Controlled Vocabulary Workshop

Ca'Foscari University
Venice, 8 May 2017

Agenda

First part of the workshop (14:00-15:00)

Introduction and work to date

Prototype implementation of the COAR Resource Type vocabulary in DSpace

COAR Resource Type Vocabulary in the Classification Server of the University of Vienna

Second part of the workshop (15:30-17:00)

Work to implement the COAR Vocabularies into existing guidelines: OpenAIRE-Literature Repositories and OpenAIRE-CRIS

Adopting the vocabularies in national and regional repositories networks: Japan, Portugal, Latin America

Resource Type 2.0: Public Input & Discussion

Next steps for COAR Controlled Vocabularies

Closure

Introduction and work to date
Isabel Bernal, DIGITAL.CSIC, CSIC

COAR AGENDA

- Advocacy & leadership
- **Repository Interoperability**
  1. COAR Vocabularies
  2. Open Metrics
  3. COAR Interoperability Project
- Repository Content
- Webinar and Discussion Series, Support &Training

Controlled Vocabularies Working Group

- **Interest Group created in 2014 with representatives of COAR member institutions and external experts**

- Editorial Board monitors trends, creates and maintains COAR vocabularies and promotes good practices. Its members cover many languages (and growing!): English, German, Czech, Chinese, Spanish, French, Dutch, Italian, Catalan, Portuguese, Japanese, Russian, Turkish

- **Goals: Spread benefits of controlled vocabularies; foster usage as they help in searching and enhance interoperability on the web**

More interoperability on the web

- Rapid growth of controlled vocabularies, particularly in the semantic web landscape.

- They are often discipline-based, focus on specific aspects of Open Access/ Open Data contents, or enrich the description of resources by looking into information like funders, projects and research institutions

- The group promotes usage of widely known international standards and avoids reinventing the wheel

WHERE ARE WE NOW?

- Version 1.1, October 2016. Previous version July 2015

- Draft Version 1.0, April 2017

RESOURCE TYPES

ACCESS MODE

RESOURCE VERSIONS

DATES TYPES

- Under investigation

- Under debate
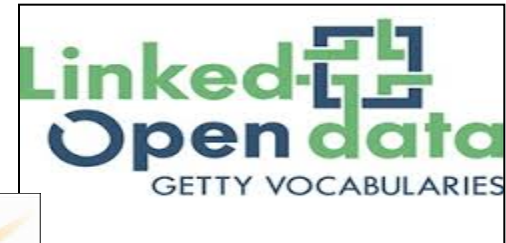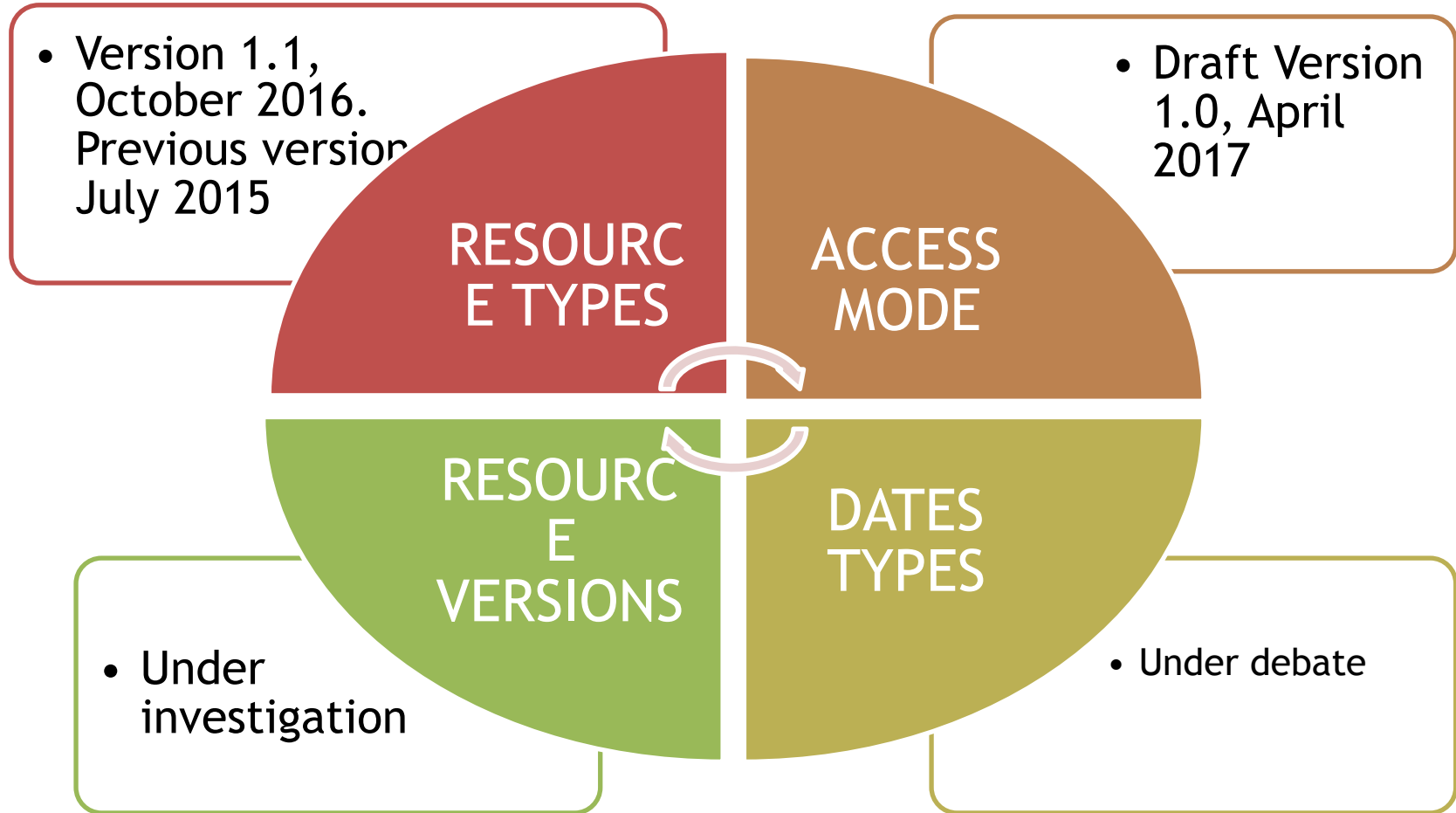
Linkage between info:eu-repo and COAR vocabularies

- **Since 2008 info:eu-repo** has been part of DRIVER (and later on, OpenAire) guidelines in order to harmonize metadata usage across repositories
- **Wide adoption internationally** (by DSPACE, Eprints repositories) and OJS-based journals
- However, **info eu initiative lacks sustainable institutional support and a truly global management team >>>> COAR Working Group** emerged, in part, to fill the gap and secure its maintenance

INFO:EU-REPO VOCABULARIES

Publication Types

Access Rights

Versions

Grant agreement identifiers

Resource Identifier schemes

Data types and value

Author Identifier

Classification schemes

Object types

COAR VOCABULARIES

Resource Types

Access mode

Version Types

Date Types

COAR Resource Types Vocabulary: Current Version

A bit about methodology

Some vocabularies mapped

FRBR-aligned Bibliographic Ontology (fabio)

Eprints Type Vocabulary Encoding Scheme

DCMI Type Vocabulary

- **Multilinguality** and support for regional differences
- **Revision of existing vocabularies and terms from info:eu-repo**
- **Matching with other vocabularies and dictionarie**s referring to the same or similar thing
- **Statistical analysis** of terms used in repository metadata (using BASE aggregator)
- **Use of VocBench**, a web-based, multilingual, collaborative development platform for managing OWL ontologies, SKOS(XL) thesauri and generic RDF datasets

Main features

- **Currently available in 12 languages, contains more than 50 concepts and uses SKOS format**
- **Hierarchical structure**: the concepts of the vocabulary are organized from a generic down to a granular level
- The vocabulary establishes relations between **broad and narrower terms** and also between **preferred and alternative terms**
- There are concepts in the vocabulary that are **mapped with concepts/terms belonging in other like minded vocabularies**
- Extends info:eu-repo vocabulary for publication types
- We are embarking on Version 2.0 soon, please have your say!

⇑ -- conference paper --

URI	http://purl.org/coar/resource_type/c_5794
Definition (en)	A conference paper that is submitted to a conference and presented to the audience. The conference paper is published in proceedings.
Preferred label (en)	conference paper
Preferred label (zh)	会议论文
Preferred label (es)	comunicación de congreso
Preferred label (it)	comunicazione a congresso
Preferred label (ru)	статья для конференции
Preferred label (ca)	comunicació de congrés
Preferred label (pt)	comunicação em conferência
Preferred label (de)	Konferenzbeitrag
Preferred label (fr)	article dans une conférence
Preferred label (nl)	conferentie bijdrage
Preferred label (ja)	会議発表論文
Preferred label (tr)	konferans belgesi
Alternative label (en)	conference papers
Alternative label (en)	conferencePaper
Alternative label (es)	presentación de congreso
Alternative label (es)	comunicacion de congreso
Alternative label (es)	presentacion de congreso
Alternative label (ru)	доклад конференции
Alternative label (ru)	тезисы конференции
Alternative label (ru)	тезисы докладов
Alternative label (ru)	доклад для конференции
Alternative label (ru)	статьи конференции
Alternative label (zh)	會議論文
Alternative label (zh)	會議文章
Alternative label (zh)	会议文章
Alternative label (pt)	comunicação em congresso
Alternative label (pt)	artigo de conferência
Alternative label (pt)	artigo de congresso
Alternative label (pt)	apresentação em congresso
Alternative label (pt)	apresentação em conferência
Alternative label (fr)	article de conférence
Alternative label (pt)	comunicação
Alternative label (ru)	статья для научной конференции
Exact match	http://purl.org/eprint/type/ConferencePaper
Exact match	http://purl.org/info:eu-repo/#semantics/conferencePaper
Broad match	http://purl.org/dc/dcmitype/Event
Broad match	http://purl.org/ontology/bibo/Conference
Broader	conference proceedings

Concept URI

Definition

Scheme

Hierarchies and mappings

Labels in different languages

http://purl.org/coar/resource_type/c_5794
conference paper

Definition
A conference paper that is submitted to a conference and presented to the audience. The conference paper is published in proceedings.

Property	Value
rdf:type	skos:Concept
skos:inScheme	http://purl.org/coar/resource_type
skos:broader	http://purl.org/coar/resource_type/c_f
skos:exactMatch	http://purl.org/eprint/type/Conference http://purl.org/info:eu-repo/#semantics/conferencePaper
skos:broadMatch	http://purl.org/dc/dcmitype/Event http://purl.org/ontology/bibo/Conference

Property	Value
dcterms:modified	2016-09-05T12:11:20Z
skos:definition	http://purl.org/coar/resource_type/def_3380

prefLabel	altLabel	Lang
conference paper	conference papers conference paper	en
会议论文	會議文章 会议文章	zh
comunicación de congreso	presentación de congreso comunicacion de congreso presentacion de congreso	es
comunicazione a congresso		it
статья для конференции	доклад конференции тезисы конференции тезисы докладов доклад для конференции статьи конференции статья для научной конференции	ru
comunicació de congrés		ca
comunicação em conferência	comunicação em congresso artigo de conferência artigo de congresso apresentação em congresso apresentação em conferência comunicação	pt
Konferenzbeitrag		de
article dans une conférence	article de conférence	fr
conferentie bijdrage		nl
会議発表論文		ja
konferans belgesi		tr

As RDF/XML | As TURTLE

Powered by Loddy

Concepts are the fundamental element in the vocabulary. When referring to concepts in repositories metadata scheme they must be accompanied by their respective URI.
It is up to the repository to include one or more labels associated with the concept and also to tag one resource with more than one concept

VocBench: Vocabulary Editing and Workflow Tool

Concept

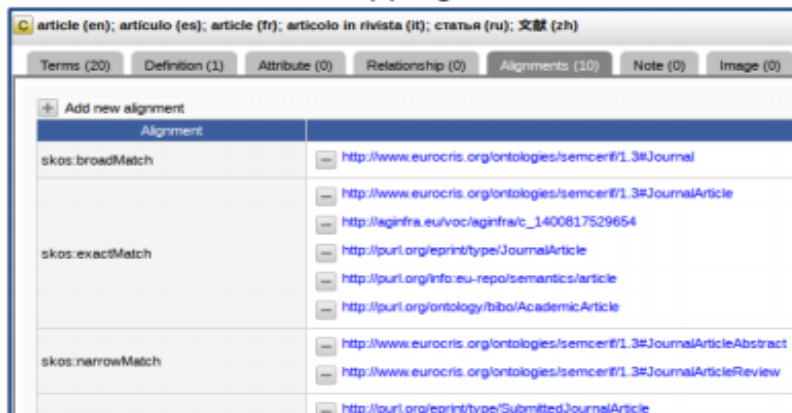VocBench VERSION 2.3 [Build 20150423] (DEVELOPMENT)

Recent changes | Concepts | Properties | Schemes | Validation | SPARQL | ICV

Concepts ☑ Show URI ☐ Show non-preferred

URI: http://purl.org/coar/resource_type/c_6501

- **C** annotation (en); comentario (es); commento (it); примечание (ru); 标注 (zh)
- **C** article (en); articulo (es); article (fr); articolo in rivista (it); статья (ru); 文献 (zh)
- **C** audiovisual (en); material audiovisual (es); materiale audiovisivo (it); 视听 (zh)
- **C** bachelor thesis (en); proyecto fin de carrera (es); tesi di laurea (it); диплом бакалавра (ru); 本科毕业论文 (zh)

Mappings

C article (en); artículo (es); article (fr); articolo in rivista (it); статья (ru); 文献 (zh)

Terms (20) | Definition (1) | Attribute (0) | Relationship (0) | Alignments (10) | Note (0) | Image (0)

+ Add new alignment

Alignment	
skos:broadMatch	http://www.eurocris.org/ontologies/semcerif/1.3#Journal
skos:exactMatch	http://www.eurocris.org/ontologies/semcerif/1.3#JournalArticle
	http://aginfra.eu/voc/aginfra/c_1400817529654
	http://purl.org/eprint/type/JournalArticle
	http://purl.org/info:eu-repo/semantics/article
	http://purl.org/ontology/bibo/AcademicArticle
skos:narrowMatch	http://www.eurocris.org/ontologies/semcerif/1.3#JournalArticleAbstract
	http://www.eurocris.org/ontologies/semcerif/1.3#JournalArticleReview
	http://purl.org/eprint/type/SubmittedJournalArticle

Multilingual Labels

C article (en); artículo (es); article (fr); articolo in rivista (it); статья (ru); 文献 (zh)

Terms (20) | Definition (1) | Attribute (0) | Relationship (0) | Alignments (10)

+ Add new term

Language	
English (en)	article (Preferred) W
	paper W
	journal article W
	journal articles W
Español (es)	articulo (Preferred) W
	articulo revisado por pares W
	articulo cientifico W
	articulo cientifico W
Français (fr)	article (Preferred) W
	article scientifique W
	article de recherche W
中文 (zh)	文献 (Preferred) W
	论文 W
	期刊文章 W
	文章 W
	期刊论文 W
Русский (ru)	статья (Preferred) W
Deutsch (de)	Wissenschaftlicher Artikel (Preferred) W
	articolo in rivista (Preferred) W

Some reflections about current version of Resource Types vocabulary

- **The hierarchical model offers many opportunities but also challenges** to grow and implement the vocabulary in a consistent way
- **Mappings with other vocabularies** are pending for some concepts
- **Not all levels of the vocabulary are equally developed**. Some levels are too granular (e.g, "report") while others are incomplete (e.g, "dataset", "software")
- First level: text, dataset, interactive resource, image, workflow, cartographic material, sound, software, other is based on DCMI and DataCite schemas
- **Definitions of concepts in English only**

- "Dataset" particularly requires a more granular representation to become useful to repositories/other infrastructures
- Some concepts refer to types of resources that are quite new for most repositories ("workflow").**Content acceptance policies greatly vary amongst repositories**
- **The diversification of "Research Outputs" today has little to do with the landscape a few years ago and Open Science paves the way to even wider variety**
- **The Editorial Board has collected much useful feedback from the international community: debate in the second half of this meeting**

Access mode vocabulary

- Basis is info:eu-repo
- **The vocabulary indicates the access status of a resource**
- 4 concepts with URIs and labels in some languages of the Working group:
  - **open access**
  - **metadata-only access** (replaces closed access term in info:eu and extends its scope)
  - **embargoed access**
  - **restricted access**
- **Pending:** is this vocabulary the right place to address "green OA" and "gold OA"?
- Send us your feedback!

Definitions in draft v1

- **Open access refers to a resource that is immediately and permanently online, and free for all on the Web, without financial and technical barriers**.
- **Embargoed access refers to a resource that is metadata only access until released for open access on a certain date**. Embargoes can be required by publishers and funders policies, or set by the author (e.g such as in the case of theses and dissertations).
- **Restricted access refers to a resource that is available in a system but with some type of restriction for full open access**. The user must log-in to the system in order to access the resource; The user must send an email to the author or system administrator to access the resource; Access to the resource is restricted to a specific community (e.g. limited to a university community)
- **Metadata only access refers to a resource in which access is limited to metadata only. The resource itself is described by the metadata, but is not directly available through the system or platform**. There is no electronic copy of the resource available (record links to a physical resource); The resource is only available elsewhere for a fee (record links to a subscription-based publisher version); The resource is available open access but at a different location (record links to a version at an open access publisher or archive); The resource is available elsewhere, but not in a fully open access format (record links to a read only, or other type of resources that is not permanent or in some way restricted)

Community outreach

**Join the Editorial Board!
All languages welcome!**

Send us your feedback about the COAR vocabularies

Send suggestions for Working Group FAQs

Check out our Related Resources page

Do you want to share your experience with controlled vocabularies in COAR webinar series?

Keep tuned about upcoming webinars & workshops

- Vocabularios Controlados para Repositorios: Objetivos y Avances del Grupo de Trabajo COAR, Isabel Bernal and Paola Azrilevich, COAR Editorial Board
- Publishing Controlled Vocabularies for Access and Reuse, Rowan Brownlee, ANDS
- Controlled Vocabularies, Jochen Schirrwagen, COAR Editorial Board, and Timo Borst, ZBW

Resource Type 2.0: Public Input & Discussion

Towards "Resource Types" Version 2.0: Some Issues for Discussion

ON THE AGENDA OF THE EDITORIAL BOARD

- Review definitions of concepts. Definition of some concepts may slightly clash with subconcepts beneath
- Include source for all definitions
- Consider translation of definitions into all working group members languages
- Complete mappings with other vocabularies

USUAL COMMENTS RECEIVED

- Publication and format types mixed
- Too many types of "reports"
- Examples of workflows in repositories
- Poor definition of "dataset" and need for developing its inner variety
- More alternative labels for some concepts (e.g, webpage for website)
- Differences between "moving image" and "video" may be unclear

Some proposals for new concepts

- Data Management Plan
- Software paper
- Conference presentation
- Blog post
- Architectural drawing
- Ephemera
- Interview transcripts
- Annotated sources
- Survey responses

- Audiobook
- Honour thesis
- Corrigendum
- Learning Object material
- Manuscript (digitized collections)
- Newspaper article
- ....what else do you think is missing in Version 1.1??

Discussion

Do you think the most important "resource types vocabularies" have been mapped?

COAR Controlled Vocabularies target Open Access repositories, Open Access journal platforms and (to some extent) CRIS (via CERIF-XML). Is there another target group / infrastructure that could benefit?

What could be the benefit(s) and problem(s) of organisin the vocabulary in a hierarchical way

Would you like suggest / rename / remove a concept and why?

What research data terms would you recommend to add (and not specific to a single discipline)?

What issues would you like to see in the Guidelines for repository managers? And in the Guidelines for IT staff?

Next steps for COAR Controlled Vocabularies
Ilkay Holt, COAR

Version Types

Term	info:eu-repo/	Definitions
AO	authorVersion	Author's Original
SMUR	submittedVersion	Submitted Manuscript Under Review
AM	acceptedVersion	Accepted Manuscript
P	N/A	Proof
VoR	publishedVersion	Version of Record
CVoR	updatedVersion	Corrected Version of Record
EVoR	updatedVersion	Enhanced Version of Record
NA	N/A	Not Applicable (or Unknown)

Version Types: what is next?

Consultation with interest group

Finalisation of concepts with editorial group

Translations of concepts and definitions

Deadline => End of the year

Date Types

Term	info:eu-repo/date	Used by
Accepted	accepted	DataCite
Available		DataCite
Copyrighted	copyrighted	DataCite, Bibo
Collected		DataCite
Created		DataCite, Bibo
embargoEnd	embargoEnd	OpenAIRE
Issued	publication	DataCite, Bibo
Submitted	submitted	DataCite, Bibo
Updated		DataCite, Bibo
Valid		Datacite

Under discussion…

Collaboration with other vocabulary services

Phaidra Classification Server

COAR Vocabularies in guidelines

OpenAIRE Literature Repository Guidelines

OpenAIRE CRIS-CERIF Guidelines

LA Referencia Guidelines

DINI Certificate for repositories in Germany

ANDS (Australian National Data Service)

Research Vocabularies Australia

Registry interoperability

Share vocabulary experiences

Collaborate on documentation

Linking services on both web sites

The way forward

Vocabularies

   Access types vocabularies 1.0 release

   Translation of definitions (to be discussed)

   Resource Types 2.0

   Version types

Diversity in the context of languages

Continue collaboration

Use cases

Webinars > trainings > workshop

Thank You

Photo credit Juliette Gibert CC BY SA