



IMPLEMENTATION OF THE COAR VOCABULARIES

in OpenAIRE Guidelines for Literature Repositories and CRIS

Pedro Príncipe & Jochen Schirrwagen



COAR Controlled Vocabulary Workshop, May 8, 2017 - Venezia

OpenAIRE guidelines - What comes next?

OpenAIRE Literature Guidelines v4:

- More granularity and flexibility regarding metadata coverage and extensions
- Keep following established standards
- PIDs for publications, authors & contributors, funder and other organizations
- Update of controlled vocabularies for resource types, subject headings & classifications, access rights, license information
- Location of fulltext files
- ...

OpenAIRE Guidelines for CRIS Managers (with EuroCRIS)

- Enable Application Profiles for CERIF-XML
- Update CERIF Semantic Vocabulary
- Improve feasibility of aggregation of CERIF entities

Building Blocks



- Controlled Vocabularies
- Fulltext Location



- Author Identifiers
- Funder Identifiers



- Bibliographic Citation



Implementation objectives

- COAR Resource Types and Access Rights vocabularies are build on and extend the vocabulary terms from “info:eu-repo”
- COAR Vocabularies replace the way how “info:eu-repo” terms were encoded in the metadata guidelines
 - by linked data compatible machine readable PURLs
 - by human readable labels with multiple language support
- COAR Vocabularies are generated as SKOS RDF-XML Files
 - requires transformation to use them in XML Schema based metadata guidelines

Utilize COAR Vocabularies in XML metadata

- transform concepts from SKOS into XML Schema
 - xs:enumeration elements take the concept URIs as values

`<http://purl.org/coar/resource_type/c_6501>`

```

a          skos:Concept ;
dc-term:modified "2016-09-05T12:00:38+00:00"^^xsd:dateTime ;
skos:altLabel  "articulo científico"@es , "期刊論文"@zh , "journal articles"@en , "article scientifique"@fr , "статья"
skos:broadMatch <http://www.eurocris.org/ontologies/semcerif/1.3#Journal> ;
skos:broader   <http://purl.org/coar/resource_type/c_3e5a> ;
skos:definition <http://purl.org/coar/resource_type/def_01dd> ;
skos:exactMatch <http://www.eurocris.org/ontologies/semcerif/1.3#JournalArticle> , <http://purl.org/info:eu-repo/sem
skos:hiddenLabel <http://purl.org/coar/resource_type/xl_81d0> ;
skos:inScheme   <http://purl.org/coar/resource_type> ;
skos:narrowMatch <http://www.eurocris.org/ontologies/semcerif/1.3#JournalArticleAbstract> , <http://www.eurocris.org/o
skos:narrower   <http://purl.org/coar/resource_type/c_2df8fbb1> , <http://purl.org/coar/resource_type/c_dcae04bc> ;
skos:prefLabel  "articolo in rivista"@it , "wetenschappelijk artikel"@nl , "Wissenschaftlicher Artikel"@de , "article
skos:related    <http://purl.org/eprint/type/JournalItem> , <http://purl.org/eprint/type/SubmittedJournalArticle> ;
skos:relatedMatch <http://purl.org/eprint/type/JournalItem> , <http://purl.org/eprint/type/SubmittedJournalArticle> ;

```

`/schema/v4" targetNamespace="http://openaire.eu/schema/v4" ele`

`RI from the COAR resource type vocabulary</xs:documentation>`

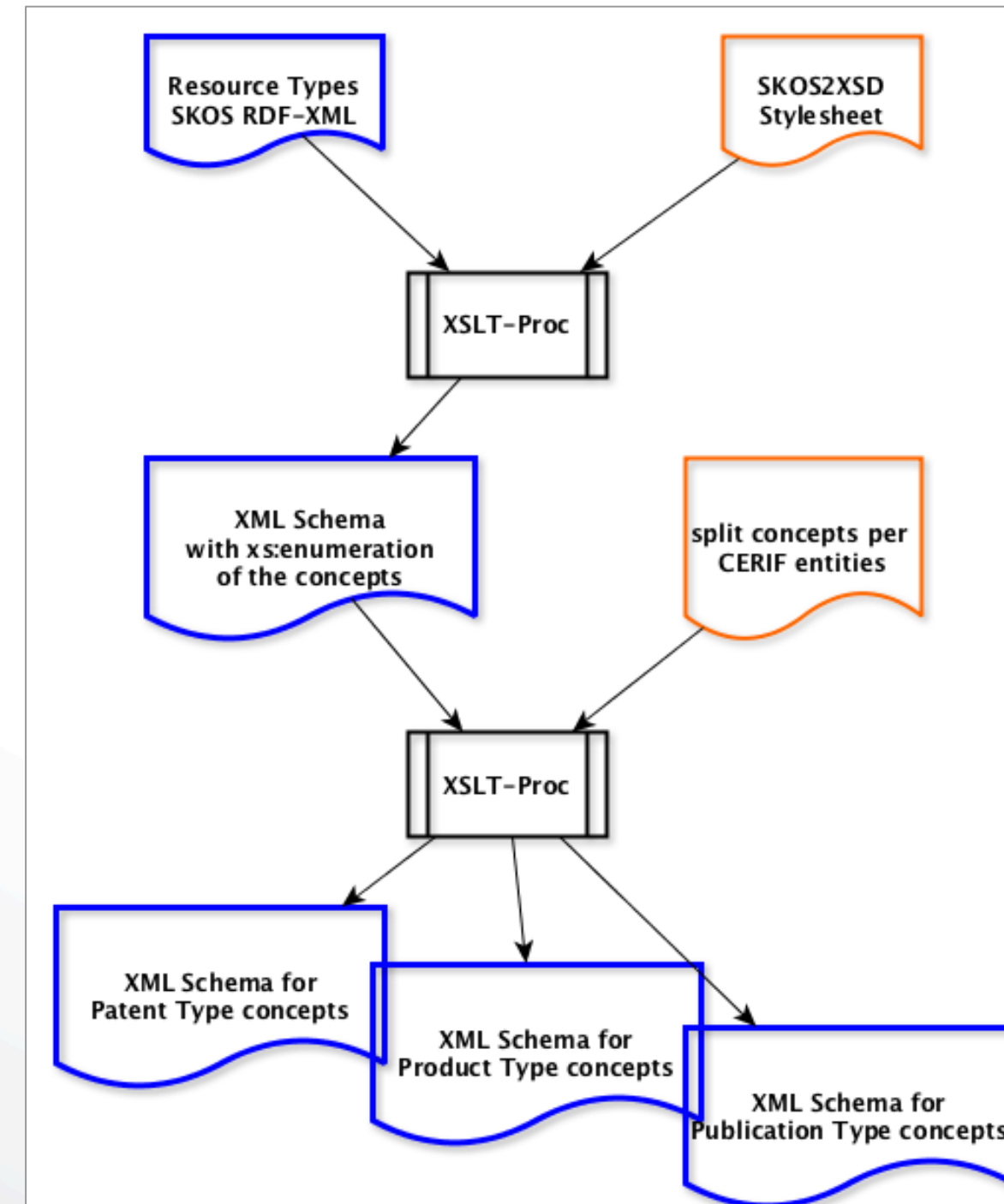
```

8      <xs:enumeration value="http://purl.org/coar/resource_type/c_1102"/><!--annotation-->
9      <xs:enumeration value="http://purl.org/coar/resource_type/c_0640"/><!--journal-->
10     <xs:enumeration value="http://purl.org/coar/resource_type/c_6501"/><!--journal article-->
11     <xs:enumeration value="http://purl.org/coar/resource_type/c_b239"/><!--editorial-->
12     <xs:enumeration value="http://purl.org/coar/resource_type/c_7a1f"/><!--bachelor thesis-->
13     <xs:enumeration value="http://purl.org/coar/resource_type/c_86bc"/><!--bibliography-->
14     <xs:enumeration value="http://purl.org/coar/resource_type/c_2f33"/><!--book-->
15     <xs:enumeration value="http://purl.org/coar/resource_type/c_3248"/><!--book part-->
16     <xs:enumeration value="http://purl.org/coar/resource_type/c_ba08"/><!--book review-->
17     <xs:enumeration value="http://purl.org/coar/resource_type/c_7ad9"/><!--website-->

```

SKOS -> XML-Schema pipeline for the CRIS Guidelines

1. generate XML schema from the SKOS representation
2. (optionally, but required for CRIS-CERIF) split the Resource Type concepts to have separated types for the CERIF-entities
3. Publication, Patent and Product



How to encode the concepts in the metadata fields?

```
25 <dc:title xml:lang="en">A title</dc:title>
26 <dc:description xml:lang="en">A description</dc:description>
27 <dc:language>eng</dc:language>
28 <dc:date>2016-11-11</dc:date>
29 <dc:type rdf:resource="http://purl.org/coar/resource_type/c_6501">journal article</dc:type>
30 <dc:format>application-pdf</dc:format>
31 <dc:identifier>http://example.org/test/article1</dc:identifier>
32 <dc:source>in print</dc:source>
33 <dc:subject>metadata guidelines</dc:subject>
34 <dc:publisher>OpenAIRE Guidelines team</dc:publisher>
35 <dc:rights rdf:resource="http://creativecommons.org/licenses/by/4.0/">CC-BY 4.0</dc:rights>
```

- provide pairs of the concept-URI and corresponding labels in desired languages
- *work in progress:*
 - the CERIF-CRIS Guidelines: <https://github.com/openaire/guidelines-cris-managers>
 - the Literature Guidelines: https://gitlab.ub.uni-bielefeld.de/openaire/lit_guidelines_v4_preview

Application Profile Preview

OpenAIRE Guidelines
literature_guidelines_v4_draft

Search docs

OpenAIRE Guidelines for Literature Repositories

Introduction
Use of OAI-PMH
Use of OAI-DC
Application Profile Overview
1. Title (M)
2. Creator (M)
3. Contributor (R)
4. FundingReference (MA)
5. AlternateIdentifier (R)
6. RelatedIdentifier (MA)
7. Embargo Period Date (M)
8. Language (R)
9. Publisher (MA)
10. Publication Date (M)
11. Publication Type (M)
12. Description (MA)
13. Format (R)
14 Resource Identifier (M)

Read the Docs
v: literature_guidelines_v4_draft

Docs » OpenAIRE Guidelines for Literature Repositories » Application Profile Overview

Application Profile Overview

This documentation uses the following namespace abbreviation:

- dc** : <http://purl.org/dc/elements/1.1/>
- datacite** : <http://datacite.org/schema/kernel-4>
- rdf** : <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
- aire** : <http://openaire.eu/schema/v4>

OpenAIRE-Field	OAI-DC Element	Refinement by vocabulary
1. Title (M)	dc:title	
2. Creator (M)	datacite:creator	
3. Contributor (R)	datacite:contributor	DataCite contributor type
4. FundingReference (MA)	datacite:fundingReference	DataCite funderIdentifier type
5. AlternateIdentifier (R)	datacite:alternateIdentifier	
6. RelatedIdentifier (MA)	datacite:relatedIdentifier	DataCite relatedIdentifier type
7. Embargo Period Date (M)	datacite:date	DataCite dateType
8. Language (R)	dc:language	IETF BCP 47, ISO 639-1,2,3
9. Publisher (MA)	dc:publisher	
10. Publication Date (M)	dc:date	

Example

```

11 <datacite:creator>
12   <datacite:creatorName>Lastname, Firstname</datacite:creatorName>
13   <datacite:nameIdentifier nameIdentifierScheme="ORCID" schemeURI="http://orcid.org/">0000-0001-5000-0007</datacite:nameIdentifier>
14 </datacite:creator>
15 <datacite:contributor contributorType="HostingInstitution">
16   <datacite:contributorName>OpenAIRE</datacite:contributorName>
17 </datacite:contributor>
18 <datacite:fundingReference>
19   <datacite:funderName>European Commission</datacite:funderName>
20   <datacite:funderIdentifier funderIdentifierType="Crossref Funder ID">http://doi.org/10.13039/501100000780</datacite:funderIdentifier>
21   <datacite:awardNumber awardURI="http://cordis.europa.eu/project/rcn/100180_en.html">123456</datacite:awardNumber>
22   <datacite:awardTitle>OpenAIRE Guidelines Project</datacite:awardTitle>
23 </datacite:fundingReference>
24 <datacite:relatedIdentifier relatedIdentifierType="URL" relationType="Continues">https://guidelines.openaire.eu/en/latest/literature/index.html</datacite:relatedIdentifier>
25 <dc:title xml:lang="en">A title</dc:title>
26 <dc:description xml:lang="en">A description</dc:description>
27 <dc:language>eng</dc:language>
28 <dc:date>2016-11-11</dc:date>
29 <dc:type rdf:resource="http://purl.org/coar/resource_type/c_6501">journal article</dc:type>
30 <dc:format>application-pdf</dc:format>
31 <dc:identifier>http://example.org/test/article1</dc:identifier>
32 <dc:source>in print</dc:source>
33 <dc:subject>metadata guidelines</dc:subject>
34 <dc:publisher>OpenAIRE Guidelines team</dc:publisher>
35 <dc:rights rdf:resource="http://creativecommons.org/licenses/by/4.0/">CC-BY 4.0</dc:rights>
36 <file accessRights="open" mimeType="application-pdf" objectType="fulltext">http://example.org/test/article1/download/pdf</file>
37 <version>draft</version>

```

Thanks

Pedro Príncipe and Jochen Schirrwagen (OpenAIRE), Jan Dvorak (EuroCRIS)



www.openaire.eu



@openaire_eu



facebook.com/groups/openaire



linkedin.com/groups/OpenAIRE-3893548



pedroprincipe@sdum.uminho.pt