W3C LBD Community Group
Minutes - Call 05/04/2022

Attendees:

- Mathias Bonduel (KU Leuven and Neanex)
- Kevin Luwemba Mugumya (University of Nottingham Malaysia)
- Katja Breitenfelder (Fraunhofer IBP/ acatech - German Academy of Science and Engineering)
- Alex Donkers (Eindhoven University of Technology)
- Philipp Dohmen (Amberg Group AG Switzerland)
- Conor Shaw (University College Dublin)
- Joel Bender (Cornell University)
- Diellza Elshani (ICD)
- Manu Argyris (Amberg Group AG Switzerland)

Date and time

- 05/04/2022, Tuesday, 15:00-16:30@UTC/ 17:00-18:30@CEST/ 08:00-09:30@PDT/ 10:00-11:30@CDT

Moderator

- Mathias Bonduel

Agenda

1. Introduction of new participants
2. Presentation **"RDF-based Object Type Libraries" (OTLs) in construction project settings** - slides
3. Discussion/Q&A

Minutes

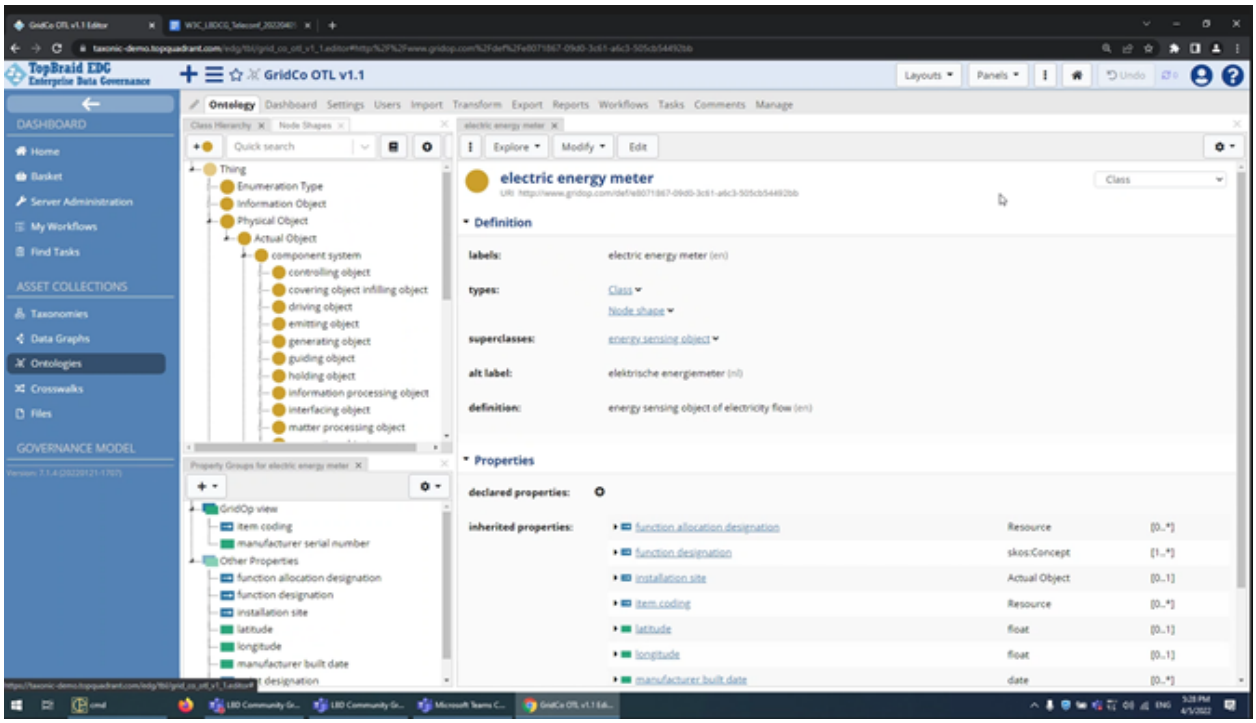
1. **Introduction of new participants**
    a. No new participants

2. **"RDF-based Object Type Libraries" (OTLs) in construction project settings**.(Jan Voskuil and Eliana Papoutsoglou, Taxonic)
    a. Presentation of OTLs
    i. Company Taxonic: 2012 founded, practical application of knowledge graphs, application in industry (infrastructure)

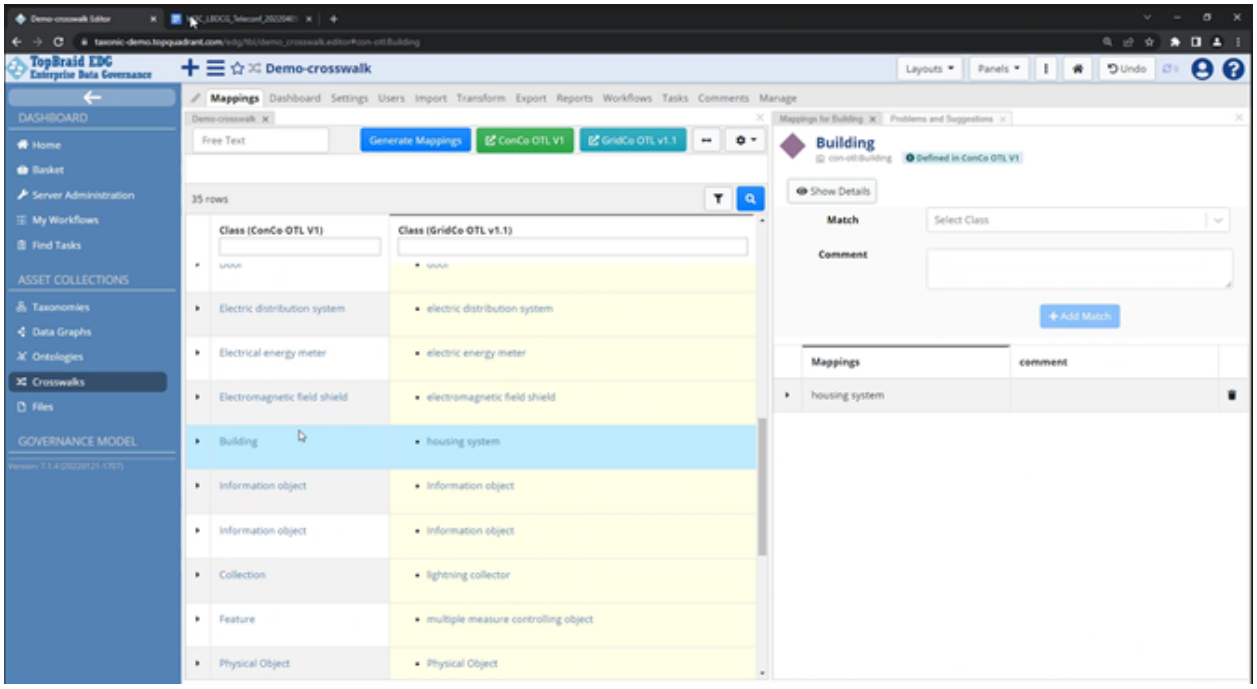
ii. Problem statement: “Datafication”: Huge number of different document types, Manual Processes etc., no exchange of data models -> knowledge graphs representing data models -> exchange of RDF based semantic data models
iii. Different perspectives when exchanging knowledge graphs
    1. Defined roles and responsibilities: principal & contractors
    2. Depending from the role-> different processes, data workflows and exchangeable models
iv. Solution: combining graphs, different classes (Conco classes, GridCo classes, ..) and “suggestions”
v. How mapping is generated: similarity of labels, adding more labels, third party taxonomies
vi. Approach for combining graphs:
    1. Project OTL (specific for the project)
    2. (Project data model ->) Project database

b. Life demo TopBraid EDG (by Taxonic)

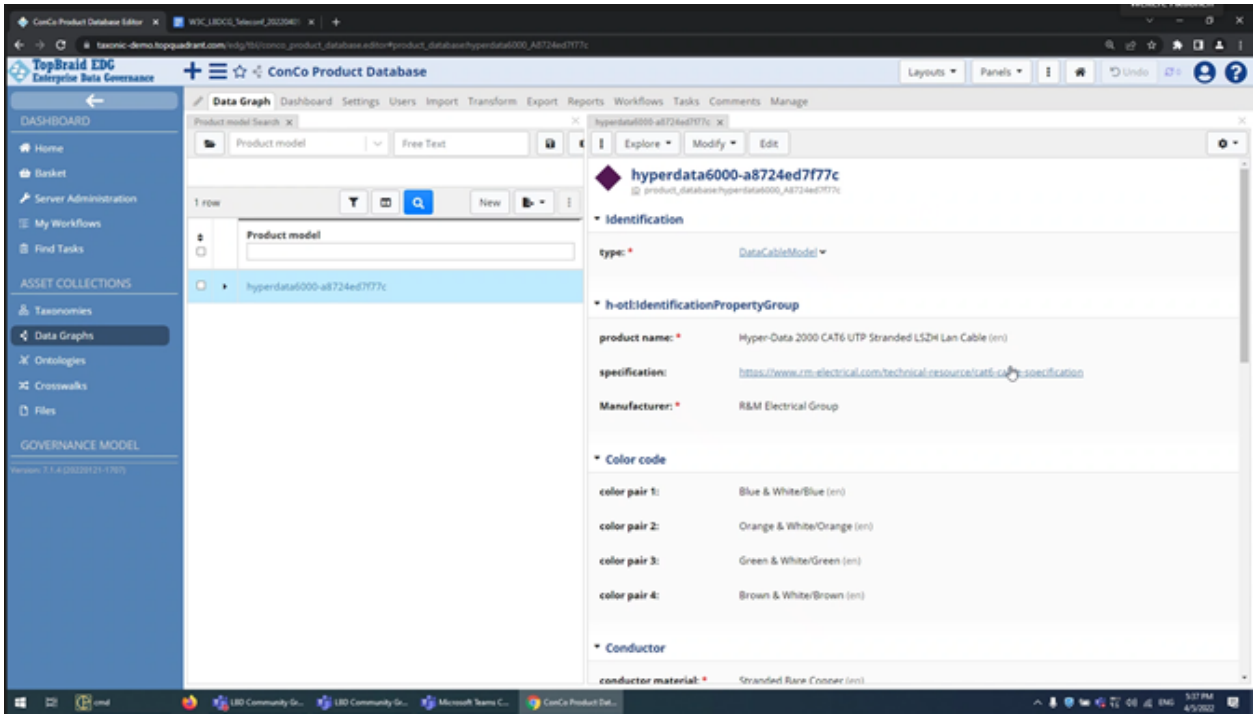
i. Role: Contractor, ontologies -> classes -> super classes, definitions, properties etc.



ii. Class mapping



iii. Product database (how an asset is supposed to be)



c. Wrap up/ Take-aways

i. As a contractor you can make fully use of the data model by aligning it to other project data models -> one standardized work process (based on mappings) -> data integration across (various) projects

3. **Q&A - Open Discussion**
    a. Q (Mathias): TopBraid EDG's Crosswalks function is used to create alignements, but which types of alignments (relation)? A (Jan): Basis for conversion on the fly, there are not so many distinctions made, but looking for "direct" relations. Enumerations is a good example, a contractor and a principle might be using them in a different way… Also transformations needed between (with SHACL f.i.). There is work to be done after the (individual) mapping, but a lot of it can be reused after time.
    b. Q (Mathias): Information requirements of the principle. Conto OTL as a leading model. How to manage it? A (Jan): Biggest challenge is to manage the complexity between different (project) models.
    c. Q (Mathias): Do you see opportunities for the contractor to share their own OTL definitions with the principle, when transferring a dataset? A (Jan): Generalizations are always possible. It can be the case that the receiver is also just not interested in things that use concepts (classes, properties) they don't know.
    d. Q (Phillip): what about using other inputs to derive ontologies from? Or use existing ontologies as a starting point? A (Jan): Book "Knowledge Graphs" (by A. Hogan and others) describes creating specific data models for knowledge graphs. The requirement capture of stakeholders for ontology modeling or data governance questions takes a lot of time.. Therefore, it is recommendable to (re)use existing ontologies of organizations (etc.) in order to demonstrate its advantages. E.g. taking a generic taxonomy as a starting point for requirement capture (for example within an enterprise).
    e. Q (Mathias): A lot of data models/ vocabularies etc. aren't shared but only used "behind the walls" of enterprises. A (Jan): People are changing, using OTLs, but most organizations are still scared to share definitions or data models considering it as sensitive data.
    f. Q (Mathias): Organizations often have put quite some time and effort into generating the OTL and they might not want other companies to benefit directly from their effort. How to support the change? A (Jan): Go first sharing success stories - more people will get involved. If others are reusing your concepts, it's also easier to do mappings afterwards. They can give feedback, which is a good thing.
    g. Q (Mathias): The OTL concept is a "quite Dutch concept" people from other regions are maybe not so familiar with. Why do you use knowledge graphs? A (Jan): Current project workflows are most frequently based on "Excel sheet" exchanges.. Compared to this, the OTL concept allows to automate the process making use of an overarching concept.
    h. Q (Mathias): What is the benefit of using Linked Data for OTLs (e.g. why not using json schemas or something else)? A (Jan) There will always be other ways to do it. It just saves us time if we apply Linked Data technologies for it. (Mathias)

Linked Data offers an interesting combination of advantages, but perhaps the most important is the fact that it's all based on (a layer of) standards.

i. Q (Mathias) A downside is that the huge amount of necessary different standards for Linked Data needed by a company might be hindering their application. How do you see their adoption being increased? One option might to have simple and generic applications that abstract the technical details of standards, another approach is to empower people to get ahead with the technology by themselves. A (Jan) training materials are crucial. There's some (mostly) theoretic material available. Eliana and I organized a training with the material published. (Mathias) interesting to share the links.

j. Q (Mathias) I often have the feeling that there's a different approach to Linked Data as a technology stack and ontologies (information models) between people who are trained computer scientists and people who come from a different background (e.g. construction engineering). As a poll question: who in today's call has a computer science education? (only Joel raises hand) I had the chance to have my education in the 70-80s, so I encountered many situations where data had to be transferred between applications. Learned to work with languages such as prolog. I believe some more modern trained computer scientists/engineers used to OOP languages such as C++ or Java will have indeed more difficulties dealing with something more abstract like ontologies/OTLs. In general computer engineers will rather focus on developing specific applications, while (some) computer scientists have learned to look a bit wider

k. (Mathias) Maybe we can conclude that OTL, an ontology on the scale of an organization, is a necessary bridge to reach a wider consensus on concepts over organizations. This is something I would like to have told to myself when I started my research. I've first looked into rather generic ontologies (following all best practices, etc) that are applicable enough in 99% of cases, but when I tried to apply those concepts with a specific organization it was a step too far as their own datasets were too different. (Jan) Indeed, organizations need to grow into it more gradually. The technology and concepts in themselves are not that easy

Next Call

- 19/04/2022, Tuesday, 15:00-16:30@UTC/ 17:00-18:30@CEST/ 08:00-09:30@PDT/ 10:00-11:30@CDT

We are interested in getting suggestions from the community about potential agenda items for the following calls. Please send your suggestions to public-lbd@w3.org, whether you have a short presentation to bootstrap the discussion, and an approximate duration you think the discussion will last.

Previous minutes

https://github.com/w3c-lbd-cg/lbd/tree/gh-pages/minutes