Report for Ontology Metadata

task group of the

Vocabulary and Semantic Services Interest Group

Dr. Biswanath Dutta[1], Dr. Clement Jonquet[2]

[1]Assistant Professor, DRTC, Indian Statistical Institute, Bangalore

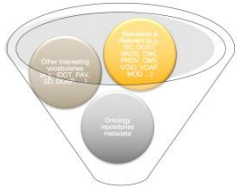
[2]Associate Professor, LIRMM, University of Montpellier

RDA P14 – Helsinki, 23-25 October 2019

A return of experience with the Ontology Metadata task group

(Work in progress)

One goal → Some work, Methodology → Experimentation & Contextualization → Generalization, Standardisation, recommandation



AgroPortal LIRMM

RDA RESEARCH DATA ALLIANCE

Interested in studying ontology metadata practices to discuss and provide recommendations

As any data, semantic resources (ontologies, thesaurus, vocabularies)…. need to be FAIR

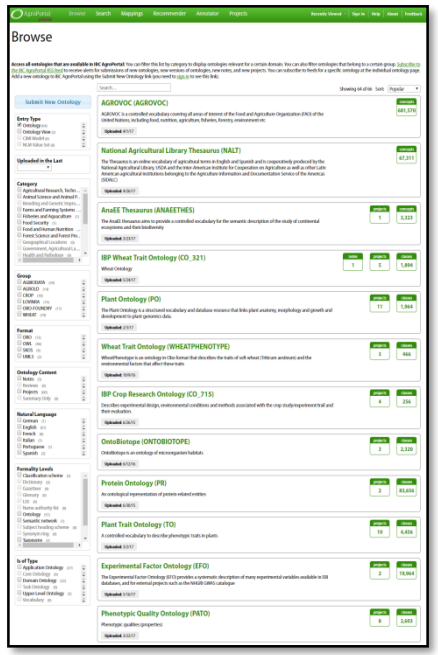
- The FAIR principles have established the importance of using standard vocabularies or ontologies to describe FAIR data and to facilitate interoperability and reuse…

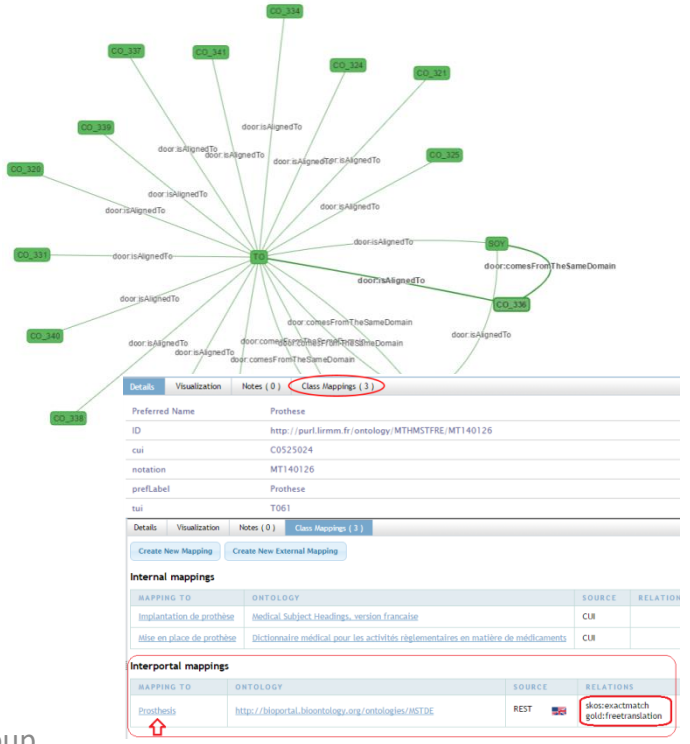
- Explosion of the number of ontologies/vocabularies

- Cumbersome to identify the ontologies, we need and manage their overlap.

Ontology repositories help to make them FAIR

Findable **A**ccessible **I**nteroperable **R**e-usable





MOD 1.4 (August, 2018)

(https://www.isibang.ac.in/ns/mod/index.html)

Stable

MOD: Metadata for Ontology Description and publication

Release August 2, 2018

This version:
http://www.isibang.ac.in/ns/mod/1.4

Latest version:
http://www.isibang.ac.in/ns/mod/1.4

Previous version:
http://www.isibang.ac.in/ns/mod/1.2
https://www.isibang.ac.in/ns/mod/1.1
https://www.isibang.ac.in/ns/mod/1.0

Revision:
1.4

Authors:
Biswanath Dutta, (Indian Statistical Institute)
Clement Jonquet, (University of Montpellier)

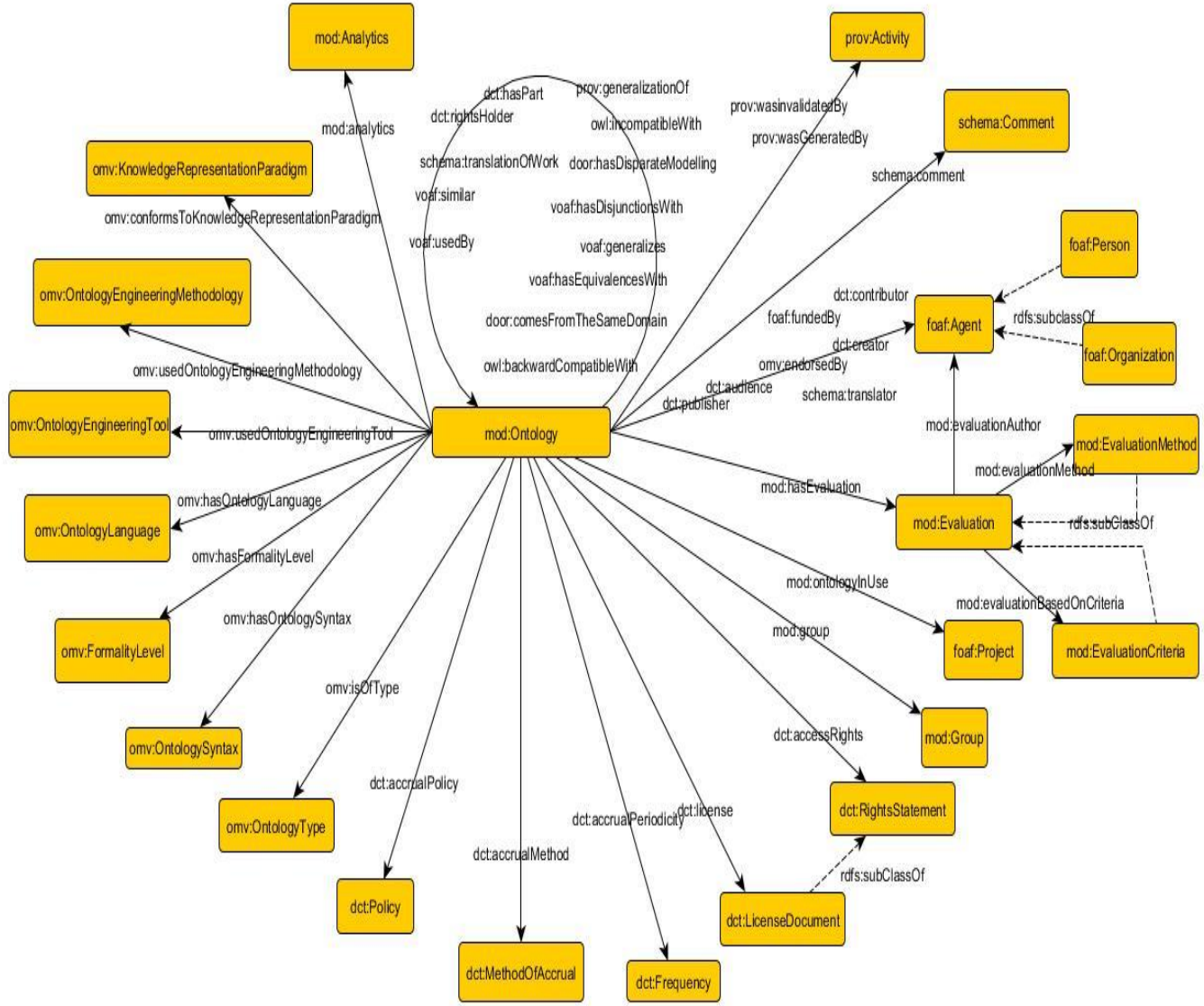
Contributors:
Anne Toulet, (University of Montpellier)
Udaya Varadarajan, (Indian Statistical Institute)

Publisher:
http://www.isibang.ac.in/

Download serialization:
Format JSON LD | Format RDF/XML | Format N Triples | Format TTL

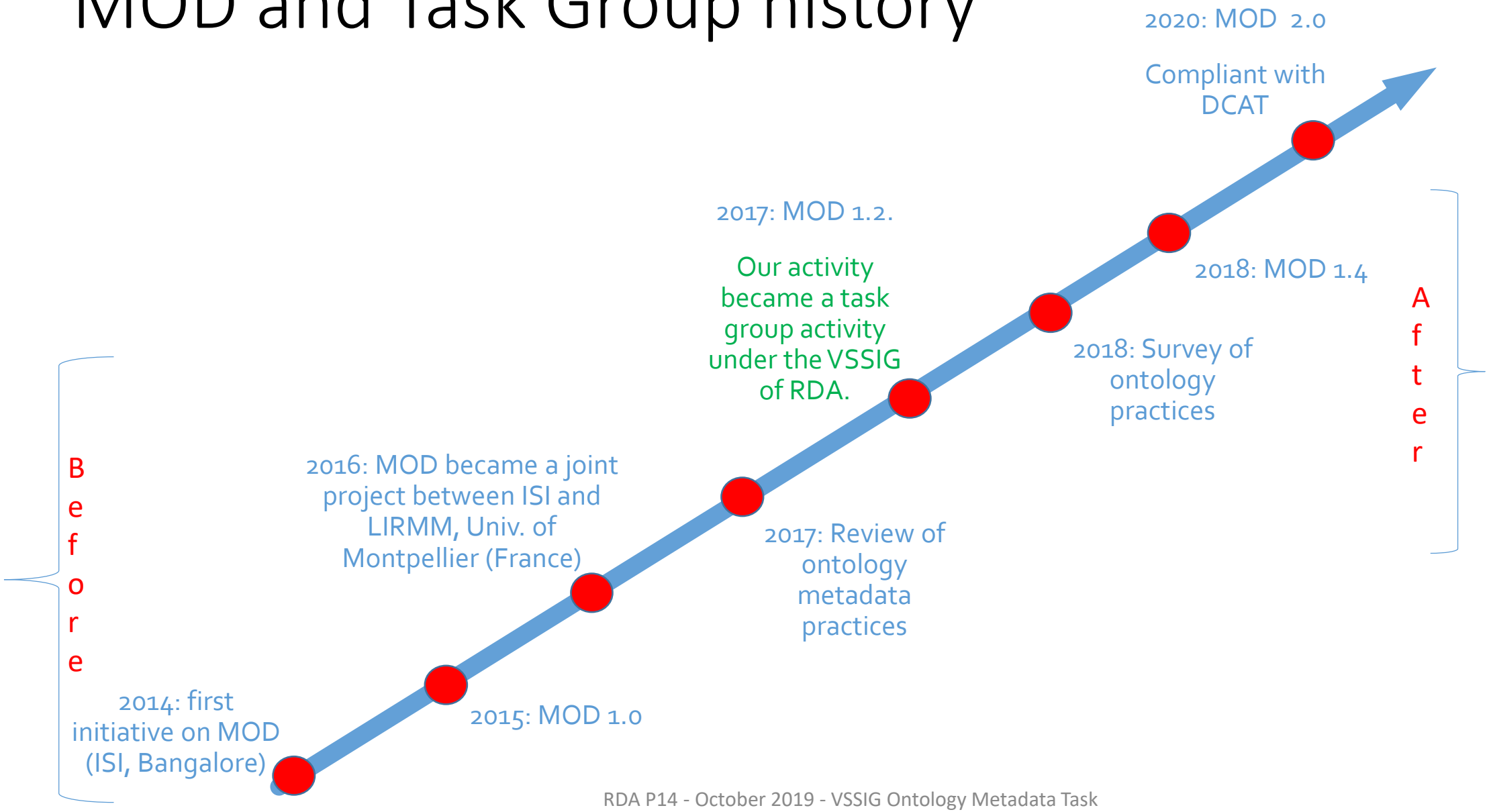
License:
License Creative Commons Attribution 4.0

Cite as:
Dutta, B., Toulet, A., Emonet, V. and Jonquet, C. (2017). New Generation Metadata vocabulary Description and Publication. In E. Garoufallou, S. Virkus, R. Siatri and D. Koutso Communications in Computer and Information Science (CCIS) 755, proceedings of 11th M Semantics Research Conference (MTSR 2017), November 28 - December 1, 2017, Talli Springer Nature, pp. 173-185.



Classes: 24
Object property: 44
Data property: 96

MOD and Task Group history



With the goal to allow to find Mr. Right ontology, enable communication between user and machine, machine and machine, ask interesting questions, and do analytics

Question: What are the ontologies endorsed by the RDA Wheat Data Interoperability Group (RDA WDI) and the National Science Foundation (NSF)?

```
SELECT DISTINCT ?Ontology ?Author

    WHERE {

{?x a mod:Ontology;
  omv:endorsedBy <https://www.rd-alliance.org/groups/wheat-data-interoperability-wg.html>;
  dct:title ?Ontology.

}

    UNION

{?x a mod:Ontology;
  omv:endorsedBy <http://dbpedia.org/resource/Category:National_Science_Foundation>;
  dct:title ?Ontology.
  }

    OPTIONAL {?xdct:creator ?Author .}
  }
```

AgroPortal Landscape page

- Display "per property"
  - Global presentation of the properties
  - Synthesis diagrams & listing

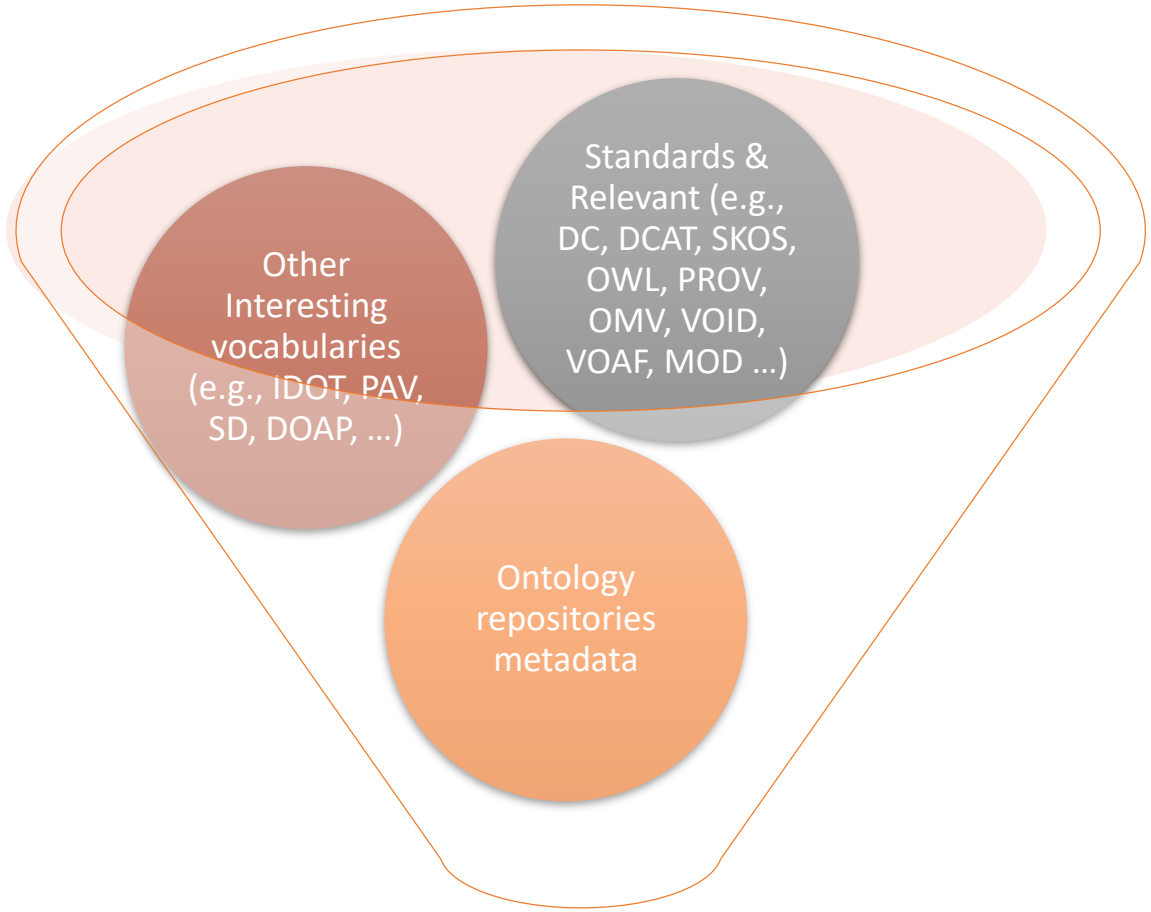
- Metadata automatically extracted from the files and authored by us and the ontology developers

- Allows to explore the agronomical ontology landscape by automatically aggregating the metadata fields of each ontologies in explicit visualizations (charts, term cloud and graphs).

Jonquet, C., Toulet, A., Dutta, B., Emonet, V.: **Harnessing the power of unified metadata in an ontology repository: the case of AgroPortal**. *Data Semantics, 2018*.

Review of ontology metadata practices

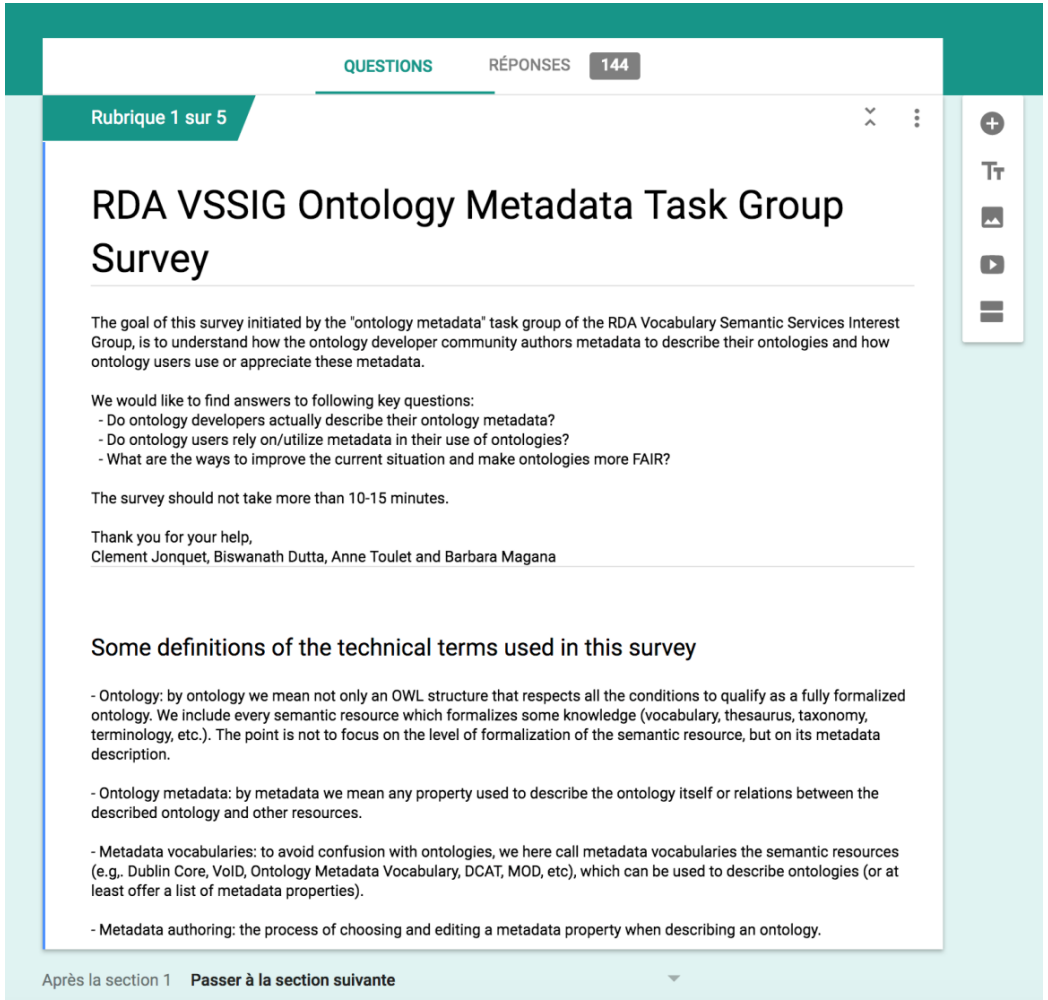
- Analysis of the use of metadata vocabularies in describing ontologies (by ontology developers)
  - 805 ontologies analyzed
- Analysis of the existing metadata vocabularies
  - 23 metadata vocabularies
- Analysis of the uses of metadata vocabularies in various ontology libraries and repositories (e.g., BioPortal, MMI)
  - 13 libraries



346 relevant properties that could be used to describe ontologies

127 used to build a new metadata model inside AgroPortal

Survey of ontology metadata practices



QUESTIONS RÉPONSES 144

Rubrique 1 sur 5

RDA VSSIG Ontology Metadata Task Group Survey

The goal of this survey initiated by the "ontology metadata" task group of the RDA Vocabulary Semantic Services Interest Group, is to understand how the ontology developer community authors metadata to describe their ontologies and how ontology users use or appreciate these metadata.

We would like to find answers to following key questions:
- Do ontology developers actually describe their ontology metadata?
- Do ontology users rely on/utilize metadata in their use of ontologies?
- What are the ways to improve the current situation and make ontologies more FAIR?

The survey should not take more than 10-15 minutes.

Thank you for your help,
Clement Jonquet, Biswanath Dutta, Anne Toulet and Barbara Magana

Some definitions of the technical terms used in this survey

- Ontology: by ontology we mean not only an OWL structure that respects all the conditions to qualify as a fully formalized ontology. We include every semantic resource which formalizes some knowledge (vocabulary, thesaurus, taxonomy, terminology, etc.). The point is not to focus on the level of formalization of the semantic resource, but on its metadata description.

- Ontology metadata: by metadata we mean any property used to describe the ontology itself or relations between the described ontology and other resources.

- Metadata vocabularies: to avoid confusion with ontologies, we here call metadata vocabularies the semantic resources (e.g,. Dublin Core, VoID, Ontology Metadata Vocabulary, DCAT, MOD, etc), which can be used to describe ontologies (or at least offer a list of metadata properties).

- Metadata authoring: the process of choosing and editing a metadata property when describing an ontology.

Après la section 1 **Passer à la section suivante**

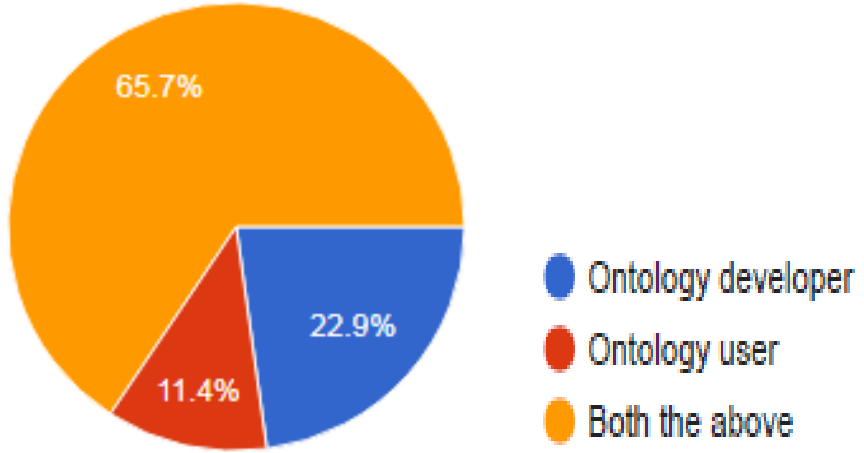
With the goal to answer the following questions:

- Do ontology developers actually describe their ontology metadata?

- Do ontology users rely on/utilize metadata in their use of ontologies?

- What are the ways to improve the current situation and make ontologies more FAIR?

Total 168 participants.



The survey report is available here:
https://zenodo.org/record/3484530#.Xa8Qe5IzZdh

Survey of ontology metadata practices: Findings

- Variety of metadata vocabularies (e.g., DC, DCT, PROV, VOiD, DCAT, Schema.org)
  - Interestingly: the only ontology specific metadata vocabulary OMV (first published in 2005) is found to be hardly used by the community
  - No existing vocabularies really covers enough aspects to be used solely

- Metadata vocabularies do not rely on one another although there is a strong overlap
  - Multiple properties to capture similar information (e.g., dc:license, and cc:license)
  - For instance, 25 properties available for dates

- Reviewed ontology libraries and repositories use some metadata elements but do not always use standard metadata vocabularies

- 16% of ontologies did not use any metadata properties, 43% use less than 10 properties
  - Properties facilitated by ontology editors are more frequent
  - Confusion of use: DC/DC Term or SKOS documentation properties

Our future goal

- Turn MOD 1.4 into an extended version, MOD 2.0, compliant with the DCAT specification (v2.0)
- Produce an “application profile” for the description of ontologies
- Discuss with the various ontology editor (e.g., Protégé, VocBench) on integration of MOD in the software
- Automatize the process of creating mod:Ontology instances from ontology libraries (e.g., BioPortal, AgroPortal, OBO Foundry)
  - Exchange the content of these libraries without changing their internal data models

Conclusion

- *FAIR: I2. (Meta)data use vocabularies that follow FAIR principles*

- We need a standard way to describe semantic resources and we need this to be adopted and shared by ontology stakeholders
  - Don’t reinvent the wheel => make a profile DCAT

- A collaborative approach is required => RDA, then maybe a W3C Recommendation

Participation, contribution

- Join us to discuss the subjects on the Slack channel

  #tg-ontology-metadata

- Follow the project on GitHub: https://github.com/sifrproject/MOD-Ontology

- Follow the project on ResearchGate: https://www.researchgate.net/project/MOD-Metadata-for-Ontology-Description-and-publication

Questions? Feedback?

Thank you

bisu@drtc.isibang.ac.in

@biswanathdutta