
Description	This ontology describes the domain of urban planning based on the OWL-based translation of the Suggested Upper Merged Ontology (SUMO), available at: http://www.ontologyportal.org/ .
Scope (Domain)	Urban Planning
Use cases (Motivation, Relevance)	The SEMANCO Energy Model is a formal ontology – specified using Web Ontology Language 2 (OWL 2) – comprising concepts captured from diverse sources including standards, use cases and activity descriptions and data sources related to the domains of urban planning and energy management. In particular it contains the terms and attributes that describe regions, cities, neighbourhoods and buildings; energy consumption and CO2 emission indicators, as well as climate and socio- economic factors that influence energy consumption.
Data sets	
Open issues/ Challenges	
Tool support	

Concept Modelling Ontology (CMO)

Name	Concept Modelling Ontology (CMO)
Author and License	Michel Böhms, Peter Bonsma, Bruno Fies Unknown license
URI	http://www.modelservers.org/public/ontologies/cmo/cmo.ttl
Description	CMO is a reusable, generic ontology (also referred to as an 'upper ontology') that enables full-power, pure semantic, concept modelling
Scope (Domain)	Generic Ontology, Top level ontology
Use cases (Motivation, Relevance)	The modelling & monitoring of energy nodes in urban areas for holistic and optimized energy management within the Odyseus project. http://www.odysseus-project.eu/ The modelling & configuration of residential districts/homes for supporting Self-organized Collective Housing (CSO) in the FP7-NMP Proficient project http://www.proficient-project.eu/
Data sets	
Open issues/ Challenges	
Tool support	Ifc2cmo http://www.resilient-project.eu/documents/35984/54543/2_ODYSSEUS.pdf

Registered Organization Vocabulary

Name	Registered Organization Vocabulary
Author and License	Unknown
URI	http://www.w3.org/ns/regorg
Description	This is a vocabulary for describing organizations that have gained legal entity status through a formal registration process, typically in a national or regional register. It focuses solely on such organizations and excludes natural persons, virtual organizations and other types of legal entity or 'agent' that are able to act. It is a profile of the more flexible and comprehensive Organization Ontology [ORG].
Scope (Domain)	Organization

Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

The Event Ontology

Name	The Event Ontology
Author and License	Yves Raimond (yves@dbtune.org), Samer Abdallah (samer.abdallah@elec.qmul.ac.uk), CC-BY Creative Commons Attribution Unported (Open) http://creativecommons.org/licenses/by/3.0/
URI	http://purl.org/NET/c4dm/event.owl
Description	The event ontology deals with the notion of reified events. It defines one main Event concept. An event may have a location, a time, active agents, factors and products, as depicted below.
Scope (Domain)	Event, time
Use cases (Motivation, Relevance)	
Data sets	Use of the ontology in other domains: • glastonbury-2011 • linked-open-data-of-ecology • rdfize-lastfm • rkb-explorer-webscience
Open issues/ Challenges	
Tool support	

km4city

Name	km4city
Author and License	paolo nesi (paolo.nesi@unifi.it) http://www.disit.dinfo.unifi.it, CC-BY-SA Creative Commons Attribution-ShareAlike Unported (Open) http://creativecommons.org/licenses/by-sa/3.0/
URI	http://www.disit.org/km4city/schema
Description	To interconnect the data provided by the Tuscany Region, the Open Data of the City of Florence, and the other Static and Real Time dataset, we started to develop a Knowledge Model, that allows to collect all the data coming from the city, related to mobility, statistics, street graph, sensors, cultural heritage, parkings, weather, services, energy, events.
Scope (Domain)	geographic locations, transportation, city, sensors, cultural heritage, services, parkings, weather, events, public structures
Use cases (Motivation, Relevance)	No use case defined, but demo mapping applications.
Data sets	This project published the transportation data for the city of Florence and geographic data for the Tuscani region (Italy) : http://log.disit.org/
Open issues/ Challenges	The information is difficult to find in a web site consisting of one page with criptic URIs.
Tool support	Tools and slides: http://www.disit.org/6056 documentation ENG: http://www.disit.org/5606 related to version 1.1 of the ontology documentation ITA: http://www.disit.org/6461 of version 1.4 of the ontology image: http://www.disit.org/6507 of version 1.4 of the ontology ontology .. the OWL and triple version http://www.disit.org/6506 mobile demonstrator. http://LOG.disit.org graph can be used to browse the knowledge model of Smart City, just an example of a Florence segment. http://log.disit.org/servic /?graph=71de8caef449ed56143aa95c8c8266ab From that, you can see the whole DISIT knowledge knowledge model for Florence, based on Km4City ontology. Link at Service Map tool: http://servicemap.disit.org API of Servicemap http://www.disit.org/6597 open source mobile tool: http://www.disit.org/659 Service Map tool: http://servicemap.disit.org a tool for developers to pose geographic queries (learn and generate code queries in an esy manner) and see the knowledge base produced by the harvesting process

Internet of Things (IoT) Ontology

Name	Internet of Things (IoT) Ontology
Author and License	Konstantinos Kotis, Unknown

URI	http://purl.org/IoT/iot
Description	Internet of Things (IoT) Ontology is a reference ontology for data integration and semantic coordination of smart entities. The aim of the ontology is to provide a clear understanding of the new research domain of IoT in respect to the need for 'true' (i.e. semantic) interoperability of smart entities and other kind of entities (control, physical) that may be plugged in it anytime, by anyone and from anyplace. The objective is not to focus in sensor and observation data descriptions as in SSN ontology, but instead to emphasize the notion of interconnected, clustered and aligned smart entities towards supporting their semantic registration, coordination and retrieval in a Web of Things.
Scope (Domain)	Internet of Things, Web of Things
Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

OpenIoT Ontology

Name	OpenIoT Ontology
Author and License	http://myr.altervista.org/foaf.rdf#me, W3C software license (Open) http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231
URI	http://openiot.eu/ontology/ns/
Description	This ontology describes abstraction of sensors and their integration with cloud computing concepts. This ontology is developed by DERI (http://www.deri.ie) for the OpenIoT project (http://openiot.eu). It is based on the alignment among the W3C Semantic Sensor Networks Incubator Group (SSN-XG) ontology, the SPITFIRE ontology and the LSM vocabulary.
Scope (Domain)	Sensors, Cloud Computing
Use cases (Motivation, Relevance)	

Data sets	
Open issues/ Challenges	
Tool support	

SPITFIRE Ontology

Name	SPITFIRE Ontology
Author and License	http://myr.altervista.org/foaf.rdf#me, Alexandre Passant, W3C software license (Open) http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231
URI	http://spitfire-project.eu/ontology/ns/
Description	This ontology describes sensors, observations, and related concepts. It also describes events and their correlations. The final aim is to support a better description of sensor context. This ontology is developed by DERI (http://www.deri.ie) for the SPITFIRE project (http://spitfire-project.eu). It is based on the alignment among the W3C Semantic Sensor Networks Incubator Group (SSN-XG) ontology, the Dolce-DnS Ultralite ontology and the Event Model F ontology.
Scope (Domain)	Sensors
Use cases (Motivation, Relevance)	https://www.itm.uni-luebeck.de/files/1213/6973/3906/IEEEComMag.pdf
Data sets	
Open issues/ Challenges	"As it was difficult to foresee the wealth of current Web applications back when the Web was created, we have to wait and see how people will use the Semantic Web of Things. It is also hard to predict if a Semantic Web of Things will be as broadly adopted as the Web is today. One indicator is that LOD has already achieved significant uptake by governments (including UK, USA), the media sector (BBC), life sciences, geo information systems, and Web companies (Freebase). Making sensor data part of this data pool is clearly beneficial as then integration with knowledge from arbitrary sources is possible. For example, sensors and their data can be linked to geographic data (correlated natural phenomena), user-generated data (social feedback), government data (census information), life-science data (causes and effects of diseases), etc." Source: https://www.itm.uni-luebeck.de/files/1213/6973/3906/IEEEComMag.pdf, 2015

Tool support	

Eurobau Utility Ontology

Name	Eurobau Utility Ontology
Author and License	CC-BY Creative Commons Attribution Unported (Open) http://creativecommons.org/licenses/by/3.0/
URI	http://semantic.eurobau.com/eurobau-utility.owl
Description	The Eurobau Utility Ontology provides utility elements for describing building materials and respective offerings from the Eurobau semantic dataspace. This ontology defines a few extensions to GoodRelations.
Scope (Domain)	Building Materials
Use cases (Motivation, Relevance)	BauDataWeb is one of the largest and richest public datasets for a well-defined vertical sector that is available on the Semantic Web. It covers a major share of the European It covers a major share of the European market.
Data sets	81 Manufacturers / Brands 19 Reseller 183 Warehouse locations 56.360 Product Models (including variants) 56.360 Product Models (including variants) 1.783.798 Offerings 95 % of the product models include rich FreeClassOWL descriptions
Open issues/ Challenges	Unkown
Tool support	Any SPARQL endpoint SPARQL queries via the OpenLink Software Virtuoso repositories at http://lod.openlinksw.com/sparql and http://linkeddata.uriburner.com/sparql

FreeClassOWL Ontology

Name	FreeClassOWL Ontology
Author and License	CC-BY Creative Commons Attribution Unported (Open) http://creativecommons.org/licenses/by/3.0/
URI	http://www.freeclass.eu/freeclass_v1.owl
Description	The FreeClass Ontology for construction and building materials and services provides classes and properties for describing products and services from the building and construction industry. It is derived from the free classification standard freeClass. For more information, see http://www.freeclass.eu/. The conversion of this ontology has been funded by the Österreichische Forschungsförderungsgesellschaft GmbH (FFG) and the Bundesministerium für Verkehr, Innovation und Technologie (BMVIT) under the FIT-IT Semantic Systems project 'myOntology' (contract number 812515). For describing the commercial aspects of respective offerings, please use the GoodRelations ontology. The FreeClassOWL Ontology is a GoodRelations-compliant ontology for describing construction and building materials and services.
Scope (Domain)	Building Materials
Use cases (Motivation, Relevance)	
Data sets	N/A
Open issues/ Challenges	
Tool support	FreeClass Semantic Search for Construction Materials: Online tool for demonstrating how the usage of Semantic Web technologies can improve a search for building and construction materials

CERISE CIM Profile for Smart Grids

Name	CERISE CIM Profile for Smart Grids
Author and License	TNO: Maarten Steen Unknown license
URI	http://ns.cerise-project.nl/energy/def/cim-smartgrid
Description	A Profile of the IEC Common Information Model (CIM) for Smart Grids, developed by the Cerise-SG project

Scope (Domain)	Smart Grids
Use cases (Motivation, Relevance)	Interoperability with a special interest in the information exchanges between smart grids and their surroundings. Creation of future proof and efficient information exchange between the energy sector, eGovernment and geo-world. It is not realistic to assume that these worlds can be easily adjusted given the mass behind it. Our approach covers two levels: technical (web services, exchange formats, protocols) and content (semantics, information models). In case of model mismatches between the different worlds, semantic model transformation services are developed. More specifically the following use cases have been analysed: - Information is exchanged within a crisis management scenario dealing with the effects of a flood on the power grid. Due to sector-interdependent effects during this disaster data sharing is essential for successful crisis management; - Energy Balancing Information Facility for facilitating the administrative balancing in a smart grid
Data sets	
Open issues/ Challenges	Laura Daniele [17-07-2015]: This is a draft version for internal use in the CERISE project. We encountered some issues with the generation this OWL profile with the CIMTool that still need to be solved. One issue is that the mapping of cardinalities in the transformation from UML to OWL is not always correct. For example: • the UML association Meter [0..1] was mapped by the CIMTool into the OWL property MeterReading.Meter exactly 1, while we would expect it to be mapped to MeterReading.Meter max 1 • the UML association Readings [0..*] was mapped by the CIMTool into the OWL property MeterReading.Readings min 1, while while we would expect it to be mapped to MeterReading.Readings min 0 To overcome the issue we are changing manually the incorrect cardinalities in the generated OWL profile, but there are many properties and this requires quite some time and effort, so some cardinalities can still be not compliant with the original UML model. Roel Stap[12-06-2015]: For gas metering the class SimpleEndDeviceFunction is defined, specialisation of EndDeviceFunction. Within this class there is a mandatory attribute defined \"kind\" of type EndDeviceFunctionKind. This last class is an enumeration of different type of metering. This class can be used can be used to distinguish between different tupe of metering, for example electric and gas metering. This means this class is mandatory, for each type of metering the type shall be defined.
Tool support	Created with TopBraid Composer

COINS Building Information Model (CBIM)

Name	COINS Building Information Model (CBIM)
Author and License	The COINS system is a publication of the COINS programme, represented by CUR Bouw & Infra, Gouda.

	The COINS system is an open standard. The contents of the standard are freely available. Reuse of the standard is not subject to any restrictions.
URI	http://www.coinsweb.nl/c-bim.owl
Description	COINS is an open BIM standard. It is complementary to standards issued by buildingSMART such as IFC, IFD Library and IDM. COINS supports the exchange of Systems Engineering information and ensures that an object tree, GIS data, 2D drawings, 3D models, IFC models and object type library can be stored in association in a database. It also provides a BIM-container interchange format. It is used by partners in building construction projects for the purpose of exchanging building information and managing building information. The first edition was published in 2010 as COINS 1.0. A first update was released as COINS 1.1 in December 2014.
Scope (Domain)	Buildings / Exchange of building information and management of building information
Use cases (Motivation, Relevance)	COINS is not describing use case but what they called “Reference frameworks”. A Reference frameworks intended as industry standards will be made available dealing with the specific issues mentioned below: • Functional specification (available) • Preparing a Design Dossier • Transferring building information • Object data management • Preparing the object structure • Testing a functional spatial schedule of requirements • Making quantity estimates (available) • Applying a library • Using construction sector libraries • Managing a building configuration
Data sets	N/A
Open issues/ Challenges	For the moment, only 2 reference frameworks are under development.
Tool support	The COINS Navigator is a reference implementation to demonstrate the principles that lie at the bottom of the COINS standardization development. The application has the following features: • creating a C-BIM model • editing all aspects of a C-BIM model • loading/saving a C-BIM model • importing/exporting a COINS Container • simulate a COINS Building Information System (CBIS) • demonstrate the COINS version management system • merging C-BIM models

	• report generation in Excel or HTML format • switch between layer view and object tree view • build and link to COINS object libraries • link to external object libraries (CROW Cheobs, BuildingSMART IFD Library, ETIM) • specify and checking a Window of Authorization • link and visualize IFC models and/or PMO models • import planning data from Primavera of MSProject • link with the VISI building management data standard (under development) The COINS Navigator can freely be downloaded, used and further distributed.

CASCADE Fiumicino Airport ontology

Name	CASCADE Fiumicino Airport ontology
Author and License	Institute Mihajlo Pupin: Sanja Vranes, Nikola Tomasevic, Marko Batic Unknown license
URI	http://jpo.imp.bg.ac.rs/cascade/airport-ontology/FCO/airportOntologyFCO_TBox.owl
Description	A full-blown ontology model of Fiumicino airport (Rome, Italy) which models a specific airport infrastructure by classifying installed technical systems relevant to the energy management aspect. It was developed by extension and population of the CASCADE Generic Facility ontology. Fiumicino airport (Rome, Italy) model (TBox) developed within EU FP7 CASCADE project
Scope (Domain)	facility management, operation, monitoring and controlling, devices/sensors
Use cases (Motivation, Relevance)	Ontologies used as part of a framework to reduce energy in airports is of particular interest because of the potential these types of buildings have. Airports consume as much energy as small cities. With successful demonstration at Fiumicino airport in Rome, the solution can be replicated in other airports around Europe, leading to potentially enormous energy savings and CO_2 emissions reduction.
Data sets	Apart from airport-internal private datasets from the Fiumicino airport, there are no other datasets that currently make use of the ontology.
Open issues/ Challenges	
Tool support	Created with TopBraid Composer

CASCADE Malpensa Airport ontology

Name	CASCADE Malpensa Airport ontology
Author and License	Institute Mihajlo Pupin: Sanja Vranes, Nikola Tomasevic, Marko Batic Unknown license
URI	http://jpo.imp.bg.ac.rs/cascade/airport-ontology/MXP/airportOntologyMXP_TBox.owl
Description	A full-blown ontology model of Malpensa airport (Milan, Italy) which models a specific airport infrastructure by classifying installed technical systems relevant to the energy management aspect. It was developed by extension and population of the CASCADE Generic Facility ontology. Malpensa airport (Milan, Italy) model (TBox) developed within EU FP7 CASCADE project.
Scope (Domain)	Airports
Use cases (Motivation, Relevance)	Even if this ontology is oriented to create a model of airport facility, it can be used also in generic buildings modelling, particularly public buildings or complexes, due to the commonality with airport sub-functions.
Data sets	
Open issues/ Challenges	The CASCADE deontology is characterized by a partial superposition with other ontologies taken into account (regarding geography or buildings).
Tool support	Created with TopBraid Composer

Energy in Buildings Ontology

Name	Energy in Buildings Ontology (EiBO)
Author and License	info@planergy.it , Unknown
URI	http://www.planergy.it/file/EiBO v1.owl
Description	The ontology developed in Planergy allow the semantic description of the phenomena inherent energy flows incoming and outgoing from a set of buildings immersed in their environment, by formaly allowing the description of : - the physical spaces (buildings and other sub objects) - the properties belonging to these physical spaces - the functionalities needed to support monitoring and measurement activities - the description of processes (administrative and economic) - the human presence in the spaces and their allocation

	- the terms used in different region to describe these spaces
Scope (Domain)	physical space, monitoring, measurement, roles, regions
Use cases (Motivation, Relevance)	It has been developed to support the administration in publishing open data related to energy performances of public buildings in Italy. It should foster the development of PPP with ESCOs
Data sets	
Open issues/ Challenges	
Tool support	The ontology is available in plain OWL

INERTIA Ontology

Name	INERTIA Ontology
Author and License	Peter Kostelnik, peter.kostelnik@gmail.com, All rights reserved / no license (No Open)
URI	http://www.inertia-project.eu/inertia/files/document/ontologies/inertia-schema.n3
Description	Ontology contains information describing the whole domain required for INERTIA pilot applications. Ontology serves as the common vocabulary used across all software components, but also serves as flexible support of describing and accessing all information and static data used in required by application logic of INERTIA pilots. Ontology describes whole location context, taxonomy of devices (Distributed Energy Resources, sensors, actuators) and occupancy model
Scope (Domain)	Location context, models of devices, occupancy model.
Use cases (Motivation, Relevance)	Based on semantic middleware prototype requirements, the design and development of INERTIA ontologies focuses mostly on the semantic model of BIM, IoTDevices and DER modelling. The Middleware is required to have access and control over different subsystems within the Local Hub. The main roles of the Semantic Based Middleware are to provide: • real-time information regarding a building's (or cluster of buildings') infrastructure and equipment • dynamic control over specific DERs Most use cases involve access to real-time information and/or dynamic control, either directly or indirectly through usage of historical databases of past events.

	Regarding explicit use of real-time data, the Inertia's Middleware is involved as a major component supporting UC 1 – Monitoring of Local Hub's Energy Data, UC 4 – Monitoring of personalized energy data and UC 2 – Automated real time control planning of the facility infrastructure based on contextual information , providing a continues stream of data about energy usage and contextual information from sensors and DERs in combination with descriptive data stemming from INERTIA ontologies. The Middleware will also allow the INERTIA system to use real time building occupancy detection from motion sensors and other contextual information that can be used as part of the background data for the spatio temporal occupancy flow models required for UC 3 – Automated real time control on building's DERs based on occupancy and scheduling information . The ability to control DERs such as HVAC and lighting from the Aggregator leve l is an integral part of in particular UC 10 – End user control of local Hub Portfolio . http://www.inertia-project.eu/inertia/files/document/deliverables/INERTIA_Deliverable_D3.1.pdf
Data sets	Example Dataset: http://www.inertia-project.eu/inertia/files/document/ontologies/event-dump.n3
Open issues/ Challenges	Unkown
Tool support	Via N3/RDF

INSPIRE Data Specification on Transport Networks

Name	INSPIRE Data Specification on Transport Networks
Author and License	Unknown
URI	http://cui.unige.ch/isi/onto/inspire-TN
Description	INSPIRE Data Specification on Transport Networks in OWL
Scope (Domain)	Transport
Use cases (Motivation, Relevance)	*Transport Networks* is defined within the INSPIRE Feature Concept Dictionary as: "The transport component should comprise an integrated transport network, and related features, that are seamless within each national border. In accordance with article 10.2 of the Directive, national transport networks may also be seamless at European level, i.e. connected at national borders. Transportation data includes topographic features related to transport by road, rail, water, and air. It is important that the features form net works where appropriate, and t hat links between different networks are established, i.e. multi-modal nodes, especially at the local level, in order to satisfy the requirements for intelligent

	transport systems such as location based services (LBS) and telematics. The transport network should also support the referencing of transport flow to enable our navigation services." [INSPIRE Feature Concept Dictionary] http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_TN_v3.0.pdf
Data sets	
Open issues/ Challenges	
Tool support	

CityGML Ontology

Name	CityGML Ontology
Author and License	Unknown
URI	http://cui.unige.ch/citygml/2.0/
Description	This OWL version of the CityGML standard has been created by (a) generating classes, properties and axioms from the CityGML 2 XML Schemas, (b) manually fixing some generation problems, (c) manually replacing every refernce to gml:xxxPropertyType by references to xxx, and (d) manually adding missing gml: classes, properties, and axioms for the geometry profile of CityGML
Scope (Domain)	City
Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

URBAMET Thesaurus

Name	URBAMET Thesaurus
Author and License	The urbamet association. No license, the Thesaurus is not available as OWL file.
URI	http://notx.documentation.developpement-durable.gouv.fr/Urbanisme/thesaurus/navigation.xhtml
Description	The URBANDATA Association is a consortium of urban information providers in European countries. It aims to improve the international exchange and dissemination of information about urban issues and to develop new products and services which will aid those processes. URBANDATA publishes the database website URBADOC which contains over 700,000 records of the literature on urban and social research, policy and practice in the countries of its five members and elsewhere. The French chapter of URBANDATA (the French association “urbamet”) has produced a Thesaurus which can be consulted on-line.
Scope (Domain)	The main subjects covered are: • Land management • Urban management • Architecture • Local government • Environment • Community facilities and amenities • Local finance • Urban infrastructure services • Housing • Pollution and conservation • Urban transportation
Use cases (Motivation, Relevance)	URBAMET is primarily intended for town planners, local elected representatives, architects and urban development professionals, as well as researchers and students, librarians and documentalists, etc
Data sets	The last two years of the databank are available with free access on this site. URBAMET can also be consulted on the Urbadoc web site, alongside the 6 other European databanks addressing these issues.
Open issues/ Challenges	Still to be converted into OWL…

Tool support	N/A

SAREF: the Smart Appliances REFerence ontology

Name	SAREF: the Smart Appliances REFerence ontology
Author and License	Laura Daniele (laura.daniele@tno.nl), Unknown
URI	http://ontology.tno.nl/saref
Description	The Smart Appliances REFerence (SAREF) ontology is a shared model of consensus that facilitates the matching of existing assets (standards/protocols/datamodels/etc.) in the smart appliances domain. The SAREF ontology provides building blocks that allow separation and recombination of different parts of the ontology depending on specific needs. The starting point of SAREF is the concept of device (e.g., a switch). Devices are tangible objects designed to accomplish a particular task in households, common public buildings or offices. In order to accomplish this task, the device performs one or more functions. For example, a washing machine is designed to wash (task) and to accomplish this task it performs the start and stop function. The SAREF ontology offers a list of basic functions that can be eventually combined in order to have more complex functions in a single device. For example, a switch offers an actuating function of type 'switching on/off'. Each function has some associated commands, which can also be picked up as building blocks from a list. For example, the 'switching on/of' is associated with the commands 'switch on', 'switch off' and 'toggle'. Depending on the function(s) it accomplishes, a device can be found in some corresponding states that are also listed as building blocks. When connected to a network, a device offers a service, which is a representation of a function to a network that makes the function discoverable, registerable and remotely controllable by other devices in the network. A service can represent one or more functions. A service is offered by a device that wants (a certain set of) its function(s) to be discoverable, registerable, remotely controllable by other devices in the network. A service must specify the device that is offering the service, the function(s) to be represented, and the (input and output) parameters necessary to operate the service. A device in the SAREF ontology is also characterized by an energy/power profile that can be used to optimize the energy efficiency in a home or office that are part of a building.
Scope (Domain)	Smart Appliances, Devices, Sensors, Actuators, Device functions, Services attached with devices
Use cases (Motivation, Relevance)	The Smart Appliances REFerence (SAREF) ontology is conceived as a shared model of consensus that facilitates the matching of existing assets in the smart appliances domain, reducing the effort of translating from one asset to another, since the SAREF ontology requires one set of mappings to each asset, instead of a dedicated set of mappings for each pair of assets. Using the SAREF ontology, different assets can keep using their own terminology and data models, but still can relate to each other through their common semantics. In other words, the SAREF ontology enables semantic interoperability in the smart appliances domain.

	The ontology is based on the fundamental principles of reuse and alignment of concepts and relationships that are defined in existing assets, modularity to allow separation and recombination of different parts of the ontology depending on specific needs, extensibility to allow further growth of the ontology, and maintainability to facilitate the process of identifying and correcting defects, accommodate new requirements, and cope with changes in (parts of) the SAREF ontology. The ontology mainly addresses the consumer (mass) market of the home, private dwellings, but also common public buildings and offices, and the standard appliances used in that environment. The appliances covered by SAREF ontology are: • Home and buildings sensors (temperature, humidity, energy meters, environmental sensors etc.) and actuators (windows, doors,). Sensors belonging to appliances are treated individually. • White goods, namely, rinsing and cleaning, cooking and baking, refrigerating and freezing, vacuum cleaning, washing and drying as well. • HVAC; heating, ventilation, air conditioning • Lighting
Data sets	The Smart Appliances reference (SAREF) ontology can be used to match the data from different organizations. Example of instances: saref_sampledata.ttl available at http://ontology.tno.nl/saref_sampledata (click on "Individuals" in the Navigation tab on the top right corner to visualize the sample data) download at http://ontology.tno.nl/saref_sampledata.ttl
Open issues/ Challenges	The SAREF ontology has been mapped on the ETSI M2M Architecture[46], and found that there is a good correlation between the ETSI M2M Architecture and SAREF's function-related device categories. The mapping with energy-related and building-related device categories is still minimal. For further implementation of SAREF into ETSI M2M, the SAREF ontology needs to be extended with ETSI M2M specific functionality, such as M2M Gateway, and Remote Management functionality.
Tool support	The SAREF ontology is expressed in RDF/OWL and serialized in Turtle (therefore, the file extension .ttl), which is a compact syntax alternative to RDF/XML. Thus, the ontology can be opened with any ontology editor, such as TopBraid Composer, Protégé and NeOn.

[46] http://www.etsi.org/technologies-clusters/technologies/m2m

DECT ULE ontology

Name	DECT ULE ontology
Author and License	TNO: Jasper Roes, Frank den Hartog, Laura Daniele, Jack Verhoosel Unknown license
URI	https://sites.google.com/site/smartappliancesproject/ontologies/dect_ule-ontology
Description	The DECT ULE ontology describes the DECT ULE HF standard, which is based on a star network topology of network entities. A HFNetworkEntity can be a HFConcentrator, which is the network's master device, or a HFDevice. There are up to thousands of devices supported by the concentrator and connected to it. The HF protocol supports several types of HF messages exchanged between network entities (i.e., commands, requests, responses), and each of these messages has a message type code. A HFMessage is structured in 3 fields (i.e., network, transport and application layers. It considers home, private dwellings, but also common public buildings and offices, and the standard appliances used in that environment. Elevators and other special equipment are not covered.
Scope (Domain)	DECT ULE HF standard; Star network topology; HF protocol;
Use cases (Motivation, Relevance)	The study covers the following interoperability use cases for Smart Appliances: - Interoperability with construction design tools (product information, product performance and product behaviour) - Interoperability with Facility Management and Energy Management Systems - Interoperability with Building Control systems - ESCO (Energy Services) systems - Interoperability with the Smart Grid
Data sets	
Open issues/ Challenges	Challenges: - Proposal for a unified ontology to be contributed to ETSI for consideration as a future standard. - Documentation of the proposed the ontology into the ETSI M2M architecture.
Tool support	

Echonet ontology

Name	Echonet ontology
Author and License	TNO (adaptation from the ECHONET consortium specification) in the frame of the Smart Appliances Study (2013/0077), License of the specification is "open to the public" only for versions 1.0 and 1.01
URI	https://sites.google.com/site/smartappliancesproject/ontologies/echonet-ontology
Description	The Echonet ontology represents Echonet device objects and their properties (Echonet: Energy Conservation and HOmecare NETwork (ECHONET) for Device Objects). A Device defines one or more DeviceObject. Device objects represent mechanical functions of a device and aim at facilitating controls and status verification through communications between devices. There are general properties applicable to any device object, such as hasOperationStatus. These general properties are defined as sub properties of the hasDeviceObjectProperty property.
Scope (Domain)	Echonet device objects, Echonet device properties, device mechanical functions, controls and status verification, device communication
Use cases (Motivation, Relevance)	Energy Conservation and homecare network for Device Objects.
Data sets	
Open issues/ Challenges	
Tool support	

EnOcean ontology

Name	Enocean: EnOcean Alliance Equipment Profile (EEP)
Author and License	Laura Daniele (laura.daniele@tno.nl), Unknown
URI	https://sites.google.com/site/smartappliancesproject/ontologies/enocean-ontology
Description	EnOcean[47] is a company that develops energy harvesting wireless sensors which are claimed to be maintenance free and flexible allowing cost reduction in buildings and industrial facilities. In 2012 this technology has subsequently been standardized as

[47] www.enocean.com

	ISO/IEC 14543-3-10.Full interoperability is guaranteed together with the EnOcean Equipment Profiles (EEPs) drawn up by the EnOcean Alliance[48]. The EnOcean Equipment Profile (EEP) contains information about devices "enabled by EnOcean", including RORG (identifies the EnOcean Radio Protocol (ERP) radio telegram type), FUNC (identifies the basic functionality of the data content), and TYPE (identifies the type of device in its individual characteristics). The Enocean ontology specifies the user data embedded in the structure of a radio telegram as defined by the EnOcean Equipment Profile (EEP). Therefore, the ontology defines an EEP_profile class. Through the hasElement property, the EEP_profile class is characterized by 3 elements: the RORG class, which represents the ERP radio telegram type using a code, for example, the value F6 represents an RPS telegram type; the FUNC class, which represents the basic functionality of the data contained in a radio telegram, for example, TemperatureSensor, AutomatedMeterReading , Detector , and HVAC_component; and the TYPE class, which represents the specific characteristics of a device type, for example, a temperature sensor with range between -10°C and 30°C (TemperatureSensor_range10Cto30C class) The ontology defines 4 types of telegrams according to the EEP profile, namely RPS, 1BS, 4BS and VLD, which are represented by the corresponding classes TelegramRPS, Telegram1BS, Telegram4BS, and TelegramVLD , respectively. Each telegram has a RORG (hasRORG property), and can have several device functions (hasFUNC property) and types (hasTYPE property). Each RORG class, FUNC class and TYPE class has a code (hasRorgCode property, hasFuncCode property and hasTypeCode property, respectively). These codes are used to assemble the 3 field code that characterizes a specific telegram.
Scope (Domain)	EnOcean, Equipment Profile, EEP, Device Types, Device Function, Sensors Function
Use cases (Motivation, Relevance)	The ontology could be utilized to model any EnOcean device/sensor/actuator under a common framework.
Data sets	-
Open issues/ Challenges	The TYPES are defined completely for the TelegramRPS and Telegram1BS classes. For the Telegram4BS class the TYPES are defined until and including the A5_10 subclass. For the TelegramVLD class the TYPES are not defined at all. For completeness, it is advised to add the remaining TYPES in the future.

[48] www.enocean-alliance.org

	The EEP document[49], which was used as a reference for the ontology, defines enumerations that are used to further characterize the specific TYPE of telegrams. These enumerations are too many and too detailed to be included in the current version of the ontology. However, the ontology could be extended in the future to cover also this aspect of the EnOcean Equipment Profile. The source used to create the ontology is a secured pdf from which the information could not be automatically copied. As a consequence, comments that could better explain the telegrams are missing in the ontology.
Tool support	The Enocean ontology is expressed in RDF/OWL and serialized in Turtle (therefore, the file extension .ttl), which is a compact syntax alternative to RDF/XML. Thus, the ontology can be opened with any ontology editor, such as TopBraid Composer, Protégé and NeOn.

FAN FPAI ontology

Name	FAN FPAI ontology
Author and License	Unknown
URI	https://sites.google.com/site/smartappliancesproject/ontologies/fan-ontology
Description	The Fanfpai ontology describes the resources (appliances) used in the Flexible Power Application Infrastructure (FPAI) . These resources are defined in the Resource Abstraction Interface (RAI class), which is used to express the energetic flexibility that appliances can offer and how this flexibility should be exploited. The RAI is an interface layer between: the Resource Abstraction Layer (RAL class) that monitors and controls the appliances and knows how much flexibility they can offer. The RAL consists of two main components: the resource manager (ResourceManager class) and the resource driver (not considered in this ontology); the energy apps (EnergyApp class) that are typically provided by a third party and exploit the flexibility that appliances have to offer. An energy app is only interested in exploiting energetic flexibility and not in the details of a specific appliance, such as a washing machine, for instance.
Scope (Domain)	appliances, household appliances, Flexible Power Application Infrastructure, FPAI
Use cases (Motivation, Relevance)	
Data sets	

[49] http://www.enocean-alliance.org/eep/

Open issues/ Challenges	
Tool support	

FIEMSER ontology

Name	Friendly Intelligent Energy Management Systems in Residential Buildings Data Model
Author and License	Laura Daniele (laura.daniele@tno.nl), Juan Pérez Project Coordinator (juan.perez@tecnalia.com), Unknown
URI	https://sites.google.com/site/smartappliancesproject/ontologies/fiemser-ontology
Description	The Fiemser ontology describes the main classes of the Energy-focused BIM model and WSN-related data that are part of the FIEMSER data model. The ontology describes the building space organization in terms of the Building, BuildingPartition, BuildingSpace and BuildingZone classes. A building partition defines a part of a building managed by either a dweller (e.g., a flat) or a facility manager (e.g., a common building area). A building space defines the physical spaces of the building. A building zone defines a functional area in the building that will be controlled as a unique zone. A building consistsOf some building partitions, a building partition consistsOf some building spaces, a building zone consistsOf some building spaces. The Fiemser ontology also describes the devices (Device class) used in the building in terms of HomeEquipment and ControlledDevice. A HomeEquipment is any home appliance or mechanism to increase building energy efficiency, such as Generator, Load, Mechanism and Storage. Generators represent devices that provide part of the energy required by the building, for example, PV (of type ElectricalGenerator) and Boiler (of type ThermalGenerator). Loads represent devices that consume energy and offer a service to the user, for example, TV (of type ElectricalLoad) and Radiator (of type ThermalLoad). Mechanisms represent devices that are installed in the home to increase its energy efficiency, but don not generate or consume energy by themselves, for example, a Blinder. Storage devices represent devices that store energy and can be used to provide convenient energy management strategy, for example, Battery (of type ElectricalStorage) and Tank (of type ThermalStorage). A ControlDevice represents a device directly connected to the FIEMSER control infrastructure and used to monitor and/or control the environment and its appliances. A control device consistsOf some ControlComponent that can be a hardware component (Sensor or Actuator or CommDevice) and a software component. An Actuator is any actuating hardware installed in a control device, such as a Dimmer, Switch and Controller. A Sensor can be a MeasurementSensor (e.g., thermostat) or StateSensor (e.g., presence). A communication device (CommComponent) identifies the communication devices used for data exchange and uses a specific Network protocol (NetProtocol class).
Scope (Domain)	BIM, WSN, Building space description, Climate, Location, Devices, Devices in a building, Energy Consumption, Home Usage Profile, Price, Device Schedule

Use cases (Motivation, Relevance)	FIEMSER FP7 European R&D project's[50] objective was the development of an innovative energy management system for existing and new residential buildings, which pursues the increase of the efficiency of the energy used and the reduction of the global energy demand of the building, but without penalizing the comfort levels of the users. The core motivation is the minimization of the energy demand from external resources and the management of local energy consumption/production/storage. Since special emphasis was given on the interoperability with architectural CAD tools and building energy simulation tools, the gbXML data model was selected as reference data model for the FIEMSER development.
Data sets	The specific sub-models used to create the FIEMSER data model belong to the following corresponding categories of data: Environmental and Contextual data (ENV), Energy-focused Building Information Model (BIM), Data from sensors (WSN), User Preferences (USR), Resources scheduling data (SCH), Advices (ADV), Energy Performance Indicators (EPI), and User access right (RGH).
Open issues/ Challenges	The Fiemser ontology describes the main classes of the Energy-focused BIM model and WSN-related data that are part of the FIEMSER data model. Although also the other 6 models of the FIEMSER data model contain relevant information, it was not possible to include them in the current version of the ontology. It is therefore advised to do so as part of future work. The source used to create the ontology is a secured pdf from which the information could not be automatically copied. As a consequence, comments that could better explain the ontology may be missing.
Tool support	The FIEMSER ontology is expressed in RDF/OWL and serialized in Turtle (therefore, the file extension .ttl), which is a compact syntax alternative to RDF/XML. Thus, the ontology can be opened with any ontology editor, such as TopBraid Composer, Protégé and NeOn.

FIPA Device Ontology

Name	FIPA Device Ontology: Foundation for Intelligent Physical Agents BDevice Ontology Specification
Author and License	Laura Daniele (laura.daniele@tno.nl), gateways@fipa.org, All rights reserved / no license (No Open)
URI	https://sites.google.com/site/smartappliancesproject/ontologies/fipa.ttl
Description	In 2002, the then existing FIPA Gateways TC published an ontology for describing devices and their properties.

[50] www.fiemser.eu

	The FIPA ontology describes a device ontology that aims at enabling interoperability between software agents, as defined by the FIPA Device Ontology Specification. This ontology can be used by agents when communicating about devices: when agents pass profiles of devices to each other, these profiles can be validated using the information contained in this ontology. The main class of the ontology is the Device class, which defines a device and its general properties. A device has some InfoDescription, such as the name, vendor and version of the product under consideration, and has some hardware and software properties. Software properties include the details of the device's operating system (hasOperatingSystem), such as its name, vendor and version. Hardware properties are the type of connection that the device uses (hasConnection), the amount of memory that it requires (hasMemory), the user interfaces offered by the device (hasUserInterface), and the type of central processing unit (hasCPU). The connection type is expressed in terms of name, vendor and version of the connection provider (hasConnectionInfo). The MemoryTypeDescription class defines the unit of measure of the memory (hasMemoryUnit), and its usage type, namely application, storage, or both application and storage (hasMemoryUsageType). The UIDescription class defines the information that characterize the screen of the device (hasScreen), such as its width (hasWidth), height (hasHeight), resolution (hasResolution), and the measurement units (hasWidhtHeightUnit). The ontology also defines the RequestDeviceInfo function that can be used in the FIPA framework by an agent to make a query to request the device information contained in the ontology.
Scope (Domain)	Device, Device interoperability, Device description, Profile, Software agents,
Use cases (Motivation, Relevance)	The FIPA ontology can be used by agents when communicating about devices. Agents pass profiles of devices to each other and validate them against the FIPA ontology. The profiles come in handy for example in a situation where memory- or processing-intensive actions take place; agent A1 can ask agent A2 whether device D has enough capabilities to handle some task A1 has in mind.
Data sets	-
Open issues/ Challenges	The OWL version of the FIPA ontology has been created according to the FIPA device ontology specification[51]. This specification refers to some classes defined in other FIPA ontologies, namely the FIPA-Nomadic-Application and FIPA-Agent-Management ontologies. These ontologies have not been translated to OWL. However, the Fipa ontology can be extended to consider the FIPA-Nomadic-Application by using the AgentPlatform class, and the FIPA-Agent-Management ontologies by using the QoS class.
Tool support	The FIPA ontology is expressed in RDF/OWL and serialized in Turtle (therefore, the file extension .ttl), which is a compact syntax alternative to RDF/XML. Thus, the ontology can be opened with any ontology editor, such as TopBraid Composer, Protégé and NeOn.

[51] http://www.fipa.org/specs/fipa00091/SI00091E.html

Hydra Basic Device Information ontology

Name	HYDRA ontology: Heterogeneous physical devices in a distributed architecture ontology
Author and License	Dr. Markus Eisenhauer Project Coordinator (markus.eisenhauer@fit.fraunhofer.de), Unknown
URI	https://sites.google.com/site/smartappliancesproject/ontologies/hydra-ontology
Description	Although there are several ontologies developed in the Hydra project, the Hydra Basic Device Information ontology has been included consisting of the following modules: i) Basic Device Information, ii) Device Services, iii) Device Events, iv) Device Malfunctions, v) Device Capabilities and vi) State Machine. The Basic Device Information module represents general device information. The HydraDevice is the main ontology class, which is further divided in the PhysicalDevice and the SemanticDevice classes. Physical and semantic devices share common device properties, such as deviceId or inLocation, but have different semantic interpretation and behaviour. The HydraDevice class refers to the InfoDescription class using the info property. The InfoDescription class contains basic information about device friendlyName, manufacturer data, i.e., manufacturerName and manufacturerURL, and device model data, i.e., modelName, modelDescription and modelNumber. An important part of the basic device information is the representation of device type modelled as sub classes of the PhysicalDevice concept, such as SensorDevice, ActuatorDevice, MediaDevice and MobileDevice. Further, the hasEmbeddedDevice property of the SemanticDevice class recursively refers to HydraDevice concept. This property enables the creation of models of composite devices, such as in case of the HeatingSystem device, which can be, for example, composed of Thermometer and Pump devices.
Scope (Domain)	Physical device, Device information, Device Modelling, Device Services, Device Malfunctions
Use cases (Motivation, Relevance)	HYDRA aims to interconnect devices, people, terminals, buildings, etc., not only providing interoperability at a syntactic level, but also at a semantic level. Hydra relies on semantic descriptions/annotations to expose device capabilities (using ontologies) so that applications can understand these capabilities and use them.
Data sets	-
Open issues/ Challenges	The proposed Hydra device services model represents one possible approach to service modelling and may be subject to further investigation and research related to possible existing and future semantic service mark-up standards (such as WSMO) and the system architecture requirements. Ontology changes can be caused from user requirements on changes to structure and classification; in Hydra this would be the developer users' requirements. The changes can also be induced by changes in the underlying domain objects being modelled by the ontology, in Hydra; this would be changes in device capabilities, in security protocols etc.

Tool support	The ontology can be opened with any ontology editor, such as TopBraid Composer, Protégé and NeOn.

SmartCoDE ontology

Name	SmartCoDE ontology
Author and License	, Unknown
URI	https://sites.google.com/site/smartappliancesproject/ontologies/smartcode-ontology
Description	The Smartcode ontology presents a classification of Energy using Products (EuPs) into seven categories, namely variable services (VARSVC class), thermal services (THMSVC class), schedulable services (SCDSVC class), event-timeout services (ETOSVC class), charge control (CHACON class), complete control (COMCON class), and custom control (CUSCON class). These products have some parameters, such as Configuration, OnlineInput and SensorInput. Each product is characterized by an energy management strategy (hasEnergyManagementStrategy property) and its cost profile can be of interest of not for energy management purposes (isCostProfileInteresting property).
Scope (Domain)	Energy, classification of energy products, energy measurement
Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

Mirabel ontology

Name	Mirabel ontology
Author and License	SAP AG's MIRABEL team is led by Dr. Gregor Hackenbroich, whose main research interests concern the management of structured and unstructured data as well as the integration of events into business software, and by Dr. Henrike Berthold, whose main interests lie on Business Intelligence and modern architectures for data management systems. http://www.mirabel-project.eu/ (no other contact info), All rights reserved / no license (No Open)

URI	https://sites.google.com/site/smartappliancesproject/ontologies/mirabel-ontology
Description	The Mirabel ontology defines how actors can express their energy flexibility for a specific device with respect to amount, time and price in user preferences. Each device has an energy profile that describes the amount of energy consumed and/or produced over a time span. A flex offer is issued by an actor and combines the user preferences with the corresponding device energy profile.
Scope (Domain)	actors, energy flexibility, user preferences, energy profile, energy flexibility
Use cases (Motivation, Relevance)	This ontology gives a semantically better view on the flexibility concept and its meaning in relation to the building on the one hand and the smart grid on the other hand. Moreover, this ontology forms the basis for a vocabulary that can be published via the web and used to connect IT systems from various stakeholders in the energy domain that handle supply and demand of energy.
Data sets	
Open issues/ Challenges	EU project result. Still maintained?
Tool support	

Stream Annotation Ontology – SAO

Name	Stream Annotation Ontology - SAO
Author and License	Institute for Communication System, University of Surrey: Sefki Kolozali http://creativecommons.org/licenses/by/3.0/
URI	http://iot.ee.surrey.ac.uk/citypulse/ontologies/sao/sao.rdf
Description	aims to semantically represent the features of a stream data. It allows publishing content-derived data about IoT streams and provides concepts such as StreamData, Segment, StreamAnalysis on top of the TimeLine concepts. Timeline Ontology extends OWL-Time with various timelines (e.g.\ universal or discrete), temporal concepts, such as Instant, and Interval, and interval relationships. The SAO uses the broad definition of the StreamEvent concept in order to express an artificial classification of a time region, corresponding to a particular stream data. It also extends the sensor observations described in SSN Ontology ssn:Observation through a concept, StreamData, that allows to describe Segment or Point linked to time intervals or time instants. Below is the depiction of the workflow of the SAO Ontology.
Scope (Domain)	Internet of Things, stream data

Use cases (Motivation, Relevance)	Representing IoT data streams is an important requirement in semantic stream data applications, as well as in knowledge-based environments for Smart Cities. The project had identified 101 smart city scenarios and related use cases (http://www.ict-citypulse.eu/scenarios/) in cooperation with partner cities and city cooperation (City Stakeholder Group) and derived a set of requirements for a smart city framework based on proposed use cases, references in the field and "on site" workshops together with city partners.
Data sets	
Open issues/ Challenges	
Tool support	

Adapt4EE Ontology

Name	Adapt4EE Ontology
Author and License	Dr. Dimitrios Tzovaras Project Coordinator (Dimitrios.Tzovaras@iti.gr), Unknown
URI	http://www.adapt4ee.eu/adapt4ee/results/ontologies.html
Description	The Adapt4EE[52] ontology constitutes a formal model for enterprise energy performance measuring, monitoring and optimization.Adapt4EE semantic enterprise model treats, learns and manages the enterprise environment as an intelligent agent, perceives environmental state using multi-type sensors and information modalities. The Adapt4EE Data Model incorporates business processes and occupancy data. I The TTL files of the overall Adapt4EE Ontology have been utilized for the scope of the Adapt4EE project. The files include: Adapt4EE Building Information Model (BIM) Adapt4EE Business Process Model (BPM) Adapt4EE Common Information Model Adapt4EE Device Model Adapt4EE Event Model Adapt4EE KPI Model Adapt4EE Occupancy Model Adapt4EE Units Mode.
Scope (Domain)	BIM, BPM, Device, Events, Occupancy, Building Automation, Building Performance optimization, energy efficiency
Use cases (Motivation,	The Adapt4EE Enterprise Models allow for the proactive identification of optimum local adaptations of enterprise utility operations, based on predictions of possible occupancy patterns and respective business operations and energy profiles.

[52] www.adapt4ee.eu

Relevance)	The semantic coverage and subsequently the potential usage of the ontology is partially overlapping with the results from the HYDRA project.
Data sets	The model has been calibrated during the training phase based on sensor data captured during operation and then applied and evaluated in real-life every day enterprise operations
Open issues/ Challenges	-
Tool support	The Adapt4EE ontology is serialized in Turtle (therefore, the file extension .ttl), which is a compact syntax alternative to RDF/XML. Thus, the ontology can be opened with any ontology editor, such as TopBraid Composer, Protégé and NeOn.

ROUTE - Route Ontology of Urban Transportation Entities

Name	ROUTE - Route Ontology of Urban Transportation Entities
Author and License	Diarmuid Ryan (diarmuid.ryan@ucdconnect.ie), Achilleas Psyllidis (A.Psyllidis@tudelft.nl), Oudom Kem (oudom.kem@emse.fr), Matthew Horrigan (matthew.horrigan@ucdconnect.ie), CC By 3.0
URL	http://labs.geodata.gov.gr/en/dataset/urban-transportation-routes-athens
Description	This ROUTE ontology describes public urban transportation routes. It also describes concepts pertinent to trip services, pickup and drop-off types, time intervals, frequency, geographical information about stops, among other related concepts.
Scope (Domain)	Athens, services, bus stops, stop times, transportation network
Use cases (Motivation, Relevance)	Athens, services, bus stops, stop times, transportation network
Statistics	None provided
Questions	Available as a zip file

9 Collected datasets

9.1 Gap analysis

The availability of open linked data related to energy in general is scarce. There are some online portals offering relevant data which is largely not open (e.g. data from *Eurostat*), and of which only a small part specifically addresses the energy domain. Such example is www.engagedata.eu which offers some 253 datasets tagged with the keyword 'energy', however, a closer inspection reveals that not all data is in an open format (e.g. *rdf*) or freely available, with some of the provided links leading to data with restricted access. Similarly, www.publicdata.eu has more than a thousand hits relating to energy, the majority of them provided in formats like *xls*, *csv* and *html*.

Popular portals such as www.datahub.io also offer a variety of datasets that are potentially interesting for Ready4SmartCities, but only a few of them are open (a general energy-related search returned ca. 630 results, of which only 12 were *rdf+xml*, and 7 *api/sparql*).

A portal concentrated solely on offering open linked data online and for free is hitherto not available to our knowledge. www.smartcity.linkeddata.es is the first of its kind that offers linked open datasets with immediate overview of their availability, form, license, etc. However, due to the lack of organizations publishing their data as linked and open, the catalogue experiences slow growth in terms of new content being uploaded on the website. Feedback through the online survey used to screen for new datasets is rare, and the involvement of the community identified in WP1 seems to be harder compared to ontologies. Possible ways to increase interest and participation with respect to datasets are discussed in Part C *Conclusions*.

The most relevant data for this project seems to be resulting from different initiatives/projects, such as the *Energy efficiency assessments and improvements* dataset, a comprehensive dataset that demonstrates the power of linked open data by covering assessments from Sweden and the US. Of the identified datasets, *Linked Clean Energy Data* is perhaps the most comprehensive, as it covers domains such as policy and regulatory country profiles, key stakeholders, project outcome documents, thesaurus, renewables, energy efficiency, climate change.

With 18 datasets it is impossible to perform a meaningful analysis due to the low number of datasets. The aim is to identify data that belongs to domains not yet covered in order to achieve certain diversity and make recommendations with regards to datasets for Energy Measurement and Validation.

Specifically for the domain of energy management systems interoperability, there are high demands regarding security and privacy issues. Also, there are rather complex data structures and a huge amount of data so that it seems that there is a natural barrier for publishing data on the web. In that respect, there are still a lot of open questions to be discussed and solved. Additionally, there is still a lack of clear business cases for data owners to open their data and to justify additional efforts to transfer and host the data in the web. All these circumstances might explain why there is only very few open linked data available. In general, found datasets are either results of research projects or somehow driven by public authorities. From industry a natural interest is driven by marketing use cases, i.e. provision of open data to advertise their products. Accordingly, they typically focus on unique selling features instead of providing neutral and comparable product descriptions.

Our preliminary conclusion about availability of open datasets in the area of Energy Management Systems and Energy Measurement and Validation is quite disappointing. The following section summarizes the result of our research and, not claiming to give a complete picture of the current situation, it shows the challenges of providing a critical mass of data to be a sound basis to build new applications or point of information.

9.2 List of datasets

The European Building and Construction Materials Database for the Semantic Web

Name	The European Building and Construction Materials Database for the Semantic Web
Author and License	Andreas Radinger, Martin Hepp, Otto Handle unknown license (data mapped from the Eurobau database available at http://eurobau.com/)
URL	http://semantic.eurobau.com/sitemap.xml (for fetching all data) http://semantic.eurobau.com/eurobau-utility.owl (ontology) http://linkeddata.uriburner.com/sparql (public SPARQL endpoint) http://eurobau.com/ (source)
Description	Major dataset of the European building and construction materials market for the Semantic Web on the basis of the GoodRelations Web Vocabulary for E-Commerce. (see http://semantic.eurobau.com/)
Scope (Domain)	Construction Materials
Use cases (Motivation, Relevance)	Comparison of products? Search for products
Statistics	81 Manufacturers / Brands 19 Resellers 183 Warehouse locations 56.360 Product types (including variants) 1.783.798 Offerings 95 % of the product models include rich FreeClassOWL descriptions
Questions	

Daily Global Weather Measurements, 1929-2009 (NCDC, GSOD)

Name	Daily Global Weather Measurements, 1929-2009 (NCDC, GSOD)
Author and License	National Climate Data Center (NCDC) unknown license
URL	http://aws.amazon.com/datasets/Climate/2759;

	http://www7.ncdc.noaa.gov/CDO/cdoselect.cmd?datasetabbv=GSOD&countryabbv=&georegionabbv=
Description	A collection of daily weather measurements (temperature, wind speed, humidity, pressure, &c.) from 9000+ weather stations around the world. Historical data are generally available for 1929 to the present, with data from 1973 to the present being the most complete.
Scope (Domain)	Climate
Use cases (Motivation, Relevance)	The US National Climatic Data Center has been collecting weather data at stations around the globe since 1929. In particular, the Global Summary of the Day contains samples of surface weather data like rainfall, temperature, wind speed, etc.
Statistics	9000+ monitored weather stations ca. 20 field names with types (integer, float, boolean) and description (e.g. measurement – miles, Fahrenheit, milibars, knots, inches)
Questions	The dataset can only be used within the United States. The bulk data is quite large (20GB) and is therefore not quickly obtainable/downloadable. A demo/snippet of the data would be helpful for organisations seeking to explore and make use of it.

Repener building energy

Name	Repener building energy
Author and License	Álvaro Sicilia et.al. Creative Commons Attribution
URL	http://arcdev.housing.salle.url.edu/repener/sparql
Description	Integrated information of the Spanish territory, regarding energy certification, building monitoring, and geographical data
Scope (Domain)	energy efficiency, energy certification
Use cases (Motivation, Relevance)	
Data sets	
Open issues/ Challenges	
Tool support	

Enipedia

Name	Enipedia
Author and License	TU Delft
URL	http://enipedia.tudelft.nl/wiki/Main_Page
Description	Enipedia is an active exploration into the applications of wikis and the semantic web for energy and industry issues. Through this we seek to create a collaborative environment for discussion, while also providing the tools that allow for data from different sources to be connected, queried, and visualized from different perspectives.
Scope (Domain)	energy and industy issues
Use cases (Motivation, Relevance)	
Data sets	http://enipedia.tudelft.nl/wiki/Special:SparqlExtension
Open issues/ Challenges	
Tool support	

Linked Clean Energy Data

Name	Linked Clean Energy Data
Author and License	Florian Bauer, Renewable energy & energy efficiency partnership, http://www.reeep.org/ OGL license (UK Open Government License)
URL	www.reegle.info/downloads/latest_reegle_dump.nt
Description	A comprehensive set of linked clean energy data on several domains.
Scope (Domain)	Policy and regulatory country profiles, key stakeholders, project outcome documents, thesaurus, renewables, energy efficiency, climate change
Use cases (Motivation,	Apart from helpful documentation like project outcomes and a thesaurus, the data give insight into other domains relevant to the work in Ready4SmartCities, such as stakeholders, as well as climate data. Energy efficient measures that meet the regulations

Relevance)	and policies of the respective country also need to be taken into consideration when planning any energy efficiency related activities.
Statistics	
Questions	

State Energy Data System (SEDS)

Name	State Energy Data System (SEDS)
Author and License	U.S. Energy Information Administration (EIA) unknown license The data collected by EIA surveys forms (http://www.eia.gov/survey/) are for the most part not proprietary and available. For users eager to dive deeper there are assembled tools to access searchable databases.
URL	Assembled tools are available to customize searches, view specific data sets, study detailed documentation, and access time-series data. • http://api.eia.gov/ Application Programming Interface (API) is a machine readable format which can serve all customers for free, though a registration key is needed for access. (For further information see: http://www.eia.gov/developer/) • http://www.eia.gov/beta/api/bulkfiles.cfm The bulk download facility provides the entire contents of each major API data set in a single ZIP file. • http://www.eia.gov/tools/models/datatools.cfm Additional set of data tools for exploiting data from different domains.
Description	The State Energy Data System (SEDS) is the source of the U.S. Energy Information Administration's (EIA) comprehensive state energy statistics. SEDS is aimed to create historical time series of energy production, consumption, prices, and expenditures by state for analysis and forecasting purposes. (For further information see: http://www.eia.gov/state/seds/)
Scope (Domain)	Consumption, Prices and Expenditures, Production
Use cases (Motivation, Relevance)	There are many use cases for smart cities where energy data system is of relevance: • Historical time series of energy production / consumption, prices and expenditures • Energy Analysis • Exploitation of data for prediction purposes
Statistics	408,000 electricity series organized into 29,000 categories

	30,000 State Energy Data System series organized into 600 categories 115,052 petroleum series and associated categories 11,989 natural gas series and associated categories 132,331 coal series and associated categories (released Feb 25, 2014) 3,872 Short-Term Energy Outlook series and associated categories (released May 27, 2014) 368,466 Annual Energy Outlook series and associated categories (released May 27, 2014)
Questions	
Name	State Energy Data System (SEDS)

Energy efficiency assessments and improvements

Name	Energy efficiency assessments and improvements
Author and License	Department of Energy http://www.eia.gov/consumption unknown license
URL	data-gov.tw.rpi.edu/raw/10/data-10.nt.gz
Description	This is a linked dataset (in RDF) for demonstrating the power of linked data, through linking data about energy efficiency assessments from Sweden and the US. Additionally, the dataset links to other linked data sources in Sweden, such as the SNI-codes and LKF-datasets from Statistics Sweden (SCB). The data itself is constructed by transforming and re-publishing parts of three existing open datasets; results from the PFE and EKC projects at the Swedish Energy Agency, and the IAC assessment and recommendation database.
Scope (Domain)	Energy efficiency assessment, measures for energy efficiency improvements, saved energy, cost
Use cases (Motivation, Relevance)	The dataset contains information primarily about suggested (and/or implemented) measures for energy efficiency improvements, including data about the amount of energy saved, costs involved, the nature of the improvement and measure taken, as well as basic information of the assessed organisation.
Statistics	
Questions	

Residential Energy Consumption Survey

Name	Residential Energy Consumption Survey
Author and License	Department of Energy CC-BY-SA Creative Commons Attribution-ShareAlike Unported (Open)
URL	http://www.eia.gov/consumption/
Description	Survey (RECS), which is conducted every four years, provides national statistical survey data on the use of energy in residential housing units including physical housing unit types, appliances utilized, demographics, fuels, and other energy use information. This dataset (i.e., the full RECS dataset) is very large in size and may require specialized software to open on your computer
Scope (Domain)	Residential energy consumption data , energy consumption , energy use , Household use of energy , data , federal data download , national , housing , appliances , RECS data , energy , federal datasets , energy data , statistics
Use cases (Motivation, Relevance)	
Statistics	
Questions	

Housing market indicators

Name	Housing Market Indicators
Author and License	ODC@communities.gsi.gov.uk, http://www.nationalarchives.gov.uk/doc/open-government-licence/
URL	http://opendatacommunities.org/data.rdf
Description	A dataset of indicators of the state of the UK housing market, including affordability, ownership and supply, Right to Buy, dwelling sock, empty homes, housing waiting lists, net supply and tenure
Scope (Domain)	housing market, indicators,
Use cases (Motivation, Relevance)	Public and open access to local data in UK.

Statistics	20 datasets listed: 1. Administrative geography – discontinued: 342730 triples 2. Additional Affordable Dwellings: 222720 triples 3. Domestic Energy Performance Certificates Lodged on Register - By Floor Area: 80370 triples 4. Council Tax Band D Average: 22592 triples 5. Domestic Energy Performance Certificates Lodged on Register - By Energy Efficiency Rating: 500080 triples 6. Domestic Energy Performance Certificates Lodged on Register - By Environmental Impact Rating: 500080 triples 7. Civil Parish Council Tax Level Data: 467334 triples 8. Council Tax Requirement: 234200 triples 9. Council Tax Chargeable Dwellings: 21296 triples 10. Dev - Local Authorities: 79152 triples 11. Dev - Local Authority Buildings: 3520 triples 12. Dev - Local Authority Services: 363435 triples 13. Duty owed, but no accommodation secured: no information about triples 14. Enterprise Zones: 4110 triples 15. Fire Authorities: 444 triples 16. Collection of council tax and non-domestic rates: 198816 triples 17. Average weekly social rent of new PRP general needs lettings, 2012/2013, England, District By Number of Bedrooms: 12728 triples 18. Administrative geography: 3535460 triples 19. Administrative geography data from Ordnance Survey: 53799 triples 20. Council Tax Estimated Collection Rate: 11979 triples
Questions	

INERTIA Ontology dataset instance

Name	INERTIA Ontology dataset instance
Author and License	Peter Kostelnik (peter.kostelnik@gmail.com), Creative Commmons Attribution-NonCommercial 2.0 Generic (CC BY-NC 2.0)
URL	http://www.inertia-project.eu/inertia/files/document/ontologies/dataset-iti-building.n3
Description	Complete ontology instance used in 2nd year project review. Dataset describes whole location context for pilot building together with device equipment. More specifically, the data selected to be published comprise of a set of event-based data collected during one representative day from the multi-sensorial infrastructure deployed at the main INERTIA project's pilot site (a tertiary building with offices and a kitchen at CERTH premises in Thessaloniki, Greece).
Scope (Domain)	Location context, models of devices, consumption data, environmental data, occupancy model

Use cases (Motivation, Relevance)	The dataset example which is publicly available can be utilized as a simple instantiation for the INERTIA ontology.[53] In general, the data produced during the whole pilot implementation in CERTH premises include real-time and event-based information about distributed energy resources (DERs) consumption behaviour, environmental conditions inside and outside the pilot (temperature, humidity etc.), applied and automated control actions in the DERs as well as building occupants, and group-based and individual detection (RFID-Radio Frequency Identification detection system). The event-based data are recorded towards optimal and automated decision making in real-time without compromising users needs and comfort.
Statistics	The event-based dataset selected provides a representative example of events generated during one day: 2014-12-03. The dataset contains 97507 sensor events (environmental sensors, power consumption sensors, device actuators, etc.) available as semantic information.
Questions	-

Number of dwellings by tenure and district in the UK

Name	Number of dwellings by tenure and district in the UK
Author and License	contactus@communities.gsi.gov.uk, Unknown
URL	http://opendatacommunities.org/data/housing-market/dwelling-stock/tenure
Description	This dataset covers the years 2009 to 2013 and shows district level information with a tenure breakdown between local authority, Private Registered Providers (PRPs, formerly known as Housing Associations or Registered Social Landlords), other public sector and private sector. Figures for 2012 and 2013 are provisional. Private Registered Provider stock Information on PRP stock prior to 2012 comes from the Tenant Services Authority (TSA) Regulatory and Statistical Return (RSR). From April 2012, the TSA has become part of the Homes and Communities Agency (HCA) and information on PRP stock is now published in their annual Statistical Data Return (SDR). The SDR (and the RSR in the past) is completed by all PRPs every year in one of two variants; with PRPs owning or managing fewer than 1000 properties completing a shorter, less detailed form than those owning or managing 1000 or more properties. Other public sector dwellings 'Other' public sector dwellings follow the Census definition of a dwelling and include dwellings owned by any public sector body other than lower-tier local authorities (district councils, unitary authorities, metropolitan district councils and London boroughs) or Private Registered Providers (housing associations). This category includes dwellings owned by government departments (e.g. Ministry of Defence) and other public sector agencies (e.g. the NHS, the Forestry Commission, the Prison Service or county councils). Please note that it includes

[53] http://www.inertia-project.eu/inertia/files/document/ontologies/inertia-schema.n3

	dwellings that are vacant even if they are scheduled for demolition at a future date. Private sector stock Private sector stock is split into owner-occupied (OO) and private rental sector (PRS). There is no direct measure of either of these tenures due to the difficulty of collecting this private information and the relatively fluid interchange between these two parts of the private dwelling stock. The current methodology calculates an estimate of the PRS using information from the Labour Force Survey (LFS) and English Housing Survey (EHS). This data was derived from Table 100, available for download as an Excel spreadsheet. For fuller information please see the 'Dwelling Stock Estimates:2013, England' statistical release available in PDF format.
Scope (Domain)	towns, cities, dwellings, government, UK, national
Use cases (Motivation, Relevance)	Public and open access to local data in UK.
Statistics	68538 triples
Questions	

Impact indicator: energy efficiency of new build housing in the UK

Name	Impact indicator: energy efficiency of new build housing in the UK
Author and License	Department for Communities and Local Government (http://opendatacommunities.org/data/transparency/impact-indicators/energy-efficiency-new-builds) License: OGL http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2/
URL	http://opendatacommunities.org/data.rdf
Description	Average Standard Assessment Procedure energy rating score. How the figure is calculated: The sum of SAP energy rating scores for each new home for which an energy performance certificate has been issued in the reporting period, divided by the number of new homes for which a certificate has been issued. It is the average of the large number of scores calculated for new dwellings during the reporting period. Why is this indicator in the business plan? This is a key housing measure for which DCLG has policy responsibility. It monitors the energy efficiency of new build homes. How often is it updated? Quarterly Where does the data come from? National Energy Performance Certificate Register. Published figures are available here. What area does the headline figure cover? England Are further breakdowns of the data available? Yes, can be split by dwelling type. What does a change in this indicator show? An increase in this indicator would show an average increase in the energy efficiency of new homes. The average SAP rating is expected to gradually rise over the long-term as a growing proportion of new homes are completed to the 2010 Building Regulations standard, which requires more energy efficient new homes. Time Lag. Published within two months of the end of the reporting period. Next available

	update. To be confirmed. Type of Data. Official Statistics. Robustness and data limitations. Average figures are volatile due to a number of factors including the small number of new homes being assessed, the mix of dwelling types, the mix of heating systems used in new developments and the location of those developments. Links to Further Information https://www.gov.uk/government/organisations/department-for-communities-and-local-government/series/code-for-sustainable-homes-statistics#publications Contact Details CorporatePerformance@communities.gsi.gov.uk
Scope (Domain)	energy efficiency, housing, new buildings, impact indicator
Use cases (Motivation, Relevance)	The sum of SAP energy rating scores for each new home for which an energy performance certificate has been issued in the reporting period, divided by the number of new homes for which a certificate has been issued. It is the average of the large number of scores calculated for new dwellings during the reporting period. An increase in this indicator would show an average increase in the energy efficiency of new homes. The average SAP rating is expected to gradually rise over the long-term as a growing proportion of new homes are completed to the 2010 Building Regulations standard, which requires more energy efficient new homes.
Statistics	Data (from England) comes from National Energy Performance Certificate Register.
Questions	

Vehicle Traffic Data, Provided by City of Aarhus in Denmark

Name	Vehicle Traffic Data, Provided by City of Aarhus in Denmark
Author and License	Daniel Puschmann Centre for Communication Systems Research (CCSR) University of Surrey, UK email: d.puschmann@surrey.ac.uk,, Unknown
URL	http://iot.ee.surrey.ac.uk:8080/datasets.html#traffic
Description	A collection of datasets of vehicle traffic, observed between two points for a set duration of time over a period of 6 months (449 observation points in total), a CityPulse EU FP7 project initiative.
Scope (Domain)	traffic data, sensor measurements,temperature conditions, location nodes
Use cases (Motivation, Relevance)	Traffic monitoring for the purposes of the CityPulse EU FP7 project.
Statistics	The data is available in raw (CSV) and semantically annotated format (RDF Triple Language Turtle format) and the whole dataset consists of 3 batches depicting the different

	time periods of traffic data collection, while each one of them can be downloaded separately. Batch 1: February 2014 - June 2014 (http://iot.ee.surrey.ac.uk:8080/datasets/traffic/traffic_feb_june/index.html) Batch 2: August 2014 - September 2014 (http://iot.ee.surrey.ac.uk:8080/datasets/traffic/traffic_june_sep/index.html) Batch 3: October 2014 - November 2014 (http://iot.ee.surrey.ac.uk:8080/datasets/traffic/traffic_oct_nov/index.html) Metadata for Observation Points and Cross-observation point data are provided.
Questions	Vehicle Traffic Data, Provided by City of Aarhus in Denmark

Parking Data Stream, Provided by City of Aarhus in Denmark

Name	Parking Data Stream, Provided by City of Aarhus in Denmark
Author and License	Daniel Puschmann Centre for Communication Systems Research (CCSR) University of Surrey, UK email: d.puschmann@surrey.ac.uk,, CC-BY Creative Commons Attribution Unported (Open) http://creativecommons.org/licenses/by/4.0/
URL	http://iot.ee.surrey.ac.uk:8080/datasets.html#parking
Description	A datastream with parking data provided from the city of Aarhus.
Scope (Domain)	parking data, transportation data, parking lots
Use cases (Motivation, Relevance)	Parking monitoring for the purposes of the CityPulse EU FP7 project.
Statistics	There are a total of 8 parking lots providing information over a period of 6 months (55.264 data points in total). Data selected from May 22nd 2014 - November 4th 2014.
Questions	Available as CVS and Turtle

Pollution Data, Provided by City of Aarhus in Denmark

Name	Pollution Data, Provided by City of Aarhus in Denmark

Author and License	Daniel Puschmann Centre for Communication Systems Research (CCSR) University of Surrey, UK (d.puschmann@surrey.ac.uk), CC-BY Creative Commons Attribution Unported (Open) http://creativecommons.org/licenses/by/4.0/
URL	http://iot.ee.surrey.ac.uk:8080/datasets/pollution/index.html
Description	Pollution datastreams from the city of Århus from August to October 2014. This dataset includes simulation data of one sensor for each of the traffic sensor at the exact location of this traffic sensor.Pollution values are provided for carbon_monoxide, nitrogen_dioxide, sulfure_dioxide, particulate_matter and ozone index levels according to http://en.wikipedia.org/wiki/Air_Pollution_Index. For the pollution mockup stream one sensor has been simulated for each of the traffic sensor at the exact location of this traffic sensor. The data is measured using Air Quality Index[54] metric (449 observation points in total). The data is available in raw (CSV) and semantically annotated format using the citypulse information model. The stream generation works as follows: each sensor measurement (e.g. carbon dioxide) is initially assigned a value between 25 and 100. Every 5 minutes, the values will be updated as follows: if the value was below 20 before, it will now be the last value + random integer between 1 and 10 if the value was higher than 210, it will now be the last value - random integer between 1 and 10 else the value will be last value + a random integer between -5 and 5 This way the measurements do not erratically jump between low and high values and represent a more realistic stream but still won't go out of bounds (unrealistically low or high values)
Scope (Domain)	air pollution data, environmental values, sensor measurements, city pollution, location nodes, citypulse
Use cases (Motivation, Relevance)	The CityPulse webpage[55] offers a number of semantically annotated datasets collected from partners of the CityPulse EU FP7 project and relevant resources for smart city data. Visitors and potential stakeholders can use the menu on the left to access these resources.
Statistics	August 2014 - October 2014 generated data (not real measurements) 449 observation points in total
Questions	-

[54] http://en.wikipedia.org/wiki/Air_Pollution_Index

[55] http://iot.ee.surrey.ac.uk:8080/index.html

Weather Data, Provided by City of Aarhus in Denmark

Name	Weather Data, Provided by City of Aarhus in Denmark
Author and License	Daniel Puschmann Centre for Communication Systems Research (CCSR) University of Surrey, UK (d.puschmann@surrey.ac.uk), CC-BY Creative Commons Attribution Unported (Open) http://creativecommons.org/licenses/by/4.0/
URL	http://iot.ee.surrey.ac.uk:8080/datasets.html#weather
Description	A collection of datasets of weather observations from the city of Aarhus. Collected measurements from February 2014 - June 2014 and August 2014 - September 2014. Weather data values: Dew point in degrees Celsius, Humidity (percentage), Pressure in mBar, Temperature in degrees Celsius, Wind direction in degrees, Wind speed in kilometers per hour (kph)
Scope (Domain)	weather data, environmental values, Dew point, Humidity, Pressure, Temperature, Wind direction, Wind speed, location nodes
Use cases (Motivation, Relevance)	The CityPulse webpage[56] offers a number of semantically annotated datasets collected from partners of the CityPulse EU FP7 project and relevant resources for smart city data. Visitors and potential stakeholders can use the menu on the left to access these resources.
Statistics	February 2014 - June 2014 and August 2014 - September 2014
Questions	-

Energy time-series mapping from University of Southampton

Name	Energy time-series mapping from University of Southampton
Author and License	J.Barker@soton.ac.uk, Open Government Licence: http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2/
URL	http://data.southampton.ac.uk/dataset/energy-time-series-map
Description	This dataset maps buildings to their energy use time-series. Some buildings may have multiple time series, some are shared between two buildings. This data is provided by the role of Energy Manager in the BEMS team in buildings and estates.
Scope (Domain)	energy time-series map
Use cases (Motivation,	Open data portal development, university of Southampton

[56] http://iot.ee.surrey.ac.uk:8080/index.html

Relevance)	
Statistics	146 triples
Questions	

Linked geodata dataset

Name	Linked geodata dataset
Author and License	AKSW research group from Universität Leipzig. Open Database License (ODbL) : http://opendatacommons.org/licenses/odbl/1.0/
URL	http://linkedgeodata.org/ and http://linkedgeodata.org/Datasets
Description	The data set comprises all the Open Street Map data converted in RDF. It uses the lgdo ontology for describing data extracted from Open Street Map. It is accessible through REST, SPARQL end points, dumps and be navigated through a specific map layout. The data is interlinked with DBpedia and Geo Names.
Scope (Domain)	Geographic data covering the whole world
Use cases (Motivation, Relevance)	Linked Geo Data can serve as a crystallisation point for future spatial web data integration, since it provides unique URIs and exposes its content as Linked Data. Mappings to DBpedia were established already and other knowledge bases are likely to be interlinked with LGD in the future.
Statistics	LinkedGeoData consists of more than 3 billion nodes and 300 million ways and the resulting RDF data comprises approximately 20 billion triples. Unfortunately, last version seems from 2014.
Questions	

Conclusions

The aim of work package 2 was to identify the knowledge and data that can support interoperability in energy management systems by collecting and assessing relevant ontologies, vocabularies and standards, as well as relevant datasets and alignments.

The work has been carried out in cooperation with work package 3 leading to shared efforts in developing the underlying methodology and the provision of a general tool support, namely the ontology and dataset catalogue, the pitfall scanner as well as the alignment tool. Identified resources have been shared between both work packages and collected in the online catalogues. At the end of the project, a total number of 70 ontologies and 18 datasets from relevant domains have been published by following LOD principles, and alignments among them have been explored.

The developed online catalogue of ontologies and datasets is equipped with filtering features and provides a SPARQL endpoint so that users can query the RDF version of the catalogue. In addition, in order to provide a more detailed assessment (e.g., related to good modeling practices), the OWL ontologies available on the Web are evaluated by OOPS! (OntOlogy Pitfall Scanner!), an on-line application used to identify pitfalls in ontologies.

The Alignment server has filled the need for interoperability by providing an extensive network of 317 curated alignments between 42 ontologies covering the core ontologies of the domain. Such alignments may be used for transforming queries across datasets or importing some data under another ontology.

Links from the dataset catalogue to the ontology catalogue have been created and included in the web portal. In addition the ontology, dataset and alignment catalogues have been connected in the following way:

- Connection from the ontology pages to the alignment server, and vice-versa
- Connection from the dataset pages to the ontologies within the catalogue and outside.

Concerning alignments, the Alignment server will be maintained online and we plan to improve its content. This involves adding new ontologies to be aligned, exploiting other matchers if necessary and, above all, having alignment curation by specialists of the domain. This last activity will contribute to better evaluation of the alignment results. In turn, this may require technical improvements in the alignment server to support curation.

In addition to the these activities, greater effort has been directed at stakeholders and users of the project results, to further assist them in making use of the collected knowledge. In particular, WP2 was also supporting activities towards availability of more open datasets of our built environment. Main barriers have been identified and a strategy is described to increase awareness of BIM-LOD and overall willingness to publish own datasets. An important step is to standardize the ifcOWL ontology as a reference for further developments. This proposal was very well received by the BIM-LOD community and the buildingSMART organisation. Technical details are already solved by the community and it is expected that such ifcOWL standard will be available in 2016.

Another important goal is to agree on use cases that should show the benefits and the relationship to existing developments. It should identify business opportunities that can trigger further developments. One of these use cases was chosen as a show case. It was broken down into a general data publication process that is in line with the IDM/MVD methodology of buildingSMART and follows the guidelines developed in WP4. Although this use case is not directly related to the topics of energy efficiency it is believed that it can act as an important advertising vehicle for BIM-LOD.

Overall, the work carried out in work package 2 and 3 provides a solid basis for any stakeholder wishing to take advantage of linked data by providing the necessary tools in the form of a comprehensive catalogue with available ontologies and datasets. This technical basis combined with the comprehensive guidelines produced as part of work package 4 enables stakeholders to produce Linked Data and raises awareness of the opportunities it offers Smart Cities towards becoming interoperable.

References

[Beetz et al. 2005] Beetz, J., Leeuwen, J.P. van & Vries, B. de (2005). *An Ontology Web Language Notation of the Industry Foundation Classes.* In R.J. Scherer, P. Katranuschkov & S.-E Schapke (Eds.), Proceedings of the 22nd CIB W78 Conference on Information Technology in Construction, (pp. 193-198). Dresden: Technical University of Dresden.

[Beetz et al. 2009] Beetz, J., Leeuwen, J.P. van & Vries, B. de (2009). IfcOWL: *A case of transforming EXPRESS schemas into ontologies. Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 23 (Special Issue 01), 89-101

[Bizer, Heath, & Berners-Lee 2009] Bizer, C., T. Heath, & T. Berners-Lee (2009). *Linked Data – the story so far.* International Journal on Semantic Web and Information Systems 5, 1–22.

[Brickley, 2004] Brickley, D. (2004). *RDF vocabulary description language 1.0: RDF schema*. http://www. w3. org/tr/rdf-schema/.

[Euzenat and Shvaiko, 2013] J. Euzenat, P. Shvaiko, Ontology matching, 2nd ed., Springer, Heidelberg (DE), 2013

[García-Castro et al, 2014] García-Castro, R., Poveda-Villalón, M., Radulovic, F., Gómez-Pérez, A., Euzenat, J., Priego-Roche L.M., Vogt, G., Robinson, S., Birov, S., Fies, B. (2013). *Deliverable 3.1: Strategy for Energy Management System Interoperability*. READY4SmartCities project.

[Liebich et al. 2013] Liebich T., Stuhlmacher K., Weise M., Guruz R., Katranuschkov P., Scherer R. J. (2013). *Information Delivery Manual Work within HESMOS -A descriptive approach to defining Information Delivery Manuals-*. Deliverable D+ of the HESMOS project.

[Poveda-Villalón et al., 2012] M. Poveda-Villalón, M.C. Suárez-Figueroa, A. Gómez-Pérez. *Validating ontologies with OOPS!*. 18th International Conference on Knowledge Engineering and Knowledge Management. Galway, Ireland. 8-12 October 2012.