

Jochen Schirrwagen
Bielefeld University

OpenAIRE Guidelines

Application Profiles

METADATA ELEMENTS and CONTROLLED VOCABULARIES for REPOSITORIES
Hamburg | 16 May 2018

OpenAIRE's e-infrastructure Commons & Contentprovider

- 24 M documents
- > 1000 compatible content provider

Compatibility Status

401: OpenAIRE 3.0
96: OpenAIRE 2.0+
29: OpenAIRE 2.0
514: OpenAIRE Basic
21: OpenAIREdata 2.0

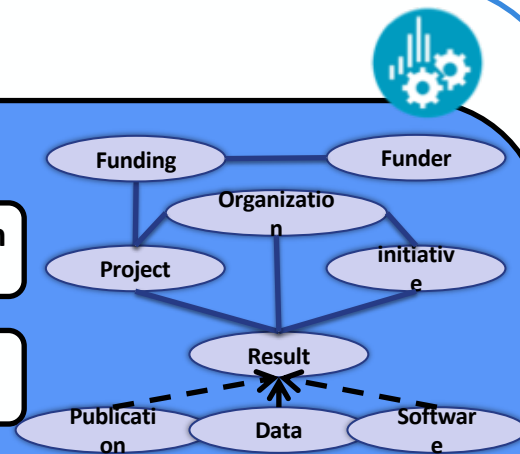
INFO SPACE SERVICES

Cleaning

De-duplication

Validation

Enrichment
By inference



CONTENT PROVIDERS



Publications
repositories

Research Data
repositories

CRIS
systems

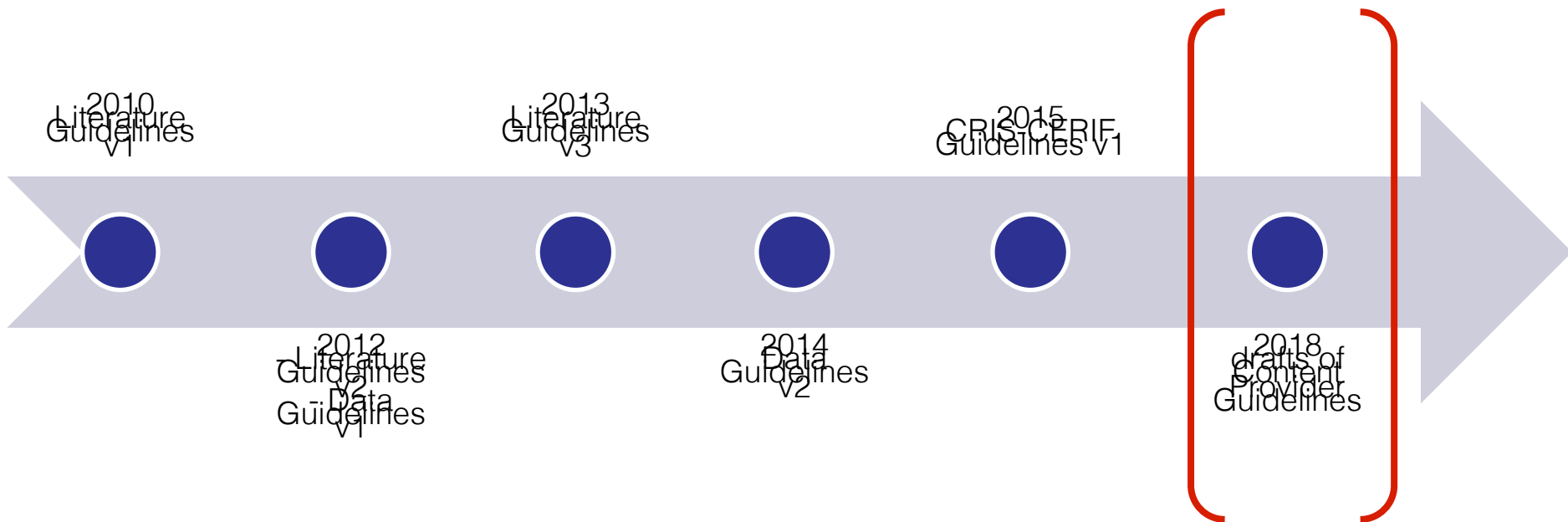
Registries
(e.g. projects)

OA
Journals

Software
Repositories



Evolution of OpenAIRE-Guidelines



FROM Guidelines for Data Providers TO Guidelines for Open Science Content Providers

RESEARCH LITERATURE

Thematic Repositories
Institutional Repositories
E-journals

RESEARCH SOFTWARE

Software Repositories
Catch-all Repositories

RESEARCH DATA

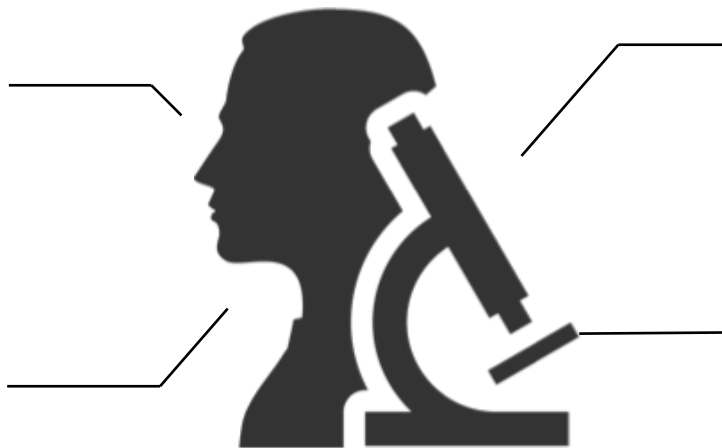
Data Repositories
Catch-all Repositories
Institutional & thematic
repositories

Other RESEARCH Products

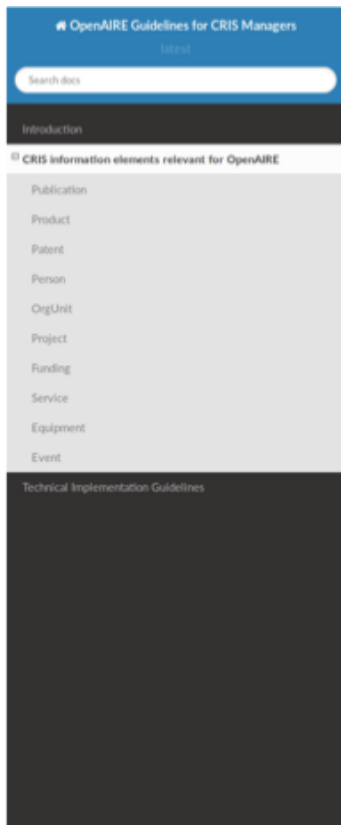
Catch-all Repositories

RESEARCH INFORMATION

CRIS

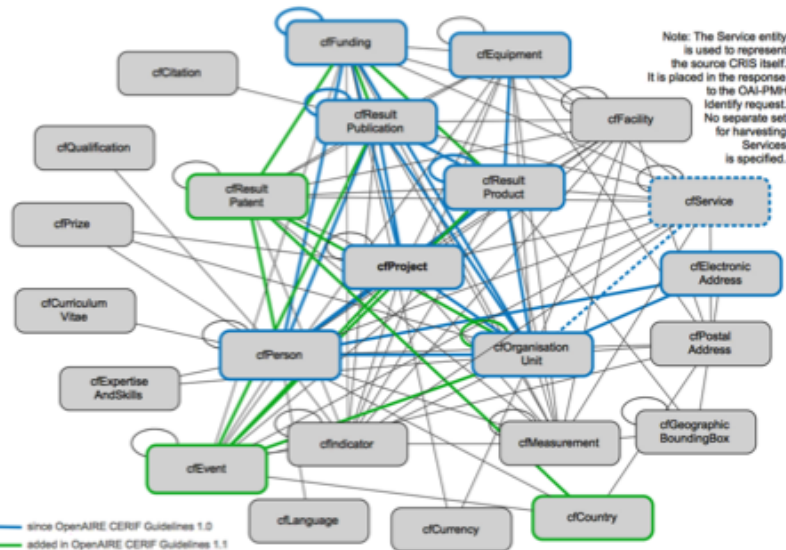


Draft OpenAIRE Guidelines for CRIS Managers v1.1



CRIS information elements relevant for OpenAIRE

CERIF is a comprehensive model for the research domain. Only a subset of that information is relevant for OpenAIRE. For example, Publications, Products (including research datasets and software), and Events fall in the scope of the OpenAIRE profile, while Prizes and Geographic bounding boxes do not. The following diagram shows the part of the CERIF model is being used in this version of the OpenAIRE Guidelines:



Draft Guidelines for Software Repository Manager

[Docs](#) » OpenAIRE Guidelines for Software Repository Managers

[Edit on GitHub](#)

OpenAIRE Guidelines for Software Repository Managers

These guidelines describe the application profile for Software Repository managers to be compatible with OpenAIRE.

- [Introduction](#)
 - [Aim](#)
 - [OpenAIRE \(Open Access Infrastructure for Research in Europe\)](#)
 - [Rationale](#)
 - [Acknowledgements & Contributors](#)
- [Application Profile Overview](#)
- [Acknowledgements & Contributors](#)

Next ➔

© Copyright 2017, OpenAIRE. This work is licensed under Creative Commons Attribution 4.0 International. Revision 93b621b3.



Application Profile Overview

This documentation uses the following namespace abbreviation:

- `datacite` : <http://datacite.org/schema/kernel-4>

OpenAIRE-Field	Metadata Element
1. Identifier (M)	datacite:identifier
2. Author (M)	datacite:creator
3. Contact Person (O)	datacite:contributor
4. Contact Group (O)	datacite:contributor
5. Name (M)	datacite:title
6. Repository (R)	datacite:publisher
7. Software Type (M)	datacite:resourceType
8. Subject (R)	datacite:subject
9. Release date (R)	datacite:date
10. Description language (O)	datacite:language
11. Alternate identifier (O)	datacite:alternateIdentifier
12. Landing page (R)	datacite:alternateIdentifier
13. Distribution location (R)	datacite:alternateIdentifier
14. Documentation (R)	datacite:relatedIdentifier
15. RelatedIdentifier (O)	datacite:relatedIdentifier
16. Programming Language (R)	datacite:format
17. Version Number (R)	datacite:version
18. Access Rights (M)	datacite:rights
19. Licence Condition (R)	datacite:rights
20. Description (R)	datacite:description
21. Tool (R)	datacite:description
22. Funding Reference (R)	datacite:fundingReference
23. Distribution form (O)	datacite:description
24. Development status (O)	datacite:description

Draft Guidelines for “Other Research Products”

[Docs](#) » OpenAIRE Guidelines for Other Research Products (ORP) Repository Managers

[Edit on GitHub](#)

OpenAIRE Guidelines for Other Research Products (ORP) Repository Managers

These guidelines describe the application profile for ORPs Repository managers to be compatible with OpenAIRE.

- [Introduction](#)
- [Application Profile Overview](#)

Next ➞

© Copyright 2017, OpenAIRE. This work is licensed under Creative Commons Attribution 4.0 International. Revision b0f665c3.

Built with [Sphinx](#) using a [theme](#) provided by [Read the Docs](#).

Application Profile Overview

This documentation uses the following namespace abbreviation:

- `datacite` : <http://datacite.org/schema/kernel-4>

OpenAIRE-Field	Metadata Element
1. Identifier (M)	datacite:identifier
2. Author (M)	datacite:creator
3. Contact Person (O)	datacite:contributor
4. Contact Group (O)	datacite:contributor
5. Name (M)	datacite:title
6. Repository (R)	datacite:publisher
7. Type (M)	datacite:resourceType
8. Subject (R)	datacite:subject
9. Release date (R)	datacite:date
10. Description language (O)	datacite:language
11. Alternate Identifier (O)	datacite:alternateIdentifier
12. Landing page (R)	datacite:alternateIdentifier
13. Distribution location (R)	datacite:alternateIdentifier
14. Documentation (R)	datacite:relatedIdentifier
15. RelatedIdentifier (R)	datacite:relatedIdentifier
16. Format (O)	datacite:format
17. Version Number (R)	datacite:version
18. Access Rights (M)	datacite:rights
19. Licence Condition (R)	datacite:rights
20. Description (R)	datacite:description
21. Tool (R)	datacite:description
22. Funding Reference (R)	datacite:fundingReference

OpenAIRE Guidelines for Literature Repository Managers v4

Identified needs to upgrade from different perspectives

- **Funder perspective**

- Link funding information with research results

- **Author and Reader perspective**

- Link authors and contributors with their research output and ease name disambiguation

- **Service provider perspective**

- Avoid overloading of oai_dc metadata
- Make maintenance and mappings of controlled vocabularies easier by help of identifiers
- Make identification of resources easier (e.g. for TDM)
- Improve alignment with other regional repository networks
 - **Agree on a shared set of metadata properties and controlled vocabularies**
 - **Examples: LA Referencia, JAIRO**

Draft OpenAIRE Guidelines for Literature Repository Managers v4

...

APPLICATION PROFILE OVERVIEW

OpenAIRE-Field	Metadata 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 (R)	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	
11. Publication Type (M)	dc:type	COAR Resource Type Vocabulary
12. Description (MA)	dc:description	
13. Format (R)	dc:format	
14 Resource Identifier (M)	dc:identifier	
15. Access Rights (M)	dc:rights	COAR Access Right Vocabulary
16. Source (R)	dc:source	
17. Subject (MA)	dc:subject	
18. License Condition (R)	dc:rights	
19. Relation (O)	dc:relation	
20. Coverage (R)	dc:coverage	
21. Publication Version (R)	aire:version	
22. Fulltext Location (M)	aire:file	

Approach

- **Build an application profile based on established and widely used metadata schemes in repositories**
 - [Dublin Core and DataCite v4.1](#)
- **Allow for additional properties when needed**
- **Align with other repository networks**

Metadata Property about author and contributor identifier

- Support the provision of author identifiers in metadata

```
<datacite:creator>  
  <datacite:creatorName>Wallentin, Carl-Johan  
  </datacite:creatorName>  
  <datacite:nameIdentifier  
    nameIdentifierScheme="ORCID"  
    schemeURI="http://orcid.org">  
    http://orcid.org/0000-0003-1983-9378  
  </datacite:nameIdentifier>  
</datacite:creator>
```

Metadata Property about funding information

- Comply with funder mandates
- Make use of CrossRef Funding Data
- In some cases we may need an attribute „fundingstream“

```
<datacite:fundingReference>
  <datacite:funderName>European Commission</datacite:funderName>
  <fundingStream>H2020 Marie Skłodowska-Curie Actions</fundingStream>
  <datacite:funderIdentifier
    funderIdentifierType="Crossref Funder ID">
  </datacite:funderIdentifier>
  <datacite:awardNumber
    awardURI="http://cordis.europa.eu/project/rcn/195983_en.html">660668
  </datacite:awardNumber>
  <datacite:awardTitle>ACT against AMR</datacite:awardTitle>
</datacite:fundingReference>
```

Metadata Property about License and Access rights

- Access rights – how can I access the resource
- License – what am I allowed to do with the resource
- Use URIs whenever possible

```
<dc:rights
  rdf:resource="http://purl.org/coar/access_right/c_abf2">
  open access
</dc:rights>
<dc:rights
  rdf:resource="http://creativecommons.org/licenses/by-nc/4.0/">
  Creative Commons Attribution-NonCommercial
</dc:rights>
```

Metadata Property about Fulltext Location

- URL of one or more fulltext files
- Attributes about access right, mimetype, objecttype

```
<file  
  accessRights=http://purl.org/coar/access\_right/c\_abf2  
  mimeType="application/pdf">  
    http://europepmc.org/articles/PMC5574022?pdf=render  
</file>
```

Controlled Vocabularies

- **Re-use and adaptation of controlled values used in the DataCite schema**
 - E.g. identifier types, role types, relation types
- **Controlled Vocabularies defined by the COAR community**
 - E.g. resource types, access rights, version types
- **Controlled Vocabularies defined in OpenMinTeD Guidelines**
 - E.g. licenses

COAR Resource Types and Access Rights



- **General aspects**

- Unique identification of vocabulary concepts
- Improved granularity
- Multilingual Support
- Implemented in SKOS

- **Resource Types**

- http://vocabularies.coar-repositories.org/documentation/resource_types/

- **Access Rights**

- http://vocabularies.coar-repositories.org/documentation/access_rights/



Sample record

```
<resource xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:datacite="http://datacite.org/schema/kernel-4"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:vc="http://www.w3.org/2007/XMLSchema-versioning"
xmlns="http://namespace.openaire.eu/schema/oaire/"
xsi:schemaLocation="http://namespace.openaire.eu/schema/oaire/ ../schemas/openaire-v4.xsd">
  <datacite:creator>
    <datacite:creatorName>Wallentin, Carl-Johan</datacite:creatorName>
    <datacite:nameIdentifier nameIdentifierScheme="ORCID" schemeURI="">http://orcid.org/0000-0003-1983-9378</datacite:nameIdentifier>
  </datacite:creator>
  <datacite:fundingReference>
    <datacite:funderName>European Commission</datacite:funderName>
    <fundingStream>H2020 Marie Skłodowska-Curie Actions</fundingStream>
    <datacite:funderIdentifier funderIdentifierType="Crossref Funder ID"></datacite:funderIdentifier>
    <datacite:awardNumber awardURI="http://cordis.europa.eu/project/rcn/195983_en.html">660668</datacite:awardNumber>
    <datacite:awardTitle>ACT against AMR</datacite:awardTitle>
  </datacite:fundingReference>
  <datacite:alternateIdentifier alternateIdentifierType="DOI">10.1002/chem.201701589</datacite:alternateIdentifier>
  <datacite:relatedIdentifier relatedIdentifierType="EISSN" relationType="IsPartOf">1521-3765</datacite:relatedIdentifier>
  <dc:title xml:lang="eng">Redox-Neutral Dual Functionalization of Electron-Deficient Alkenes</dc:title>
  <dc:description xml:lang="eng">Visible-light photoredox catalysis ...</dc:description>
  <dc:language>eng</dc:language>
  <dc:date>2017-05-10</dc:date>
  <dc:type rdf:resource="http://purl.org/coar/resource_type/c_6501">journal article</dc:type>
  <dc:identifier>http://europepmc.org/articles/PMC5574022</dc:identifier>
  <dc:publisher>John Wiley and Sons Inc.</dc:publisher>
  <dc:rights rdf:resource="http://purl.org/coar/access_right/c_abf2">open access</dc:rights>
  <dc:rights rdf:resource="http://creativecommons.org/licenses/by-nc/4.0/">Creative Commons Attribution-NonCommercial</dc:rights>
  <dc:subject>acyl radicals</dc:subject>
  <dc:subject>carboxylic acids</dc:subject>
  <dc:subject>cascade transformation</dc:subject>
  <dc:subject>multicomponent reactions</dc:subject>
  <dc:subject>photoredox catalysis</dc:subject>
  <file accessRights="http://purl.org/coar/access_right/c_abf2" mimeType="application/pdf" objectType="">http://europepmc.org/articles/PMC5574022?pdf=render</file>
```



<https://www.openaire.eu>

<https://guidelines.openaire.eu>



[@openaire_eu](https://twitter.com/openaire_eu)



facebook.com/groups/openaire



jochen.schirrwagen@uni-bielefeld.de