Report for Ontology Metadata task group of the

Vocabulary and Semantic Services Interest Group

Dr. Biswanath Dutta¹, Dr. Clement Jonquet²

¹Assistant Professor, DRTC, Indian Statistical Institute, Bangalore

²Associate Professor, LIRMM, University of Montpellier

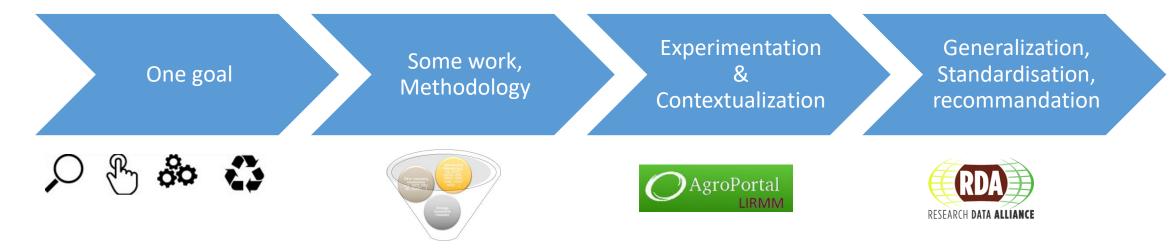
RDA P14 – Helsinki, 23-25 October 2019





A return of experience with the **Ontology Metadata** task group

(Work in progress)

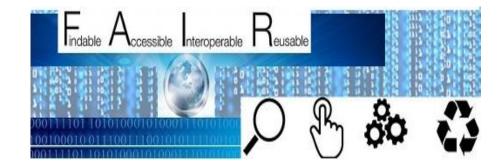


Interested in studying ontology metadata practices to discuss and provide recommendations



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





Re-usable

AgroPortal Browse	Search Mappings Recommender Annotator Projects Remaily Ver	wed = Sign in Telp A	beet Feelback
Browse			
	IIIC Agrivitents into can liter this list by caregory to diplay antiologies relevant for a certain domain. You can also filter antiologies are version of contrologies, new notice, and new projects. You can subscribe to field-for a give the signal process of the signal process field.	ies that belong to a certain gro colic ontology at the individual	up Subscribe to Londology page
	Seath	Showing 64 of 66 Sort	
Submit New Ontology	AGROVOC (AGROVOC)		
Entry Type If: Ontology no. Ontology New On Othillood I on N M Whate Ser on	AGROVC is constrained excellulary converse of annexed assess of the Food and Agriculture Organization (FAC) of the Unided Nations, including load, methion, agriculture, filteries, forestry, environment etc. Splanded 4502		
	National Agricultural Library Thesaurus (NALT)		caceta
Uploaded in the Last	The Theorems is an online exclusing of approximate terms in trights and "paints and is cosporatively produced by the National Approximate (Davy COSA) and the times research entities for Consections on Approximate an Well as other Later American-approximate institutions belonging to the Approximate Information and Documentations Serviced the American ODINC)		67,311
Category Agricultural feasurch, Techn + Animal Science and Animal P.	Splanded #35-0		
Intending and Genetic Improv Interns and Farming Spatems Probeties and Aquacubure Proof and Human Numition Frond and Human Numition Front Science and Farest Prov. Generative/aclocations	AnsEE Thesaurus (ANAEETHES) The katel documents are to possile a controlled enclosely for the senarit's description of the study of contenend encoderers and their bandwatey (senarity area)	99973 1	0000001 3,323
Government, Agricultural La	IBP Wheat Trait Ontology (CO_321) Wheat Onslay:	ana pajeta 1 5	Circus 1,894
Addictation pre	Plant Ontology (PO) The Plant Country is a mutual excluding and database scores that inits plant anatomy, acceptology and positil and development topic are provided in the	944/5 11	rieses 1,964
Format 000 (12 2 001 (22 2 001 (22 2 001 (22 2 001 (22 2 001 (22 2 001 (22 2 001 (22 2 001 (22 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	[Network 2017] Wheat Trait Ontology (WHEATPHENOTYPE) Whortwork Jona and the for the trait do not the trait of with electric follows and trace and the minimum and Jona of effort the trait do	3 90943	Geren 465
Ontology Content Notes on Devices on Projects no Summary Only in	BP Crop Research Ontology (CO_715)	Segen 4	254
Natural Language	Describes experimental design, environmental conditions and methods associated with the corp study/operimentitical and their evaluation.		
Tornality Levels	OntoBiotope (ONTOBIOTOPE) Ordebruges are including of encougeness habitas		
Clasificationademe (r) Dictionary (r) Gautteor (r) Gautte	Protein Ontology (PR) At ontological representation of patient existing Inputate Clinics	prijers 2	83,655
Conclopy on Semantic network in Subject heading scheme in Sponyment in Subsective in	Plant Trait Ontology (TO) A consult-functionary to the observation of the second secon	94945 19	(bras) 4,456
h of Type Application Ontology (c) Descal Detailing (c) Task Omdogy (c) Upport and Ontology (c) Upport and Ontology (c) Upport and Ontology (c) Upport and Ontology (c)	Experimental Factor Ontology (IFO) The Eperimental rate Codeg (CFO) (20 modes systematic description of many experimental variables and able to 10 distance, and in entrol projects such as the WHOR COMS CODE years (Special State)	projects 2	riteren 19,954
	Phenotypic Quality Ontology (PATO) Resurgic qualics projection	jajen. I	Chine 2,603

API Documentation

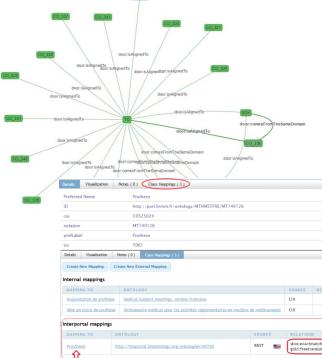
 Provide Comparing Control (Charges, Classes, etc) and initial endocring Beach, Annotatic, Ryssening the AP lusing a web browser (Charges and Perlois, evolvel, weight while II. does not plant by our systematic that are control endocring the annotation of the AP lusing a web browser (Charges and Perlois, evolvel, weight while II. does not plant by our systematic that are control endocring the annotation of the AP lusing a web browser (Charges and Perlois, evolvel, weight while II. does not plant by our systematic that are control endocring the annotation of the AP lusing a web browser to endocring the annotation of the AP lusing a systematic that are control endocring the annotation of the AP lusing a systematic that are control endocring the annotation of the AP lusing a systematic that are control endocring the annotation of the AP lusing a systematic that are control endocring the annotation of the AP lusing a systematic that are control endocring the annotation of the AP lusing a systematic that are control endocring the annotation of the AP lusing a systematic the are endocring the annotation of the AP lusing a systematic the are endocring the annotation of the AP lusing a web browser to response the AP lusing a systematic the are endocring and annotation of the AP lusing a web browser to response the AP lusing a systematic the are endocring a systematic the are endocring a systematic the are endocring the are endocring

SPARQL httpd server v1.1.5-122-{

KB ontologies_api

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT * WHERE { ?s ?p ?o } LIMIT 10





MOD 1.4 (August, 2018)

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

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, (<u>University of Montpellier</u>) Udaya Varadarajan, (<u>Indian Statistical Institute</u>)

Publisher:

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

Download serialization:

Format JSON LD Format RDF/XML Format N Triples Format TTL

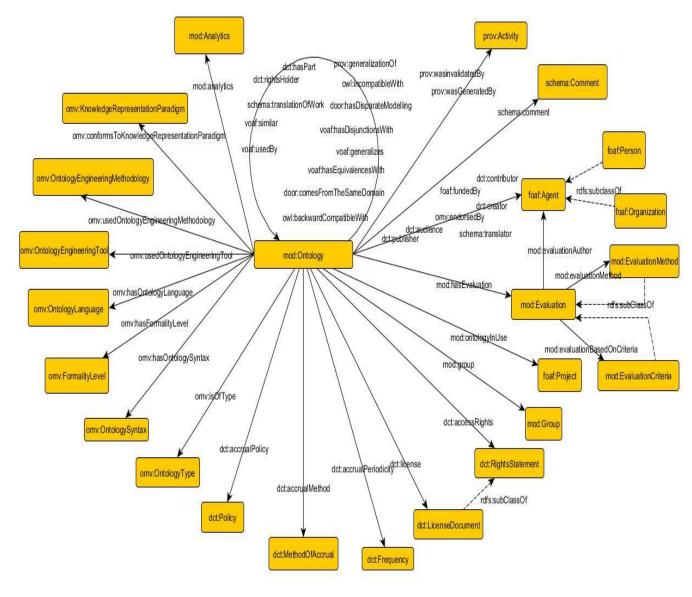
License:

License Creative Commons Attribution 4.0

Cite as:

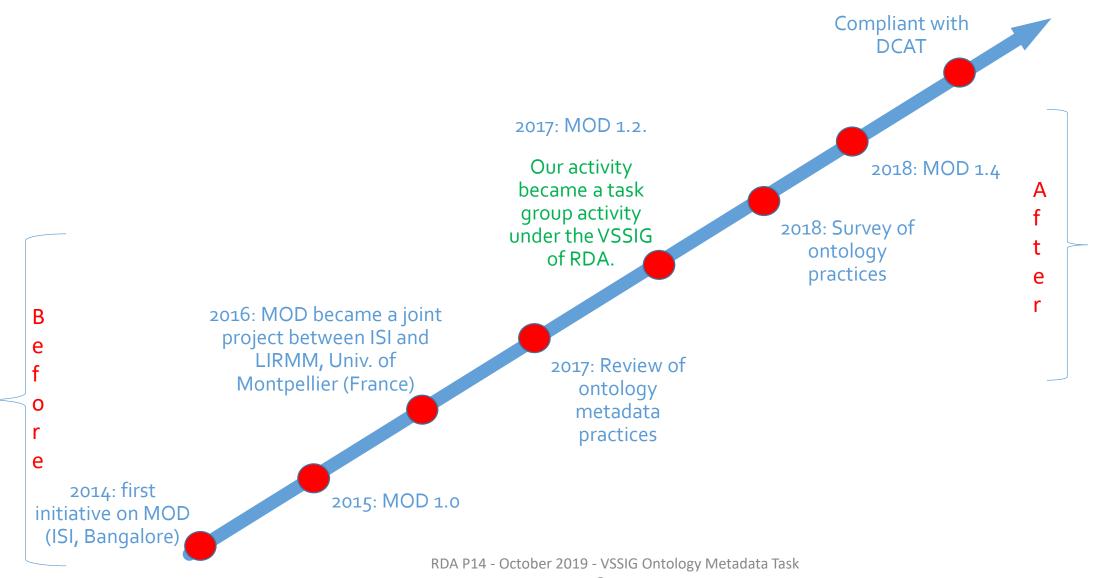
Dutta, B., Toulet, A., Emonet, V. and Jonquet, C. (2017). New Generation Metadata vocabulary Description and Publication. In E. Garoufallou, S. Virkus, R. Siatri and D. Koutso Communications in Computer and Information Science (CCIS) 755, proceedings of 11th M Semantics Research Conference (MTSR 2017),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





Classes: 24 Object property: 44 Data property: 96

MOD and Task Group history



2020: MOD 2.0

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

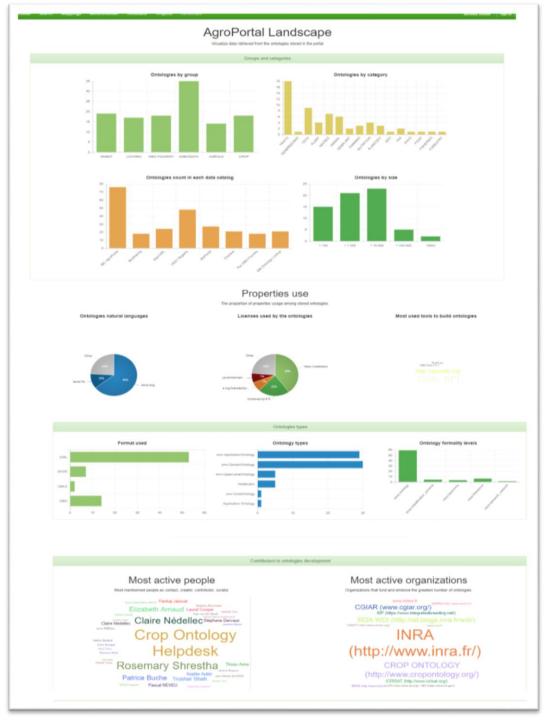
WHERE {

```
{?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).

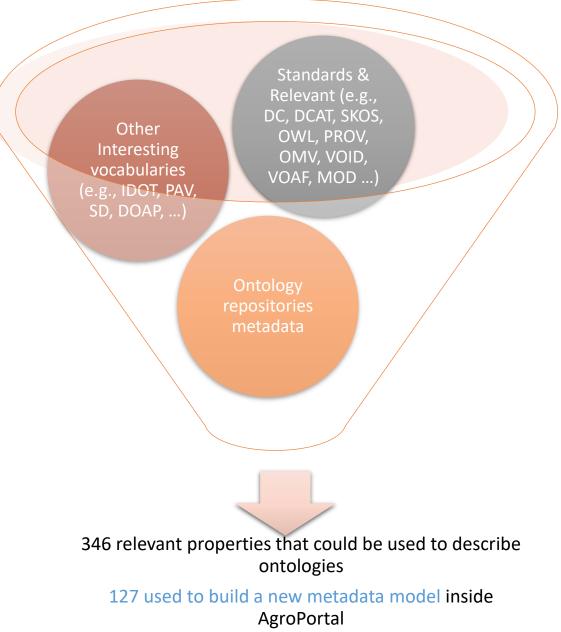


Jonquet, C., Toulet, A., Dutta, B., Emonet, V.: Harnessing the power of unified metadata in an ontology repository: the case of AgroPortal. *Data Semantics*, 2018.

RDA P14 - October 2019 - VSSIG Ontology Metadata Task Group

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

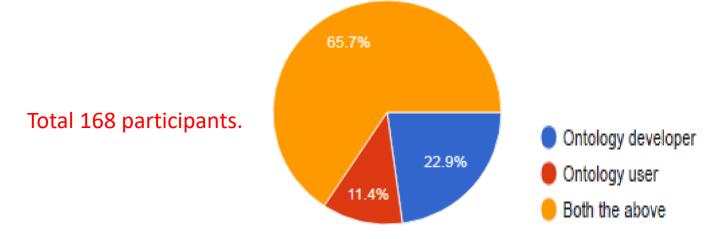


Survey of ontology metadata practices

QUESTIONS	RÉPONSES	144				
Rubrique 1 sur 5				×	: (Ð
					1	T
RDA VSSIG Ontology	Metada	ata Task	Group		Į	•
Survey						
The goal of this survey initiated by the "ontology metada Group, is to understand how the ontology developer com ontology users use or appreciate these metadata. We would like to find answers to following key questions	munity authors n					•
 Do ontology developers actually describe their ontolog Do ontology users rely on/utilize metadata in their use What are the ways to improve the current situation and 	of ontologies?	s more FAIR?				
The survey should not take more than 10-15 minutes. Thank you for your help, Clement Jonquet, Biswanath Dutta, Anne Toulet and Bar	bara Magana					
Some definitions of the technical ter	ms used in	this survey				
 Ontology: by ontology we mean not only an OWL struct ontology. We include every semantic resource which for terminology, etc.). The point is not to focus on the level of description. 	nalizes some kno	wledge (vocabulary,	thesaurus, taxonom	ıy,		
- Ontology metadata: by metadata we mean any property described ontology and other resources.	vused to describe	e the ontology itself o	or relations between	the		
- Metadata vocabularies: to avoid confusion with ontolog (e.g., Dublin Core, VoID, Ontology Metadata Vocabulary, I least offer a list of metadata properties).						

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?



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: <u>https://github.com/sifrproject/MOD-Ontology</u>
- Follow the project on ResearchGate: <u>https://www.researchgate.net/project/MOD-Metadata-for-Ontology-Description-and-publication</u>

Questions? Feedback?

Thank you

bisu@drtc.isibang.ac.in

@biswanathdutta





RDA P14 - October 2019 - VSSIG Ontology Metadata Task Group