LDAC 2014

Joint Workshop on
Linked Data in Architecture and Construction
(2nd LDAC workshop & 6th eeSemantics VoCamp)

Workshop Report

http://www.linkedbuildingdata.net/LDAC2014/
Aalto University and Tekla
Espoo, Finland
May 26-27, 2014

Abstract
The topic of this event was why and how to use the Web of Data technologies for building information management. The interest to this question arose from the expected benefits such as the ability to easily link building information models to and from external data sources, to open the models to new use cases and applications, to enable truly decentralized publication of models, to support loosely coupled interoperation through cross-model linking, and to utilize reasoning and other services developed in Semantic Web research. The topic was studied from two different perspectives: The VoCamp focused on use cases and vocabularies for the Web of Building Data, and the LDAC workshop studied the technical solutions.

Workshop Chair
Seppo Törmä
Department of Computer Science and Engineering, Aalto University, Espoo, Finland

Organising Committee
Seppo Törmä
Department of Computer Science and Engineering, Aalto University, Espoo, Finland
Pieter Pauwels
Department of Architecture and Urban Planning, Ghent University, Ghent, Belgium
Jakob Beetz
Department of the Built Environment, Eindhoven University of Technology, Eindhoven, Netherlands
Matthias Weise
AEC3 Deutschland GmbH, München, Germany

Introduction to the LDAC2014 Joint Workshop & VoCamp

This report summarizes the presentations and discussions held during the 6th Vocabulary Camp (VoCamp) on "Linked Data in Architecture and Construction", which took place at Tekla Headquarters in Espoo, Finland, from May 26 to May 27, 2014. The event was organised by Aalto University, supported by the Ready4SmartCities project.

The VoCamps series is an initiative of the European Commission carried out with the support of the ADAPT4EE project. According to its original intention, a VoCamp is an informal event where experts meet to deliver lightweight vocabularies and corresponding ontologies for the Semantic Web (Web-of-Data). The goal of every VoCamp is to bring together 20 to 30 experts to discuss the state of the art about a specific topic and to come up with a proposal for new or enriched vocabularies.

This VoCamp was cohosted with the 2nd LDAC workshop, a workshop that was first organised in 2012 and that focuses on the usage and role of linked data in the context of architecture, engineering and construction (AEC). The LDAC workshop typically gathers researchers working on this specific topic, thereby aiming to bring together diverse ideas about ways in which linked data and semantic web technologies can enhance information exchange in the AEC domain. Because of this focus, the LDAC workshop matches perfectly with the purpose and context of the VoCamp series.

The workshop was motivated by the following questions:

1. *What use cases could benefit from the Web of Data technologies*, such as decentralized publication, access over the Web, ontologies, and linking across datasets?
2. *What functionalities do the Web of Data technologies provide* to practical needs of building information management? What functionalities are missing and are there major challenges?
3. *How should the Web of Data technologies be combined with BIM in detail*? For instance:
    - When the IFC schema is converted to OWL, what is the appropriate OWL profile, representation of collections, mapping of datatypes, or handling of property name clashes?
    - What information is lost in the conversion process and how can/should ifcOWL be enriched?
    - How should datasets and linksets be organized, described, hosted, accessed, and managed?
    - What ontologies besides IFC are useful or necessary for relevant use cases?

The workshop was organised as a practically oriented research meeting. In the first day the participants split into a *Use Case track* and a *Platform track*. The Use Case track was organized as a VoCamp (Vocabulary camp) that aimed at identifying a set of use case scenarios and related technical requirements. The use cases were expected to cover the publication and access to models, linking of models to external data (such as geography, infrastructure, building products/materials, and sensors), versioning, change management, knowledge acquisition, and reasoning. The Platform track studied the capabilities and options of a Web-based BIM platform, including the details of ifcOWL and the existing converters. The two tracks came together in the second day when mapping the technical requirements with the capabilities of the platform. The goal was to promote a shared understanding within participants about the application of Web of Data to building information management, and the gaps to address by future research.

The VoCamp series

The VoCamps series is an initiative of the European Commission started with the support of the ADAPT4EE project and now supported by the READY4SmartCities project. A vocabulary camp (VoCamp) is an informal event where experts get together to deliver lightweight vocabularies and corresponding ontologies for the Semantic Web (Web of Data). The goal of a VoCamp is that 20-30 experts discuss the state of the art about a specific topic and come up with a proposal for new or enriched vocabularies.

Until now, six VoCamps have taken place in the period 2012-14:

- 1st VoCamp, 15-16 November 2012. Energy Efficiency Modelling. Organised by CERTH/ITI, Thessaloniki, Greece.

- 2nd VoCamp, 14-15 February 2013. Building Information Models (BIM). Organised by the European Commission, Brussels, Belgium.

- 3rd VoCamp, 13-14 June 2013. EupP (Energy using and producing products) Management. Organised by DFKI, Kaiserslautern, Germany.

- 4th VoCamp, 13-14 February 2014. Integrating multiple domains and scales. Organised by ARC Engineering and Architecture La Salle (FUNITEC), Barcelona, Spain.

- 5th VoCamp, 20-21 May 2014. Device & Sensor Ontologies. Organised by Fraunhofer FIT, Sankt Augustin, Germany.

- 6th VoCamp, 26-27 May 2014. Linked Data in Architecture and Construction. Organised by Aalto University and Tekla, Espoo, Finland.

This document summarizes the program and objectives of the 6th VoCamp, the presentations delivered by participants and the outcomes of the discussions.

Programme
Monday 26 May 2014

10:00 Welcome to the event - Seppo Törmä, Aalto University

10.15 Welcome by the host - Ragnar Wessman, Trimble

10.30 Overview of buildingSMART position – Thomas Liebich, AEC3

11.00 Round of introductions

11.15 Presentation of VoCamp champions
Platform VoCamp champion
Peter Bonsma, RDF.org, Bulgaria

Use Case VoCamp champion
Jaap Bakker, CB-NL, Netherlands

11.45 Position statements
Building Lifecycle Performance Optimization
James O'Donnell, University College Dublin, Ireland

Building Information Modelling and Web of Data technologies
Kris McGlinn, Trinity University Dublin, Ireland

ByggNet and BIM-based processing of building permit applications
Eilif Hjelseth, Direktoratet for byggkvalitet (DIBK), Norway

Energy efficiency of buildings
Gonçal Costa, Universitat Ramon Llull, Barcelona, Spain

Use of ontologies within the IDM methodology
Matthias Weise, AEC3, Germany

Linked Data in legacy IFC
Jakob Beetz, Eindhoven University of Technology, Netherlands

12:45 Lunch

13:45 Identifying topics/focus for the tracks

14:00 Parallel workshop tracks:

1. Use cases
   Development/refinement of detailed requirements; where do Web of Data technologies fit – benefits/challenges

2. Platform
   Analysis of ontology options, required platform functionalities, and challenges of the framework

15:30 Afternoon coffee

15:45 Checkpoint for refocusing and regrouping

16:00 Parallel workshop tracks:

1. Use cases
   Development/refinement of detailed requirements; where do Web of Data technologies fit – benefits/challenges

2. Platform
   Analysis of ontology options, required platform functionalities, and challenges of the framework

17:30 Summarising session

20:00 Workshop Dinner

Tuesday 27 May 2014

9:00 Discussion and conclusions of the Platform and Use cases sessions

10:30 Invited closing presentations

scOntology catalogue
Maria Poveda-Villalon, Universidad Politécnica de Madrid, Spain

Mefisto – Management – Führung – Information – Simulation im Bauwesen
Sven Eric Schapke, Technische Universität Dresden, Germany

Future of BIM
Leif Granholm, Tekla, Finland

Integrated BIM?
Arto Kiviniemi, University of Liverpool, United Kingdom

12:00 Lunch

13:00 Identification of open issues
Continuation of work – future activities/events

14:00 Afternoon coffee

14:15 Follow-up actions and events

15:00 Closure

The need for distributed data management in construction industry

Buildings are becoming increasingly data intensive. Buildings, especially when they are still in a construction phase, typically have diverse models or representations. They often have a requirement model, an architectural model, a structural model, a MEP model, project plans, and so forth. Furthermore, a lot of external data is available that is to some extent related to the building (Fig. 1). These datasets are distributed over various places, in the hands of multiple building stakeholders.

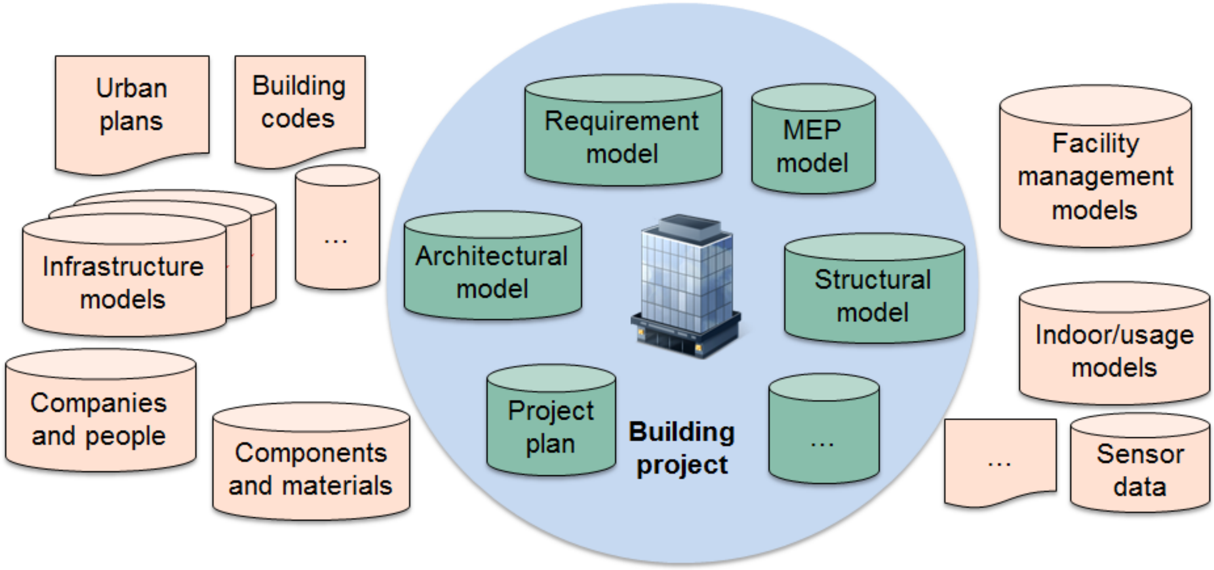

Figure Error! No sequence specified.**: Buildings are surrounded by increasing amounts of data**

From the perspective of technology, we are in the middle of a development towards Web of Data. What happened to documents in the 90's is hopefully soon to happen to data. In the beginning of Internet, *document* publication and access were considered specialist activities limited to academics or technology-oriented users. The invention of the Web in the beginning of the 90's changed everything. Document publication and access became available for everyone, and the interlinking of document provided them context and easy way to access related information. This has made the use of documents dramatically easier and created a lot of new activity and applications. The development of Web of Data technologies during the last 15 years indicates a potential to make also the *data* publication and access available to everyone and increase the value of data by linking to related datasets.

Web of Data essentially consists of the ability to identify objects with Web addresses, representation of knowledge as links between objects, to create a graph, the use of shared vocabularies to describe objects and links, and the possibility to also link data across separate, distributed datasets. *Hence, Web of Data technologies might provide useful opportunities to more easily share, manage and communicate the kinds of data that are typically dealt with in construction industry* (Fig. 1). This research question was the focal point of this LDAC workshop. Some related questions quickly arise:

- What extra functionality is available from Web of Data technologies: query functionalities, reasoning capacities, others?
- Can we use particular semantic web technologies in specific use cases?

- What ontologies are relevant besides an ontology of the Industry Foundation Classes?
- How are linked building data ontologies to be mapped onto each other? How are they to be used and organised?
- How are version histories managed? Is it a solution to work with differences and incremental updating services?

The BuildingSMART perspective: Lessons learnt and adoptable strategies

There is an interest from BuildingSMART International in investigating the possibilities of relying on semantic web technologies for the international push towards effective and efficient information management and communication in the construction industry. The idea of linked data in architecture and construction, in conjunction with the existing BuildingSMART technologies (IFC, IFD, bsDD, IDM, MVD, and so forth), has been discussed in the BuildingSMART International (bSI) meeting in Stockholm from 17 to 20 March 2014. The topic will be on the agenda again during the upcoming bSI meeting in Toronto, Canada.

Nevertheless, there are some reservations made towards the abilities of semantic web technologies. First of all, they resemble to earlier technologies and to the aims that have often been pursued using earlier technologies. For instance, semantic web technologies resemble the use of data models, frames, the aim of the BuildingSMART data dictionary (bsDD), knowledge representation, and so forth. Therefore, although there appears to be a promising appeal towards the usage of semantic web technologies, these technologies need to be put in context as well, so that lessons learnt do not need to be learnt again.

Some differences between traditional and semantic modelling can be outlined:

- Closed World Assumption (CWA) <u>versus</u> Open World Assumption
- Models are checkable for completeness and consistency because of the CWA (unknown == false) <u>versus</u> the model is always incomplete and often inconsistent (unknown data != false)
- File-based data and processes <u>versus</u> data processes distributed over the web
- File-based modelling and access to files via web services <u>versus</u> interactive information search and reasoning over distributed datasets
- Focus on *what to exchange* <u>versus</u> focus on *how to link*
- Sharing complete datasets <u>versus</u> sharing incomplete datasets

The impact of these differences needs to be assessed to find the role that semantic web technologies can play in improving data management and communication in the construction industry. As for the role of BuildingSMART, it is argued that this domain should preferably be a research-driven domain instead of a centrally managed domain. Especially considering that construction industry is still struggling to implement to Building Information Modelling (BIM) approach and the standardisation efforts provided by BuildingSMART, we should be careful not to push the usage of semantic web technologies too early and too drastically. Therefore, it is questioned to what extent a 'standard ifcOWL ontology' or a 'standard bsDD ontology' is desirable or needed. Perhaps we should first focus on trying to combine the best of two worlds (traditional modelling world + semantic modelling world) in practical and down-to-earth use cases.

Diverse perspectives on Linked Data in Architecture and Construction

> *Peter Bonsma – geometry in RDF*

By porting an IFC model of a building into an RDF data model, the knowledge in the IFC model are transferred into parametric, geometric components. For instance, a geometric instance of a T-profile in IFC can be transformed into an RDF concept with a unique resource identifier (URI), after which geometric parameters can be assigned with RDF properties. The resulting T-profile concept in the RDF data model can then be reused in various contexts. This technique relies on OWL 2 (DL) metamodelling or punning. Using this technique, multiple inheritance can be implemented on classes and associated scripts.

Additionally, an RDF representation allows to keep semantics separate from geometry. So, it is possible to represent a concept (e.g. a toy car in Lego-blocks) using an entire block of ontology concepts, and prepare the corresponding geometry relying on this block of ontology concepts.

In the PROFICIENT project, a roundtrip procedure was implemented from IFC into RDF and back into IFC.

> *Jaap Bakker – the central ontology*

In the Dutch CB-NL project, a concept library is targeted of generic reusable objects covering the complete building lifecycle. This concept library is aimed at combining a great number of distinct standards, databases, and ontologies, including STABU, NET, NLCS, BID CROW, ImGeo, AQUO, INSPIRE, CORA RioNED, ETIM, and so forth (Fig. 2).

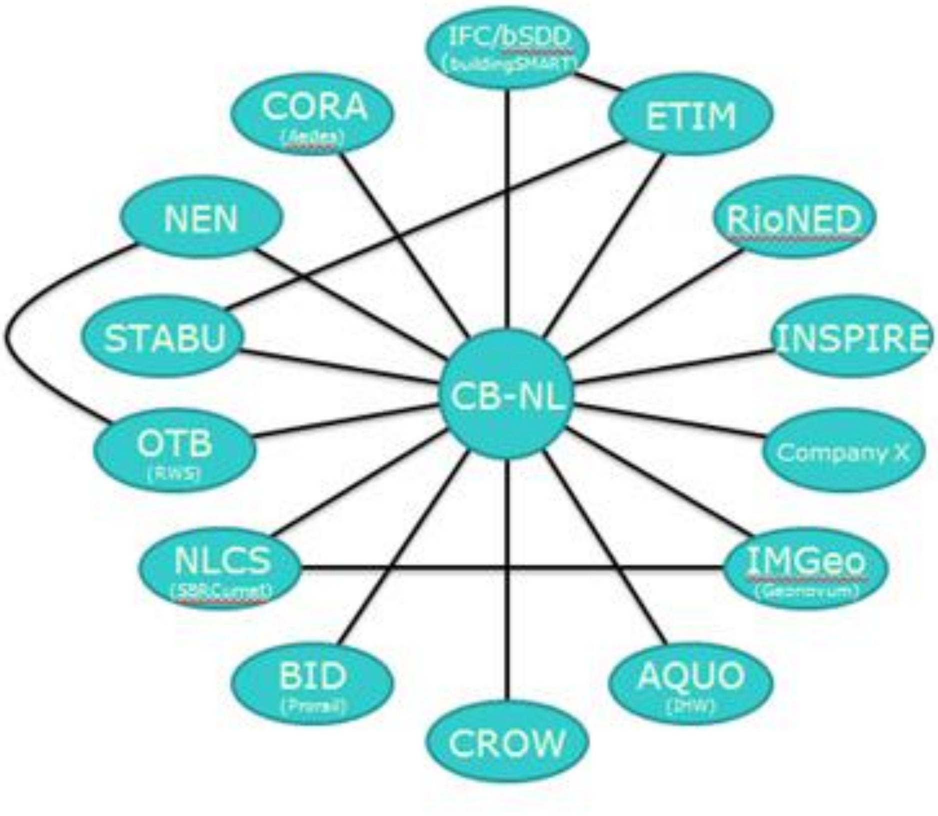

Figure 2. The CB-NL concept library acts as a central hub for all data models available in the building lifecycle.

Relying on the RDF links between the central concept library and the diverse context ontologies, data can be exchanged. The RDF data model allows providing such links. Relying on this strategy, one does not aim for one model of the world, but instead, one aims at combining contexts on the fly as is fit in particular use cases. The CB-NL in itself will become a dedicated context in the BuildingSMART Data Dictionary (bsDD).

> James O'Donnell – performance analysis

Semantic web technologies have also been tested in the implementation of a performance analysis framework aimed at a more holistic building performance operation and thus at building lifecycle performance optimization. In this implementation, diverse data sources can be directly combined with each other, providing a networked representation of the building (Fig. 3).

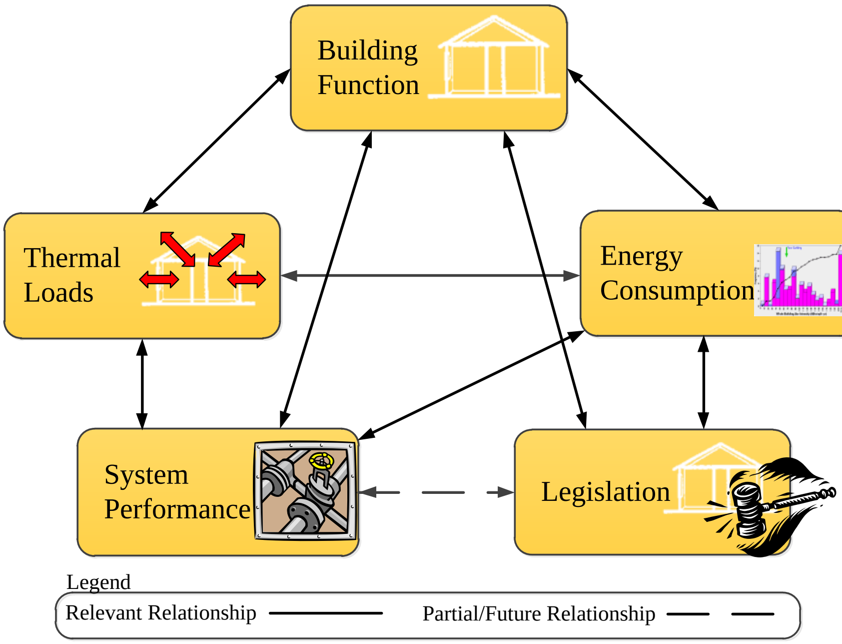

Figure 3: A holistic environmental and energy management method for building operation optimisation.

In choosing which data sources and/or concepts to link to each other, a scenario modelling method is relied upon. The scenario model describes the context of the project and how the available data should be interlinked in support of this context (Fig. 4).

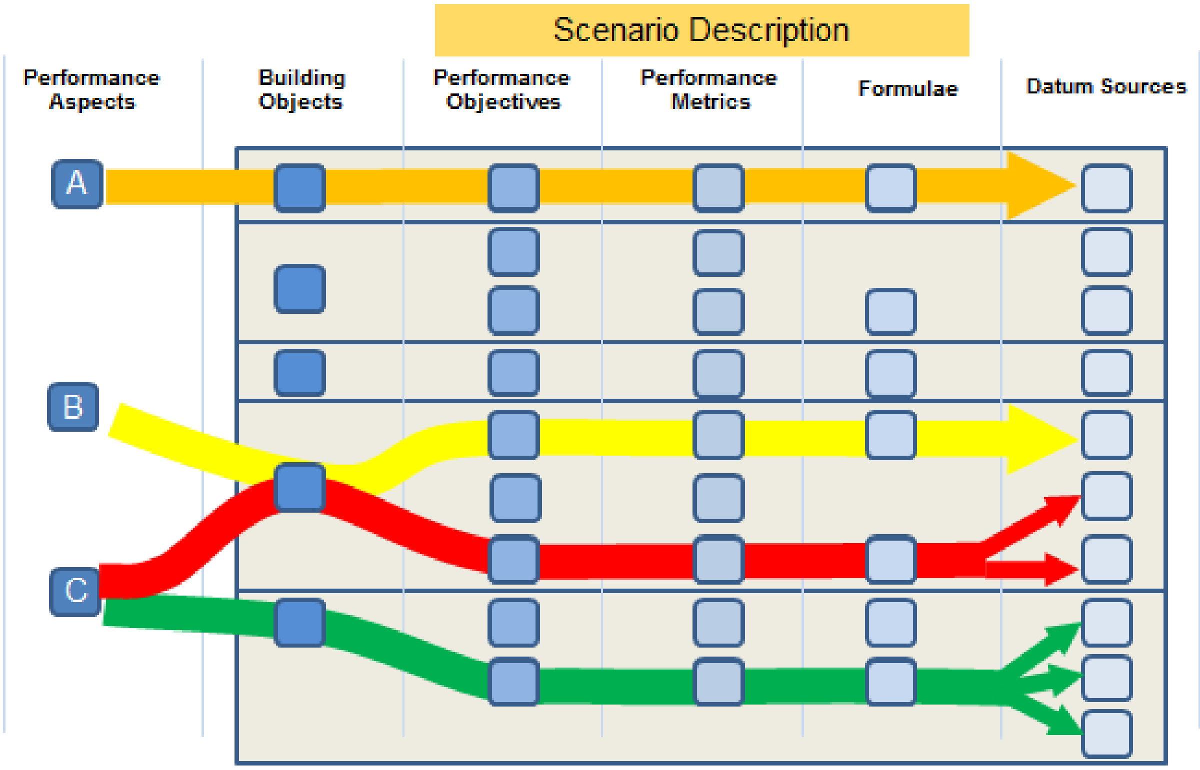

Figure 4: Diverse scenario descriptions can be modelled, each relying on a different set of data sources.

> Kris McGlinn – BIM as a part of the LOD cloud

It is possible to realise a separate hub in the linked open data (LOD) cloud including available Building Information Models (BIMs). By linking these BIM models in an RDF representation to other data in the LOD cloud (e.g. historical data, building occupant data, activity data, manager data, semantic sensor data, geographical data, device data, and so forth), additional services can be implemented that were hard to realise beforehand. Example use cases are:

- Knowledge Based Holistic Energy Management
- SWRL reasoning on historic building sensor data
- occupant activity modelling
- visualisation of rules for the end user (building manager)

For use cases relying on sensor data, it is recommended to keep the sensor data stored in a relational database (SQL) because of the amounts of data that are typically generated in such contexts. If such data are all ported into an RDF format, the performance of using these data with semantic web technologies goes down significantly. In that context, it makes sense to keep the sensor data in their original format, and copy an extract or summarising model from the data into the RDF data model for interpretation using semantic web technologies.

A number of advantages advocating the use of Web of Data technologies can be identified:

- Accessibility of data:
  Multiple data sources can be linked to provide new capabilities for data analysis for solutions; these can be deployed across platforms (using web browsers)
- Scalability of solutions
  Solutions are deployable across platforms in diverse interfaces, as all data is available in online SPARQL endpoints

A number of disadvantages advocating against the use of Web of Data technologies can be identified as well:

- User-friendliness of SPARQL
- Reliability/Accuracy of data generated
  Data sources are easily available, but the provenance of the data is often not included, leading to questions regarding reliability and accuracy of the data
- Security of data
  Who has access to which data? Are these security data stored as part of the data, or is that managed externally?
- Privacy of data
  Can data about buildings be published without interfering with privacy issues?

Assuming that workarounds exist for the disadvantages regarding reliability, accuracy, security and privacy of data, the main question remains: *how can people be convinced to use Web of Data technologies (cfr. user-friendliness)?* The sole motivators in this regard are *convincing and beneficial real world applications in practical use cases*. Producing such real world applications can be the topic and main aim of a W3C community group.

> Eilif Hjelseth – regulation compliance checking
Web of Data technologies can possibly provide additional functionalities and features to the BIM-based processing of building permit applications. The Norwegian ByggNet initiative is testing the possibilities in this direction. In this initiative, the idea is to rely on the buildingSmart Data Dictionary (bsDD), formerly known as IFD, for the automation of processing building permit applications. In this context, a set of 'shared national components' is built, including cadastre data, geographical data, and so forth, and a compliant set of 'national building regulation vocabularies' is built. These are then to be combined with an IFC ontology, so that it is possible to conclude to what extent an IFC model of a building is compliant with a defined set of national regulations in Norway.

A simple use case that is implemented provides an ontology of the Norwegian fire regulations (TEK10). These fire regulations rely on a 'building hazard class'. Although the regulation text gives a textual and tabular view that allows one to derive the hazard class of a building, a simpler decision tree can also be set up within an RDF data model, thereby allowing to derive that same hazard class of the building more rapidly and more correctly.

> Gonçal Costa – energy efficiency in smart cities
Another domain of application that could be of use as a key domain in which semantic web technologies can prove their value, is the evaluation of energy efficiency of buildings, building blocks and entire cities. A demonstrator case was set up in a publicly subsidised apartment building in Cerdanyola del Vallès, Barcelona. In this building, a number of sensors was available, allowing for a building performance analysis in terms of energy efficiency. For this demonstrator use case, a combination was made of linked geo data, ICAEN, and AEMET ontologies. Using these data sources, intuitive interfaces can be developed for the exploration of energy efficiency in a building. This test case was further repeated in case studies for Newcastle (UK), Copenhagen (Denmark) and Manresa (Spain - Fig. 5).

Figure 5: Screenshot of the SEMANCO integrated platform (http://www.semanco-project.eu/).

> Matthias Weise – Integrated Delivery Manuals

Web of Data technologies could well be relied upon to drive developments in the concept and usage of Integrated Delivery Manuals (IDMs). IDM's are typically interpreted as business cases that are 'translated' to IFC. The procedure of implementing an IDM typically consists of the following steps:

- Preliminary work
- Discover processes and actors
- Identify exchanges
- Create exchange requirements
- Extend to entity-relationship models
- Unify to model view definition
- Implementation and use

In its current situation, the communication between people is weakly captured and formalised and the communication between machines happens via an IFC model. By relying on ontologies and Web of Data technologies, domain knowledge can hopefully be captured and formalised more easily, so that the communication between people can be modelled and realised more easily. In terms of communication between machines, IFC keeps serving as a pivot ontology.

The main issues and questions in this context are:

- how to formally capture domain requirements?
- how to assess the quality of domain requirements (reasoning)?
- how to link domain requirements to IT solutions (ifcOWL pivot ontology)?
- What to add to IFC EXPRESS-derived ifcOWL ontologies to enable knowledge instead of data sharing?
- Do we need one standardised ifcOWL representation?

> Jakob Beetz – linking data in legacy IFC models

Web of Data technologies allow us to build distributed, collaborative concept libraries for the construction industry in a relatively straight-forward and easy-to-use way. In implementing Web of Data technologies, construction industry experts should consider the low-hanging fruits first and combine the best of the traditional world with the best of the semantic modelling world in the Web of Data context.

The legacy IFC format is a very good basis, and it makes therefore no sense to start this effort anew and reproduce a central, more or less agreed-upon ontology. What is harder in an IFC context, is to add extra information, which is precisely what Web of Data technologies are able to offer. So, instead of making separate specialised vocabularies and suggest them for inclusion in the EXPRESS scheme for IFC, it is possible to develop these specialised vocabularies, link them as RDF data triples and use them as is considered useful (Fig. 6). By doing so, the construction expert is not only allowed to open a building model in existing applications using the legacy IFC format, he is also allowed to extend the model with additional data using RDF links. In this context, it would be worthwhile to develop a core ifcOWL ontology as an intermediate model to link from.

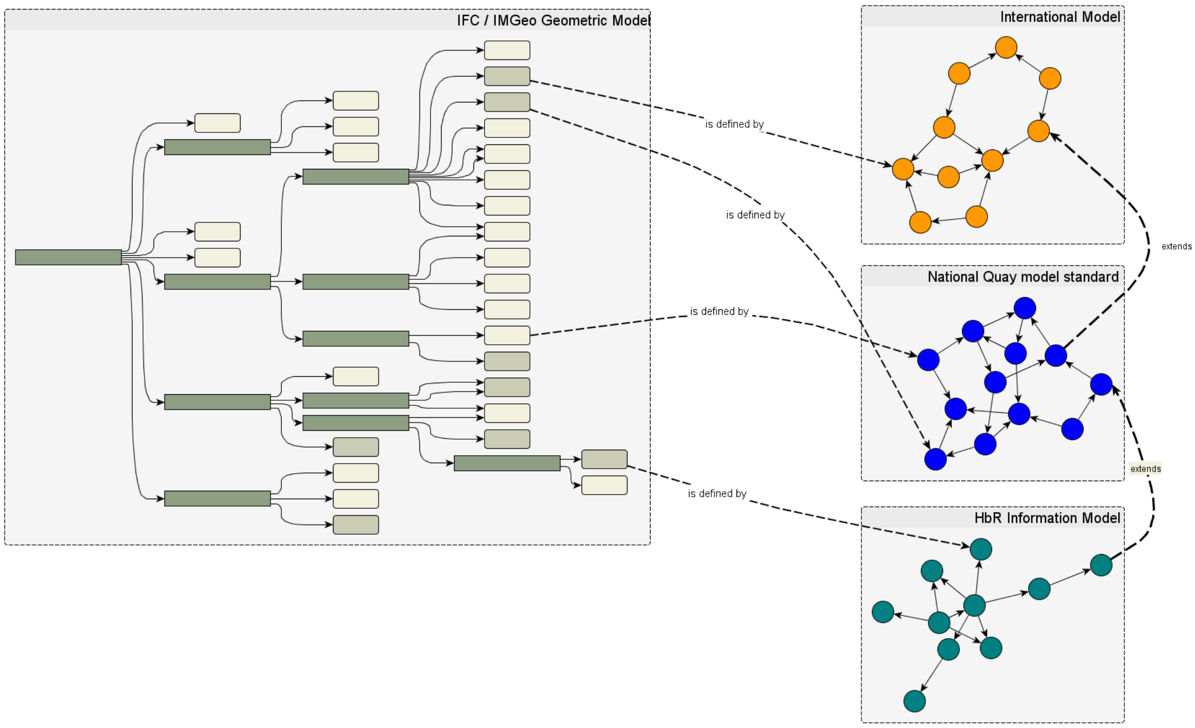

Figure 6: A core IFC model (left) can easily and in an ad hoc way be linked to concepts residing in other domains (right), which can be modelled as the end user wishes.

The Platform session

> Is there a need to standardise IFC to ifcOWL?

The IFC schema has been converted into an ifcOWL version within the DRUM project at Aalto University (http://cse.aalto.fi/en/research/groups/distributed_systems/projects/drum/). At the moment, this converter receives two different inputs. Either a conversion from EXPRESS schema (IFC2x3 or IFC4) to a corresponding OWL ontology is performed, or a conversion from an IFC physical file (SPF) into an RDF graph is performed. At the moment, the following output profiles can be followed in the conversion process of both (EXPRESS to OWL / IFC SPF to RDF):

- OWL Lite / DL
- OWL 2 RL
- OWL 2 EL
- OWL 2 QL
- OWL Full

Every one of these profiles has its advantages and its disadvantages. It would be useful if some sort of standardisation or agreement is found, so that there is at least a recommended version of an ifcOWL. Apart from this recommended ifcOWL version, the original converter application can still be provided so that alternative conversion profiles can be used in exceptional cases.

> What is the reason to provide an OWL Full ontology of the IFC schema?

OWL Full is the best conversion option for the formalisation of linked lists: lists that include an ordered collection of links to other concepts.

> How to deal with the great number of lists in the IFC schema?

The intensive usage of collections in the IFC schema in EXPRESS is one of the main concerns when considering a conversion to an OWL ontology and thus also when considering to rely on one particular ontology (OWL) profile. Several collection entities are represented in the EXPRESS schema with cardinality restrictions. For instance, the 'IfcCartesianPoint' entity has the following property:

IfcCartesianPoint Coordinates LIST [1:3] OF IfcLengthMeasure;

Depending on the OWL profile that is chosen, certain cardinality restriction might not be supported. A generic way to implement ordered collections is to rely on the 'olo:OrderedList' concept that resides in the Ordered List Ontology (http://smiy.sourceforge.net/olo/spec/-orderedlistontology.html).

> How to deal with property name clashes

Many of the properties in the IFC schema have more than one domain concept (rdfs:domain) listed. A straightforward conversion of properties therefore typically results in name clashes (two different properties with the same name). It is a possible solution to rely on 'rdfs:subPropertyOf' and convert properties hierarchically. An example property is 'overallHeight', which is available for both 'IfcWindow' and 'IfcDoor'.

> Which elements in the IFC schema cannot / are not converted?

The EXPRESS-specific elements UNIQUE, WHERE, DERIVE, and FUNCTION are currently ignored and are thus not converted in a corresponding form in the resulting ifcOWL version.

> ***Which are the requirements for standardizing ifcOWL?***

- The ifcOWL ontology should contain the (important) information of the original IFC schemas/models. A minimal list:
  - all entities
  - all properties (naming conflicts, range issue: => subPropertyOf:
  - inheritance
- The ifcOWL ontology can be aligned with non-IFC ontologies
- The ifcOWL ontology can be deployed and shared in the Internet
- The ifcOWL ontology can be processed by inference tools with acceptable performance
- (optional) The ifcOWL ontology can be converted back to SPF IFC
- The conversion procedure and its implementation (EXPRESS-to-OWL and IFC-to-RDF) should be capable of converting newer IFC versions

> ***Which are the requirements for a standard ifcOWL schema?***

- The ifcOWL ontology follows the OWL 2 RL profile (http://www.w3.org/TR/owl2-profiles/#OWL_2_RL):
  - Pro: high expressiveness, allows covering most contexts
  - Con: limited support by reasoning engines
- all entities of the original IFC EXPRESS schema are included
- all properties of the original IFC EXPRESS schema are included
  - naming conflicts and range issues are resolved with the ‘rdfs:subPropertyOf’ construct, as explained above.
- cardinalities of the original IFC EXPRESS schema are NOT included
- (undecided) all inverse properties of the original IFC EXPRESS schema are included
  - Pro: much more comfortable way to navigate the graph
  - Con: including inverse properties is redundant as the original properties are available in the canonical instance. The inverse properties can be inferred easily.
- Inheritance is included as in the original IFC EXPRESS schema
- (undecided) the domains and ranges of datatype properties are NOT included in order to make the resulting ontology less rigid/strict and in order to resolve naming conflicts and range issues (see bullet point 3 above).

This ifcOWL schema is made available for publishing. If necessary, an alternative configuration or OWL profile can be provided using the conversion available IFC-to-RDF software.

> ***What is the intended use of the ifcOWL schema?***

- linking datasets
- querying data
- publishing data

- reasoning with data

The Use Case session

During the use case session, a number of use cases was identified and explored. Through these use cases, the benefit of semantic technologies over traditional technologies should be identified. In order to do so, a formalisation of each use case has to come in place. For each use case, the following elements should be formalised / represented, so that an assessment can be made of the content and value of the use case.

- user group
- problem
- aim
- method
- datasets used
- expected outcome

The ideal environment to produce a list of key use cases and a corresponding formalisation of each use case, is a W3C group. This W3C group can be set up as a merger of two earlier initiatives, namely of the open linkedbuildingdata.net web portal and the ontology group for building data (OG4BD) mailing list.

> Communication of information during the building life cycle
There appears to be a central concern for all in the use case session, which is expected to be addressed by Web of Data technologies, namely, the exchange and communication of information in general during the whole building life cycle among diverse disciplines. It is expected that Web of Data technologies will allow a more flexible, dynamic, easy and efficient exchange of information. It is less clear what strategy should be followed to obtain maximal profit of Web of Data technologies.

- Centralised model
  This approach is currently tested by the CB-NL initiative (see Fig. 2), in which a central concept library is developed as an enabler to link diverse data sets to each other. This idea is to some extent compliant with the buildingSMART idea of a buildingSMART Data Dictionary (bsDD) and an Integrated Delivery Manual (IDM).

- Temporarily merged model
  This approach appears to be advocated by a number of research initiatives, such as the SEMANCO project, the DRUM project and the Performance Framework Tool (PFT) project. Additionally, it complies well with the buildingSMART idea of Model View Definitions (MVDs). For a specific moment in the building life cycle and a particular aim and targeted user group, a number of datasets are temporarily merged, so that the aim can be reached for the user group, and the datasets can be 'released' again afterwards.

- Decentralised, interlinked models
  This approach may be understood as a space of separate but interlinked datasets to support cross-dataset information access, aggregation of information from multiple datasets, and propagation of information across datasets. This is a completely user-

driven or industry-driven approach in which data is published in various locations and everyone simply uses whatever he or she needs in the way he or she sees fit.

It appears that the first approach (centrally shared vocabulary) resembles the effort of standardisation that has already been made with the Industry Foundation Classes (IFC), and even more with the buildingSMART Data Dictionary and the Integrated Delivery Manual (IDM) ideas. It is not entirely clear what benefits Web of Data technologies can provide to these earlier initiatives, since these Web of Data technologies seem to focus on loose links, dynamic dataset merging, and data flexibility.

Although the last approach is compliant with the way in which the Web of Data currently appears to work, but it might be less appropriate for a profit-driven industry like the construction industry, because there is little to no incentive to share data freely for anyone to use with no guarantee for a return on investment.

The second approach, in which datasets are temporarily merged for a specific user group and a specific aim, appears as the most promising and demanded approach among the workshop participants. Many of the following use cases are in line with what this information communication approach can offer. We briefly document the use cases with an exploratory formalisation according to the above formalisation method (user group, problem, aim, method, expected outcome).

> Use Case: Link a MEP piping system with a penetration in a structural wall

- user group: MEP engineers, structural engineers
- problem: MEP piping data is not available to structural engineers and structural data is not available to MEP engineers.
- aim: link instances in a MEP piping model with instances in a structural model
- method: direct links between two datasets
- datasets used: MEP instance model, structural instance model
- expected outcome: temporarily merged MEP-structural instance model

> Use Case: Link data in a scheduling system with BIM data during the building operation phase

- user group: facility managers, building owners
- problem: the facility manager and building owner has no idea how the building is used at diverse periods in time, so he cannot optimise the building to accommodate those building uses. For example, HVAC systems may be configured in direct relation to building use.
- aim: provide building use schedule data to FM for an improved building use
- method: direct links between two datasets
- datasets used: schedule data (often spreadsheets – CSV), BIM model, FM model, sensor data
- expected outcome: temporarily merged schedule-BIM instance model

> Use Case: Find and link manufacturer data

- user group: architects, designers, construction experts

- problem: which components can I use to build a certain building
- aim: find manufacturer data regarding available components, and analyse which components should best be used in the current design
- method: query available manufacturer data based on a BIM model
- datasets used: BIM model, manufacturer data
- expected outcome: more detailed BIM model

> Use Case: Find recommendations to realise building components more efficiently or in a cheaper fashion (supply chain management)

- user group: construction experts
- problem: which alternative building solutions are there for the component that I currently have
- aim: find manufacturer data regarding available components, and analyse which components should best be used as alternatives to the currently available components
- method: query available manufacturer data based on a BIM model
- datasets used: BIM model, manufacturer data
- expected outcome: changed and more detailed BIM design model

> Use Case: Link requirement specifications with performance specifications, so that the appropriate building solutions can be found more rapidly

- user group: architects, designers, construction experts
- problem: which components can I use to build a building with a particular performance level
- aim: find manufacturer data regarding available components, based on a BIM model, regulation data and performance data
- method: temporarily merged model of BIM model, regulation data and performance data + query available manufacturer data based on this temporarily merged model
- datasets used: BIM model, manufacturer data, regulation data, building performance data, sensor data
- expected outcome: changed and more detailed BIM design model

> Use Case: Link building data with building regulation data, so that a check can be made to what extent the building is compliant with applicable regulations

- user group: architects, designers, construction experts, government bodies
- problem: check compliance of building with building regulations
- aim: link building model with regulation data and perform compliance check
- method: temporarily merged model + reasoning
- datasets used: BIM model, regulation data
- expected outcome: compliance check for BIM model

From these use cases, a number of questions emerged that will need to be dealt with:

- How can datasets be linked to legacy data sources, including relational databases (SQL), images, sensor data, scans, etc.?
- What should links among datasets look like, so that the links are manageable in terms of version control?

- Measures are needed with relation to data privacy, data security and data ownership (provenance). Which data is specific to a project, which data is specific to a company and which data is part of common knowledge? Are dereferenceable URI's (namespaces) powerful enough to deal with this issue?
- In order to make 'temporarily merged models' possible, it should be made as easy as possible to find the datasets and ontologies that can be used for specific use cases. For this purpose, an open repository or portal can be set up, so that available resources can be found easily. This could be the linkedbuildingdata.net portal or the SmartCities Catalogue.

Looking beyond the horizon

> Merge between platform session and use case session

The preliminary list of use cases appears to call for a process-oriented usage of semantic web technologies. There are a considerable number of use cases that can be identified throughout the entire building lifecycle (design, construction, operation, demolition), with each phase involving very diverse disciplines and corresponding datasets. Web of data technologies are expected to allow quickly and intuitively assembling of temporarily merged models that support portions of the building life cycle phases in a dynamic and flexible fashion.

This use case background calls for a flexible and easy-to-use version of an IFC ontology in OWL. This ifcOWL then serves as a 'default' ontology in support of dynamic and flexible adoption of web of data technologies. This default ifcOWL can be provided in the form of an OWL 2 RL ontology profile of the original IFC schema. Requirement for this ifcOWL version are, as presented above:

- The ifcOWL ontology follows the OWL 2 RL profile (http://www.w3.org/TR/owl2-profiles/#OWL_2_RL):
    - Pro: high expressiveness, allows covering most contexts
    - Con: limited support by reasoning engines
- all entities of the original IFC EXPRESS schema are included
- all properties of the original IFC EXPRESS schema are included
    - naming conflicts and range issues are resolved with the 'rdfs:subPropertyOf' construct, as explained above.
- cardinalities of the original IFC EXPRESS schema are NOT included
- (undecided) all inverse properties of the original IFC EXPRESS schema are included
    - Pro: much more comfortable way to navigate the graph
    - Con: including inverse properties is redundant as the original properties are available in the canonical instance. The inverse properties can be inferred easily.
- Inheritance is included as in the original IFC EXPRESS schema
- (undecided) the domains and ranges of datatype properties are NOT included in order to make the resulting ontology less rigid/strict and in order to resolve naming conflicts and range issues (see bullet point 3 above).

Other 'flavours' of ifcOWL are certainly possible, and even supported through an open and recommended converter application. Certain flavours can be required in exceptional and high-demanding cases (reasoning purposes, big data contexts, etc.).

> Towards a standard ifcOWL ontology

In support of the platform track developments, a good test bed is required in which BIM models are publicly available in the IFC format, so that they can be used for testing the conversion into corresponding RDF graphs. This test bed can be provided in three locations.

- The demo IFC-to-RDF server at Ghent University
  http://demo.mmlab.be/IFC-repo/
- The IFC model repository at Auckland University
  http://openifcmodel.cs.auckland.ac.nz/
- The BIM server software
  http://www.bimserver.org

The development of the standard ifcOWL ontology preferably is in hands of / managed by a standardisation body, which could be either the World Wide Web Consortium (W3C) or buildingSMART.

> *W3C community group Linked Building Data*

In support of the use case track, it would be beneficial if a linked building data portal or central hub is created that provides access to recommended ontologies and data sets for building data use cases. This portal obviously relies on the standardisation efforts that are made regarding the ifcOWL ontology. Therefore, it is suggested to set up a separate W3C community group that joins the previously started initiatives 'linkedbuildingdata' (http://linkedbuildingdata.net/) and 'Ontology Group 4 Building Data (OG4BD)' (http://listserver.tue.nl/mailman/private/og4bd/). If this W3C group is set up, it could be connected to the W3C Product Modelling Incubator Group, which can be found at http://www.w3.org/2005/Incubator/w3pm/XGR-w3pm-20091008/.

The specific aim of this group is to develop, formalise and publish use cases around ifcOWL. The group will start by investigating the questions and remarks that came up during the use case session:

- How can datasets be linked to legacy data sources, including relational databases (SQL), images, sensor data, scans, etc.?
- What should links among datasets look like, so that the links are manageable in terms of version control?
- Measures are needed with relation to data privacy, data security and data ownership (provenance). Which data is specific to a project, which data is specific to a company and which data is part of common knowledge? Are dereferenceable URI's (namespaces) powerful enough to deal with this issue?
- In order to make 'temporarily merged models' possible, it should be made as easy as possible to find the datasets and ontologies that can be used for specific use cases. For this purpose, an open repository or portal can be set up, so that available resources can be found easily. This could be the linkedbuildingdata.net portal or the SmartCities Catalogue.

By setting up this W3C community group, the linked building data community becomes more easily addressable by fields related to, but not part of the building construction industry (cfr. GIS community, product data community, and so forth).

With each use case, a number of recommended ontologies or datasets is given, so that experts in construction industry can rather easily put the use case in practice. This effort can rely on work that is already done regarding the collection and management of ontology and dataset references. Examples are the SmartCities Catalogue (http://smartcity.linkeddata.es/), the DuraARK project (http://duraark.eu/), and the linkedbuildingdata web portal (http://linkedbuildingdata.net/).

The SmartCities catalogue hereby serves as a reference example:

- Goals of the SmartCities catalogue:
    - facilitate ontology reuse
    - one place to discover ontologies
        - multiple registers and search engines exist, but they are typically so generic that a construction expert has no desire to look into the details => need for a registry with building-related ontologies
    - specific for energy efficiency and smart cities
    - ontology assessment

- Population method / roadmap of the SmartCities catalogue:
    - http://goo.gl/z4TlFO: form available to submit an ontology URI, so that catalogue managers can populate the catalogue with this ontology.
    - http://goo.gl/SG0pMA: form available to submit ontology metadata, so that catalogue managers can populate the catalogue with this ontology
    - http://smartcity.linkeddata.es/: the resulting repository
        - including quality indicators (online availability, license information, ontology language)
        - including metadata (syntax, domain, natural language)

- Quality assessment and management of the SmartCities catalogue
    - manual assessment by catalogue managers (Politechnical University of Madrid)
    - automatic assessment using the OntOlogy Pitfall Scanner (OOPS - http://www.oeg-upm.net/oops/)

> *Follow-up actions and events*

- A follow-up "Building Knowledge" workshop will take place during the i-Know workshop in Graz (Austria) a day before ECPPM http://duraark.eu/events/i-know-2014-workshop-linked-building-data/
- A special issue on Linked Data in Architecture and Construction will appear in Automation in Construction. The call for papers will be distributed soon.
- A task force may be set up within buildingSMART for the promotion of a 'canonical' ifcOWL for Linked Data and other profiles.
- A W3C community group will be set up for the identification and formalisation of use cases and corresponding recommended datasets and recommended application strategies.
- A follow-up LDAC workshop will likely be set up in Eindhoven Technical University back-to-back with CIB W78 sometime in November 2015.

Acknowledgements
A sincere thank you goes to all participants of the LDAC2014 Joint Workshop on Linked Data in Architecture and Construction (2nd LDAC workshop & 6th eeSemantics VoCamp):

Anna Osello, Politecnico di Torino, Italy
Arto Kiviniemi, University of Liverpool, United Kingdom
Bruno Fiès, CSTB, France
Dag Fjeld Edvardsen, Catenda, Norway
Eilif Hjelseth, Norwegian Building Authority, Norway
Gonçal Costa, Universitat Ramon Llull, Spain
Jakob Beetz, Eindhoven University of Technology, Netherlands
James O'Donnell, University College Dublin, Ireland
Jyrki Oraskari, Aalto University, Finland
Kary Främling, Aalto University, Finland
Kris McGlinn, Trinity College Dublin, Ireland
Leif Granholm, Tekla, Finland
María Poveda-Villalón, Universidad Politécnica de Madrid, Spain
Markku Kiviniemi, VTT, Finland
Mathias Kadolsky, Technical University Dresden, Germany
Matthias Weise, AEC3, Germany
Mervi Himanen, Digital Living, Finland
Nam Vu Hoang, Aalto University, Finland
Niels Treldal, Technical University of Denmark, Denmark
Peter Bonsma, RDF, Bulgaria
Petteri Villa, Tekla, Finland
Pieter Pauwels, Ghent University, Belgium
Ragnar Wessman, Tekla, Finland
Raimar Scherer, Technical University Dresden, Germany
Raju Pathmeswaran, Birmingham City University, United Kingdom
Reijo Hänninen, Granlund, Finland
Seppo Törmä, Aalto University, Finland
Sven-Eric Schapke, think project, Germany
Thomas Liebich, AEC3, Germany
Tuomas Laine, Granlund, Finland
Vishal Singh, Aalto University, Finland

We would also like to thank Tekla for hosting this event in their headquarters in Espoo, Finland, and the European Union for supporting this event through the Ready For Smart Cities (R4SC) project.