

# Speed Dating for Building Data Vocabularies

*Full-day workshop proposal i-KNOW 2014*

*Applicants (tentative): Jakob Beetz (TU Eindhoven, NL), Matthias Weise (AEC3; DE), Pieter Pauwels (Ghent University, BE), Seppo Törmä (Aalto University, FI), Stefan Dietze (LUH, DE)*

## Introduction (skip if familiar with AEC/FM | B&C)

Over the course of several decennia, interoperability and standardization efforts in the field of Architecture, Engineering and Construction (AEC) has led to large panorama of expressive, well-documented and vendor-neutral data models, classification systems, taxonomies and other structured vocabularies. Yet – partially due to the lack of distribution and networking facilities inherent to the technological basis of the modeling means – many of these vocabularies remain in isolated silos and barely reuse established solutions from other domains, such as ontology matching, data mining, automatic ontology integration, generic data visualization, reasoning using dedicated inference engines, and so forth.

The growing need of information integration across domains of the built environment (geo | civil engineering | city planning | buildings | building components etc.) has led to a recent spike of initiatives, (EU and national-funded) research projects and standardization efforts that are using semantic web technologies to represent and transform these historically grown domain vocabularies. With an integrated view on the information of our built environment, one is able to provide end users more easily with an overall understanding of his environment, allowing him to interact with it and interactively change it into a responsible smart city and smart building environment.

**Examples of such models | vocabularies | classification and projects (many available in RDF):**

IFC, bSDD, INSPIRE, OGC Geo models, SfB(-NL), DIN 276(?), Omniclass, Uniclass, Getty AAT, BauDataWeb, CROW, OKSTRA, LandXML, CASCADE airport ontology, SIMmodel, GIS ontologies, CityGML, ETIM, gbXML etc.

## Workshop aims and scope

We propose an interactive workshop that aims at the creation of a common registry and mapping of various knowledge models, structured vocabularies, and interoperability standards from the fields of building and construction in architecture and civil engineering. Instead of aiming at a rigid axiomatic high-level pivot- or upper- ontology to which everything is mapped (à la SUO/SUMO), a light-weight approach using vocabularies with ‘softer’ semantics such as SKOS & DC will be pursued. This allows the rapid, ad hoc creation of many relations among concepts without risking to invalidate intricate description logic models while staying faithful to the LOD spirit that cherishes the emergence of synergetic effects by means of machine-readable interrelations.

## Workshop implementation approach

As a common platform/sandbox tool, the workshop will provide a quadstore database with sufficient resources (RAM/disk) to which the workshop participants will be granted access (e.g. via SPARQL endpoints). This will be initially populated with a number of different vocabulary dumps from various disparate sub-domains in building and civil engineering (architecture, building materials, roads, tunnels, bridges, pipes etc.) as well as modeling and standardization initiatives on national and international levels. Each workshop participant will be asked to take responsibility for one vocabulary she/he is familiar with or has gotten acquainted with in preparation for the workshop. Using software tools such as editors, profilers, mappers etc. freely to be chosen by the workshop participants (or suggested by the organizers) relations among vocabularies will be created during the workshop. Next to these 'traditional' digital tools (which could also be effectively used in the solitary confinement of isolated laboratory settings) we strongly encourage the use of analogue but highly interactive tools such as whiteboards, flipcharts and post-it notes to leverage the advantages of co-located synchronous collaboration settings. The results of these efforts will be immediately concurrently stored (or populated in a post-processing session) in separate yet open contexts | named graphs | reification clusters which will reside on the same machine. As a result of the workshop, the individual vocabularies and created relations will be exposed on the provided data store and can be physically taken home as a dump on USB sticks by the participants.

## The 'work' in workshop

Each workshop participant will give a short, high-level introduction to the vocabulary at hand, ideally following a suggested template prepared by the chairs. For this template, we initially suggest a breadth-first approach that starts at the high-level e.g. "my top object is 'lfcRoot', 'Object', 'Thing'" "My vocabulary has a 'part-of' relation". Each presentation will periodically be intervened by sequential bursts of activity by the participants to allow the creation of relations in a vocabulary/context of choice (VoL, VoID, plain SKOS/DC/RDF/OWL etc). Individual vocabulary sessions should not exceed 20 min depending on the number of participants / vocabulary sessions. Following these initial vocabulary sessions, small groups (2 to 3 people) start mapping or linking a chosen number of vocabularies and make sure that the resulting combined vocabulary can be communicated (via an interface with a graphical 3D representation of the 3D model). These groups are formed in a 'speed dating' programme, forcing people to collaborate and produce an end result after 20 minutes.

## Required Participants:

- Min 5, max 10-15 serious 'doers' familiar with practical modeling. In-depth knowledge in at least one engineering domain, national context, data model, vocabulary required.

Required resources:

- Web connectivity during the work shop
- Flip-boards, whiteboards, post-its, other office material
- A room for 20 people

- A data store (provided by the DURAARK consortium)

## Risks and problems:

- IPR issues of initial vocabularies. Results should be as liberal and open as possible
- If no 'tangible' results in the form of inter-vocabulary relations, profiles and a registry are produced, we have had 'just a mini-conference day'

## Example Workshop schedule:

- 09:00 Hello world & participant intro
- 09:30 intro to speed dating rules (recap of workshop intro sent out in preparation)
- 10:00 technical setup participants & test run with example vocabulary  
  - 10:45 *Break*
- 11:00 3 x 20 min speed dating sessions  
  - 12:00 *Lunch*
- 13:30 3 x 20 min speed dating sessions  
  - 14:30 *caffeine*
- 14:45 3 x 20 min speed dating sessions  
  - 15:45 *restroom*
- 16:00 wrap up, follow-up dates, medals of honor  
  - 17:00 *pub*

*This suggested outline can alternatively be restructured to allow one or two invited talks from representatives of the Semantic Web community and the built environment to attract a broader audience from both fields.*

## Potential Programme Committee:

<ul style="list-style-type: none"> <li>- Katranuschkov, Peter (TU Dresden, DE)</li> <li>- Kris McGlenn (TCDublin, IE)</li> <li>- Håvard Bell (Catenda AS, NO)</li> <li>- Bohms, Michel (TNO, NL)</li> <li>- Keane, Marcus (NUIGalway, IE)</li> <li>- Peter Bonsma (RDF Ltd., BG)</li> <li>- James O'Donnell (UCD, IE)</li> <li>- Bruno Fies (CSTB, FR)</li> <li>- Raimar Scherer (TU Dresden, DE)</li> <li>- Bram Mommers (Arcadis, NL)</li> <li>- Ed Curry (DERI, IE)</li> <li>- Martin Tamke (CITA, DK)</li> <li>- Rene Berndt (FhA, AT)</li> <li>- Ujwal Gadiraju (LUH, DE)</li> </ul>	<ul style="list-style-type: none"> <li>- Ed Corry (NUIGalway, IE)</li> <li>- Nam Vu Hoang (Aalto University, FI)</li> <li>- Karl-Heinz Häfele (KIT, DE)</li> <li>- Paul Scarponcini (Bentley, US)</li> <li>- Lieke Verhelst (Geonovum, NL)</li> <li>- Thomas Liebich (AEC 3, DE)</li> <li>- Nick Nisbet (AEC3, UK)</li> <li>- Chi Zhang (TU Eindhoven)</li> <li>- Hendro Wicaksono (KIT, DE)</li> <li>- Leandro Madrazo (ARC La Salle, ES)</li> <li>- Jean Brodeur (Natural Resource Canada)</li> <li>- Jyrki Oraskari (Aalto University, FI)</li> <li>- Lars Bjørkhaug (Catenda, NO)</li> <li>- Mathias Kadolsky (TU Dresden, DE)</li> </ul>
---	---

## Organisers:

- **Jakob Beetz** is an assistant professor at the Design Systems group of the Department of the Built Environment at the Eindhoven University of Technology. His research areas include Building Information Modeling and Semantic Web technologies.
- **Pieter Pauwels** is a postdoctoral researcher at the Department of Architecture and Urban Planning, Ghent University. He investigates how and to what extent information system support can be provided to experts in architecture and construction, thereby focusing on Building Information Modelling (BIM) and semantic web technologies.
- **Matthias Weise** is head of the software development department of AEC3, an international consulting company in the field of process optimisation in the building industry. AEC3 is frequently acting on behalf of the buildingSMART organization, the predominant standardization body in the AEC/FM field
- **Seppo Törmä** is a researcher at the Distributed Systems Group of the Department of Computer Science and Engineering at Aalto University. His fields of research include complex event processing, linked data and cognitive collaboration.
- **Stefan Dietze** is a research group leader at the L3S Research Center of the Leibniz University Hanover, Germany, where he is dealing with research in the fields of semantic & social web, knowledge-based systems and context adaptation with a focus on the use of semantic technologies for data and service integration