OpenAIRE Interoperability Guidelines for institutional thematic Repository Managers

Release 4.1-SNAPSHOT

OpenAIRE Guidelines Team

Feb 16, 2022

Contents

CHAPTER 1

Introduction

1.1 Aim

The OpenAIRE Guidelines for institutional and thematic Repository Managers 4.1 provide orientation for repository managers to define and implement their local data management policies according to the requirements of the OpenAIRE - Open Access Infrastructure for Research in Europe[1].

The OpenAIRE Guidelines were established to support the Open Access strategy of the European Commission[2] and to meet requirements of the OpenAIRE infrastructure. This new version of the Guidelines, according to the expansion of the aims of the OpenAIRE initiative and its infrastructure, has a broader scope. In fact, these Guidelines are intended to guide repository manager to expose to the OpenAIRE infrastructure **open access and non-open access publications** together with **funding information**, where applicable.

By implementing these Guidelines, repository managers will not only be enabling authors who deposit publications in their repository to fulfill the European Commission (EC) Open Access requirements, and eventually also the requirements of other (national or international) funders with whom OpenAIRE cooperates, but also incorporating their publications into the OpenAIRE infrastructure for discoverability and utilizing value-added services provided by the OpenAIRE portal.

The OpenAIRE Guidelines for institutional and thematic Repository Managers 4.0 are part of a set of OpenAIRE Guidelines that also include the OpenAIRE Guidelines for Data Archive Managers, the OpenAIRE Guidelines for CRIS managers, the OpenAIRE Guidelines for Software Repository Managers, and the Guidelines for Other Research Products Repository Managers.

1.2 What's new

In comparison with previous versions of the Guidelines, this version introduces the following major changes:

since v4.0:

- covering of FAIR principles elements
- proof of concept: enhancement of *provenance* information of with **repositoryId's** and **repositoryName's** for Aggregators

[1] http://www.openaire.eu
[2] http://ec.europa.eu/research/openscience/index.cfm?pg=openaccess

- add new value to controlled `relationType` vocabulary: **IsPublishedIn** and is adopted from DataCite Schema v4.4.
- Note on OAI-PMH repository *batch_size*

since v3.0:

- covering of FAIR principles elements
- proof of concept: enhancement of provenance information of with repositoryId's and repositoryName's for Aggregators
- an application profile and schema based on Dublin Core and DataCite incl. a new OAI-metadataPrefix
- support of identifier schemes for authors, organizations, funders, scholarly resources
- introduction of COAR Controlled Vocabularies
- compliance with the OpenAIRE Content Acquisition Policy[3], published on 05-Oct-2018.

1.3 How this document is structured

Chapter two provides a brief overview of how to configure and use OAI-PMH for OpenAIRE metadata harvesting. Chapter three describes the application profile. It assignes properties from Dublin Core and DataCite metadata schemes to OpenAIRE fields. Each OpenAIRE field is described in detail by

- the name of the field
- how it is mapped to an element in such metadata schemes
- the cardinality of the field
- definition and usage instructions with regard to allowed values in properties, sub-properties and attributes
- example(s)

1.4 Acknowledgments & Contributors

Editors

- Jochen Schirrwagen[4] (Bielefeld University, Germany)
- Miriam Baglioni[5] (CNR, Italy)
- Andreas Czerniak[6] (Bielefeld University, Germany)

Experts & Reviewers

OpenAIRE

- Pedro Principe (University of Minho, Portugal)
- Paolo Manghi (CNR, Italy)
- Aenne Loehden (Bielefeld University, Germany)
- Amelie Bäcker (Bielefeld University, Germany)
- Alessia Bardi (CNR, Italy)

LA Referencia

- Alberto Cabezas (LA Referencia, Chile)

[3] https://doi.org/10.5281/zenodo.1446407
[4] https://orcid.org/0000-0002-0458-1004
[5] https://orcid.org/0000-0002-2273-9004
[6] https://orcid.org/0000-0003-3883-4169

- Paola Azrilevich (MinCyT, Argentina)

JPCOAR

- Tomoko Kataoka (Ochanomizu University, Japan)

Antleaf

- Paul Walk (Antleaf, UK)

University of Alberta Libraries

- Sharon Farnel (University of Alberta Libraries, CA)

CARL Open Repositories Working Group

- Pierre Lasou (Université Laval, CA)

1.5 Versions

- 4.1-rc, February 2021
  - Corrections, refined examples, enhance provenance element, and adding FAIR descriptions
- 4.0, November 2018 doi:10.5281/zenodo.1299203[7]
- 4.0 draft, November 2017
- 3.0, April 2013 doi:10.5281/zenodo.1487968[8]
- 3.0, beta December 2012
  - The OpenAIRE OAI set has been renamed from `ec_fundedresources` to `openaire`.
  - New relation elements for indicating external identifiers, references and connections to datasets.
- 2.0, October 2012 doi:10.5281/zenodo.59208[9]
  - Compatibility for aggregators; extended Namespace for Project Identification
- 1.1, November 2010 doi:10.5281/zenodo.59206[10]
  - Correction of names and references; addition of license and version statement
- 1.0, July 2010 doi:10.5281/zenodo.59204[11]
  - Initial document

[7] http://dx.doi.org/10.5281/zenodo.1299203
[8] http://dx.doi.org/10.5281/zenodo.1487968
[9] http://dx.doi.org/10.5281/zenodo.59208
[10] http://dx.doi.org/10.5281/zenodo.59206
[11] http://dx.doi.org/10.5281/zenodo.59204

CHAPTER 2

Use of OAI-PMH

OpenAIRE supports a number of transfer protocols and interfaces for collecting bibliographic metadata. The usage of one of the major protocols, OAI-PMH v2.0 protocol[12], in the context of these Guidelines and its application profile is described below.

2.1 Metadata Format

OpenAIRE expects metadata to be encoded following the metadata format defined in the OpenAIRE Application Profile. The recommended metadataPrefix is `oai_openaire`. For information on how to use the individual properties, please refer to the section *Application Profile Overview*.

2.2 Metadata Content

OpenAIRE collects metadata of scientific products according to the OpenAIRE Content Acquisition Policy published at https://doi.org/10.5281/zenodo.1446407 . This includes bibliographic metadata describing open access and non-open access items.

2.3 Harvesting Batch_Size

The common convention for the harvesting *batch_size* via OAI-PMH is '100' records per request [Open Archives - OAI Flow Control[13]]. If more records are available beyond that first page with *batch_size* records, a "resumptionToken" is presented. OpenAIRE recommendation is to have a *batch_size* between **100 and 500** records per request.

A higher value of *batch_size* would be desirable if the requirements are in place.

[12] http://www.openarchives.org/OAI/openarchivesprotocol.html
[13] http://www.openarchives.org/OAI/openarchivesprotocol.html#FlowControl

2.4 Compatibility of Aggregators

Besides individual repositories and journals, also aggregators (e.g., on the national level) can become OpenAIRE compatible. In this case, additional provenance information on the original content providers harvested by such an aggregator has to be encoded for OpenAIRE on the metadata record level. In accordance with the OAI-PMH provenance guidelines[14], the provenance information has to be provided in the `about` node element of an OAI record, as displayed in the following example:

```
<about>
  <provenance xmlns="http://www.openarchives.org/OAI/2.0/provenance"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/provenance
  http://www.openarchives.org/OAI/2.0/provenance.xsd">
    <originDescription altered="true" harvestDate="2012-09-17T14:58:36Z">
      <baseURL>http://dspace.library.uu.nl:8080/dspace-oai/request</baseURL>
      <identifier>oai:dspace.library.uu.nl:1874/218065</identifier>
      <datestamp>2012-01-19T12:38:56Z</datestamp>
      <metadataNamespace>
        http://www.openarchives.org/OAI/2.0/oai_dc/
      </metadataNamespace>
      <repositoryId>opendoar:98765</repositoryId>
      <repositoryName>DSpace Library</repositoryName>
     </originDescription>
  </provenance>
</about>
```

This means information encoded in the following elements is expected (taken from https://www.openarchives.org/OAI/2.0/guidelines-provenance.htm):

- baseURL - the baseURL of the originating repository from which the metadata record was harvested.
- identifier - the unique identifier of the item in the originating repository from which the metadata record was disseminated.
- datestamp - the datestamp of the metadata record disseminated by the originating repository.
- metadataNamespace - the XML namespace URI of the metadata format of the record harvested from the originating repository.
- **originDescription - an optional originDescription block which was that obtained when the metadata record was harve**
  originDescription blocks will describe provenance over a sequence of harvests.

Each *originDescription* must also have the following two attributes which relate to the act of harvesting and any subsequent processing:

- harvestDate - the responseDate of the OAI-PMH response that resulted in the record being harvested from the originating repository.
- altered - a boolean value which must be true if the harvested record was altered before being disseminated again.

As an PoC extension for the aggregator community, we wanted to make the proposal for two more elements to reflect the non-persistent identifier of the OAI-PMH protocol for a repsoitory:

- repositoryId - structured with registry with controlled vocabulary: opendoar[15], re3data[16], doaj[17], and dris[18] and followed by the registryId (without precending zeros), eg. `dris:98765`.
- repositoryName - human readable name of the repository.

[14] http://www.openarchives.org/OAI/2.0/guidelines-provenance.htm
[15] https://https://v2.sherpa.ac.uk/opendoar/
[16] https://re3data.org
[17] https://doaj.org
[18] https://dspacecris.eurocris.org/cris/explore/dris

CHAPTER 3

Application Profile Overview

The properties of the Application Profile for OpenAIRE institutional and thematic Repository Guidelines are listed in this section. The following requirement levels for the metadata properties are used:

Mandatory (M) The property must always be present in the metadata. An empty value for the property is not allowed.

Mandatory if Applicable (MA) When the property value can be obtained it must be present in the metadata

Recommended (R) The use of the property is recommended

Optional (O) It is not important whether the property is used or not, but if used it may provide complementary information about the resource

This documentation uses the following namespace abbreviations:

- `dc`: http://purl.org/dc/elements/1.1/
- `dcterms`: http://purl.org/dc/terms/
- `datacite`: http://datacite.org/schema/kernel-4
- `oaire`: http://namespace.openaire.eu/schema/oaire/

OpenAIRE-Field	Metadata Element	Refinement by Vocabulary	FAIR
Title (M)	datacite:title	title type	
Creator (MA)	datacite:creator	name type	RDA-I3-01M
Contributor (MA)	datacite:contributor	name type contributor type	
Funding Reference (MA)	oaire:fundingReference	funderIdentifier type	
Alternate Identifier (R)	datacite:alternateIdentifier	alternateIdentifier type	RDA-F1-02M

Continued on next page

Table 1 – continued from previous page

OpenAIRE-Field	Metadata Element	Refinement by Vocabulary	FAIR
Related Identifier (R)	datacite:relatedIdentifier	relatedIdentifier type relation type resourcetype general	RDA-I3-01M RDA-I3-02M RDA-I3-03M RDA-I3-04M
Embargo Period Date (MA)	datacite:date	date type	
Language (MA)	dc:language	IETF BCP 47[19], ISO 639-3[20]	
Publisher (MA)	dc:publisher		
Publication Date (M)	datacite:date	date type	
Resource Type (M)	oaire:resourceType	COAR Resource Type Vocabulary[21]	
Description (MA)	dc:description		
Format (R)	dc:format		RDA-I1-01D
Resource Identifier (M)	datacite:identifier	identifier type	RDA-F1-01M RDA-F1-02M RDA-A1.1-01D
Access Rights (M)	datacite:rights	COAR Access Right Vocabulary[22]	RDA-A1-01M
Source (R)	dc:source		
Subject (MA)	datacite:subject		RDA-I1-01M
License Condition (R)	oaire:licenseCondition		RDA-R1.1-01M RDA-R1.1-02M RDA-R1.1-03M
Coverage (R)	dc:coverage		
Size (O)	datacite:size		
Geo Location (O)	datacite:geoLocation		
Resource Version (R)	oaire:version	COAR Version Vocabulary[23]	
File Location (MA)	oaire:file	COAR Access Right Vocabulary[24]	RDA-F3-01M
Citation Title (R)	oaire:citationTitle		
Citation Volume (R)	oaire:citationVolume		
Citation Issue (R)	oaire:citationIssue		
Citation Start Page (R)	oaire:citationStartPage		
Citation End Page (R)	oaire:citationEndPage		
Citation Edition (R)	oaire:citationEdition		
Citation Conference Place (R)	oaire:citationConferencePlace		
Citation Conference Date (R)	oaire:citationConferenceDate		
Audience (O)	dcterms:audience		

The application profile is implemented in XML Schema. The files[25] for the application profile and sample XML files[26] are part of these Guidelines and also available on the GitHub repository[27].

In the XML metadata documents the schema must be declared as follows:

```
<oaire:resource xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:datacite="http://datacite.org/schema/kernel-4"
  xmlns:oaire="http://namespace.openaire.eu/schema/oaire/"
  xsi:schemaLocation="http://namespace.openaire.eu/schema/oaire/ https://www.
↪openaire.eu/schema/repo-lit/4.0/openaire.xsd">
```

3.1 Title (M)

`datacite:title`

3.1.1 Cardinality

Mandatory

Occurrence: 1-n

3.1.2 Definition and Usage Instruction

A name or title by which a resource is known.

Allowed values, examples, other constraints

Free text.

Remarks

- adapted from DataCite MetadataKernel[28] v4.1

Property title (M, 1-n)

Use the title name as value. Repeat this property for different title types or title languages.

Attribute lang (O)

The language of the title (occurrence: 0-1)

Use the `xml:lang` attribute to indicate the language of the title. The value of the attribute should be chosen from IETF BCP 47, the IANA Language Subtag Registry[29].

[19] http://tools.ietf.org/rfc/bcp/bcp47.txt
[20] https://iso639-3.sil.org/
[21] http://vocabularies.coar-repositories.org/documentation/resource_types/
[22] http://vocabularies.coar-repositories.org/documentation/access_rights/
[23] http://vocabularies.coar-repositories.org/documentation/version_types/
[24] http://vocabularies.coar-repositories.org/documentation/access_rights/
[25] https://github.com/openaire/guidelines-literature-repositories/tree/master/schemas
[26] https://github.com/openaire/guidelines-literature-repositories/tree/master/samples
[27] https://github.com/openaire/guidelines-literature-repositories
[28] https://schema.datacite.org/meta/kernel-4.4/
[29] http://www.iana.org/assignments/language-subtag-registry

Attribute titleType (O)

The type of Title (occurrences: 0-1).

Allowed values, examples, other constraints

Controlled List Values:

- AlternativeTitle
- Subtitle
- TranslatedTitle
- Other

3.1.3 Example

```
<datacite:titles>
  <datacite:title xml:lang="en-US">
    National Institute for Environmental Studies and Center
    for Climate System Research Japan
  </datacite:title>
  <datacite:title xml:lang="en-US" titleType="Subtitle">A survey</datacite:title>
</datacite:titles>
```

3.2 Creator (MA)

`datacite:creator`

3.2.1 Cardinality

Mandatory

Occurrence: 1-n

3.2.2 Definition and Usage Instruction

The authors of the publication in priority order. May be a corporate/institutional or personal name.

Do Not Confuse With

- *Contributor (MA)*
- *Publisher (MA)*

Remarks

- adapted from DataCite MetadataKernel[30] v4.1

Property creator (M, 1-n)

Subproperty creatorName (M)

The name of the author (occurrence: 1). The format should be: family, given. Non-roman names may be transliterated according to the ALA-LC[31] schemas.

[30] https://schema.datacite.org/meta/kernel-4.4/
[31] http://www.loc.gov/catdir/cpso/roman.html

Attribute nameType (R)

The type of name (occurrence: 0-1).

Controlled list values

- Organizational
- Personal

Subproperty givenName (R)

The personal or first name of the author.

Subproperty familyName (R)

The surname or last name of the author.

Subproperty nameIdentifier (R)

Uniquely identifies an individual or legal entity, according to various schemes (occurrences: 0-n). The format is dependent upon scheme.

Note: OpenAIRE recommends including a nameIdentifierSchema such as an ORCID, ISNI, ROR, or GRID if available.

Attribute nameIdentifierScheme (M)

The name of the name identifier scheme (occurrence: 1). Mandatory if nameIdentifier is used.

Attribute schemeURI (R)

The URI of the name identifier scheme (occurrence: 0-1).

Subproperty affiliation (R)

The organizational or institutional affiliation of the creator (occurrence: 0-n).

3.2.3 Example

```
<datacite:creators>
  <datacite:creator>
    <datacite:creatorName>Evans, R.J.</datacite:creatorName>
    <datacite:affiliation>Institute of Science and Technology</
↪datacite:affiliation>
    <datacite:nameIdentifier nameIdentifierScheme="ORCID"
                    schemeURI="http://orcid.org">
      1234-1234-1234-1234
    </datacite:nameIdentifier>
  </datacite:creator>
</datacite:creators>
```

3.3 Contributor (MA)

`datacite:contributor`

3.3.1 Cardinality

Mandatory if applicable

Occurrence: 0-n

3.3.2 Definition and Usage Instruction

The institution or person responsible for collecting, managing, distributing, or otherwise contributing to the development of the resource.

Do Not Confuse With

- *Publisher (MA)*
- *Creator (MA)*
- *Funding Reference (MA)*

Remarks

- adapted from DataCite MetadataKernel[32] v4.1

Property contributor (MA, 0-n)

Attribute contributorType (M)

The type of contributor of the resource (occurrence: 1). Mandatory if *contributor* is used.

Controlled list values

- ContactPerson
- DataCollector
- DataCurator
- DataManager
- Distributor
- Editor
- HostingInstitution
- Producer
- ProjectLeader
- ProjectManager
- ProjectMember
- RegistrationAgency
- RegistrationAuthority
- RelatedPerson
- Researcher
- ResearchGroup

[32] https://schema.datacite.org/meta/kernel-4.4/

- RightsHolder
- Sponsor
- Supervisor
- WorkPackageLeader
- Other

and additionally to the DataCite contributor types above is the controlled vocabulary **Contributor Roles Taxonomy** (CRediT) from National Information Standards Organization (NISO)[33] with additionally

- Conceptualization
- FormalAnalysis
- FundingAcquisition
- Investigation
- Methodology
- Validation
- Visualization

Subproperty contributorName (M)

The name of the contributor (occurrence: 1). Mandatory if *Contributor* is used.

Attribute nameType (R)

The type of name (occurrence: 0-1).

Controlled list values

- Organizational
- Personal

Subproperty familyName (O)

The surname or last name of the contributor (occurrence: 0-1).

Subproperty givenName (O)

The personal or first name of the contributor (occurrence: 0-1).

Subproperty nameIdentifier (R)

Uniquely identifies an individual or legal entity, according to various schemes (occurrence: 0-n).

Attribute nameIdentifierScheme (M)

The name of the name identifier scheme (occurrence: 1). Mandatory if *nameIdentifier* is used.

[33] http://credit.niso.org

Attribute schemeURI (R)

The URI of the name identifier scheme (occurrence: 0-1).

Subproperty affiliation (R)

The organisational or institutional affiliation of the contributor.

3.3.3 Example

```
<datacite:contributors>
        <datacite:contributor>
          <datacite:contributorName>Evans, R. J.</datacite:contributorName>
        <datacite:contributor>
        <datacite:contributor>
          <datacite:contributorName>International Human Genome Sequencing␣
↪Consortium</datacite:contributorName>
        </datacite:contributor>
</datacite:contributors>
```

3.4 Funding Reference (MA)

`oaire:fundingReference`

3.4.1 Cardinality

Mandatory if applicable

Occurrence: 0-n

3.4.2 Definition and Usage Instruction

Information about financial support (funding) for the resource being registered.

Usage Instruction

An authoritative list of projects is exposed by OpenAIRE through OAI-PMH[34] (DEPRECATED, since January 2021 and see more detailed information here[35]) and REST[36] as documented at http://api.openaire.eu, and available for all repository managers. Values include the project name and project ID. The projectID equals the Grant Agreement identifier or Award number.

Remarks

- introduced as `info:eu-repo/grantAgreement` in previous versions of the OpenAIRE Guidelines
- adopting fundingReference element and subproperties from DataCite MetadataKernel v4.1 which replaces the `info:eu-repo/grantAgreement` syntax.
- adding subproperty fundingStream to this application profile

Property fundingReference (MA, 0-n)

Repeat this property to indicate several different funders and projects.

[34] http://api.openaire.eu/oai_pmh?verb=ListRecords&set=projects&metadataPrefix=oaf
[35] https://develop.openaire.eu/farewell-oai.hmtl
[36] http://api.openaire.eu/search/projects

Subproperty funderName (M)

Name of the funding provider (occurrence: 1). Mandatory if *FundingReference* is used.

Subproperty funderIdentifier (R)

Unique identifier of the funding entity (occurrence: 0-1).

Attribute funderIdentifiertype (R)

Type of the unique identifier of the funding entity (occurrence: 0-1).

Controlled list values

- ISNI
- GRID
- Crossref Funder
- ROR

see also Crossref Funder Registry[37]

Subproperty fundingStream (O)

Name of the funding stream (optional) (occurrence: 0-1).

Subproperty awardNumber (MA)

Project grantId or awardNumber (occurrence: 1).

Attribute awardURI (R)

URI of the project landing page provided by the funder for more information about the award (grant) (occurrence: 0-1).

Subproperty awardTitle (R)

Title of the project, award or grant (occurrence: 0-1).

3.4.3 Example

Examples utilizing all fields:

```
<oaire:fundingReferences>
  <oaire:fundingReference>
   <oaire:funderName>SNSF</oaire:funderName>
   <oaire:funderIdentifier funderIdentifierType="ISNI">http://www.isni.org/isni/
↪0000000106723101</oaire:funderIdentifier>
   <oaire:fundingStream>International short research visits</oaire:fundingStream>
   <oaire:awardNumber awardURI="http://p3.snf.ch/project-151094">151094</
↪oaire:awardNumber>
   <oaire:awardTitle>Amygdala fMRI and social cognition in patients with␣
↪unilateral MTLE and Urbach-Wiethe disease</oaire:awardTitle>
```

(continues on next page)

[37] https://www.crossref.org/services/funder-registry/

(continued from previous page)

```
  </oaire:fundingReference>
</oaire:fundingReferences>
```

3.5 Alternate Identifier (R)

`datacite:alternateIdentifier`

3.5.1 Cardinality

Recommended

Occurrence: 0-n

3.5.2 Definition and Usage Instruction

An identifier or identifiers other than the primary Identifier applied to the resource being registered. This may be any alphanumeric string which is unique within its domain of issue. May be used for local identifiers. AlternateIdentifier should be used for another identifier of the same instance (same location, same file).

Remarks

- adapted from DataCite MetadataKernel[38] v4.1

Property alternateIdentifier (R, 0-n)

Value of the atlernate identifier.

Attribute alternateIdentifierType (M)

The type of the AlternateIdentifier (occurrence: 1). Mandatory if *AlternateIdentifier* is used.

The type value is *suggested* in the following list:

Controlled list values

- `ARK` – Archival Resource Key
- `arXiv` – arXiv.org identifier
- `bibcode` – Astrophysics Data System bibliographic codes; bibcodes can be resolved via http://adsabs.harvard.edu/abs/bibcode
- `DOI` – Digital Object Identifier
- `EAN13` – European Article Number, now renamed International Article Number, but retaining the original acronym, is a 13-digit barcoding standard which is a superset of the original 12-digit Universal Product Code (UPC) system.
- `EISSN` – International Standard Serial Number (electronic Version)
- `Handle` – Handle
- `IGSN` – International Geo Sample Number; a 9-digit alphanumeric code that uniquely identifies samples from our natural environment and related sampling features.
- `ISBN` – International Standard Book Number

[38] https://schema.datacite.org/meta/kernel-4.4/

- ISSN – International Standard Serial Number; a unique 8-digit number used to identify a print or electronic periodical publication.
- ISTC – International Standard Text Code; a unique "number" assigned to a textual work. An ISTC consists of 16 numbers and/or letters.
- LISSN – The linking ISSN or ISSN-L enables collocation or linking among different media versions of a continuing resource.
- LSID – Life Science Identifiers; a unique identifier for data in the Life Science domain. Format: urn:lsid:authority:namespace:identifier:revision
- PISSN – International Standard Serial Number (print version)
- PMID – PubMed ID
- PURL – Persistent Uniform Resource Locator
- UPC – Universal Product Code is a barcode symbology used for tracking trade items in stores. Its most common form, the UPC-A, consists of 12 numerical digits.
- URL – Uniform Resource Locator
- URN – Uniform Resource Name
- WOS – Web of Science accession number

3.5.3 Example

```
<datacite:alternateIdentifiers>
   <datacite:alternateIdentifier alternateIdentifierType="URL">http://someUrl</
↪datacite:alternateIdentifier>
</datacite:alternateIdentifiers>
```

3.6 Related Identifier (R)

`datacite:relatedIdentifier`

3.6.1 Cardinality

Recommended

Occurrence: 0-n

3.6.2 Definition and Usage Instruction

An identifier of a related resource other than the primary Identifier applied to the resource being registered.

Remarks

- adapted from DataCite MetadataKernel[39] v4.1

Property relatedIdentifier (R, 0-n)

Use the related identifier as value. Repeat this property for each related identifier.

[39] https://schema.datacite.org/meta/kernel-4.4/

Attribute relatedIdentifierType (M)

The type of the RelatedIdentifier (occurrence: 1). Mandatory if *RelatedIdentifier* is used.

Controlled list values

- `ARK` – Archival Resource Key
- `arXiv` – arXiv.org identifier
- `bibcode` – Astrophysics Data System bibliographic codes; bibcodes can be resolved via http://adsabs.harvard.edu/abs/bibcode
- `DOI` – Digital Object Identifier
- `EAN13` – European Article Number, now renamed International Article Number, but retaining the original acronym, is a 13-digit barcoding standard which is a superset of the original 12-digit Universal Product Code (UPC) system.
- `EISSN` – International Standard Serial Number (electronic Version)
- `Handle` – Handle
- `IGSN` – International Geo Sample Number; a 9-digit alphanumeric code that uniquely identifies samples from our natural environment and related sampling features.
- `ISBN` – International Standard Book Number
- `ISSN` – International Standard Serial Number; a unique 8-digit number used to identify a print or electronic periodical publication.
- `ISTC` – International Standard Text Code; a unique "number" assigned to a textual work. An ISTC consists of 16 numbers and/or letters.
- `LISSN` – The linking ISSN or ISSN-L enables collocation or linking among different media versions of a continuing resource.
- `LSID` – Life Science Identifiers; a unique identifier for data in the Life Science domain. Format: urn:lsid:authority:namespace:identifier:revision
- `PISSN` – International Standard Serial Number (print version)
- `PMID` – PubMed ID
- `PURL` – Persistent Uniform Resource Locator
- `UPC` – Universal Product Code is a barcode symbology used for tracking trade items in stores. Its most common form, the UPC-A, consists of 12 numerical digits.
- `URL` – Uniform Resource Locator
- `URN` – Uniform Resource Name
- `WOS` – Web of Science accession number

Attribute relationType (M)

Description of the relationship of the resource being registered (A) and the related resource (B) (occurrence: 1). Mandatory if *RelatedIdentifier* is used.

Controlled list values

- `IsCitedBy` (indicates that B includes A in a citation)
- `Cites` (indicates that A includes B in a citation)
- `IsSupplementTo` (indicates that A is a supplement to B)
- `IsSupplementedBy` (indicates that B is a supplement to A)
- `IsContinuedBy` (indicates A is continued by the work B)

- `Continues` (indicates A is a continuation of the work B)
- `IsDescribedBy` (indicates A is described by B)
- `Describes` (indicates A describes B)
- `HasMetadata` (indicates resource A has additional metadata B)
- `IsMetadataFor` (indicates additional metadata A for a resource B)
- `HasVersion` (indicates A has a version B)
- `IsVersionOf` (indicates A is a version of B)
- `IsNewVersionOf` (indicates A is a new edition of B, where the new edition has been modified or updated)
- `IsPreviousVersionOf` (indicates A is a previous edition of B)
- `IsPartOf` (indicates A is a portion of B;may be used for elements of a series)
- `HasPart` (indicates A includes the part B)
- `IsReferencedBy` (indicates A is used as a source of information by B)
- `References` (indicates B is used as a source of information for A)
- `IsDocumentedBy` (indicates B is documentation about/explaining A)
- `Documents` (indicates A is documentation about/explaining B)
- `IsCompiledBy` (indicates B is used to compile or create A)
- `Compiles` (indicates B is the result of a compile or creation event using A)
- `IsVariantFormOf` (indicates A is a variant or different form of B, e.g. calculated or calibrated form or different packaging)
- `IsOriginalFormOf` (indicates A is the original form of B)
- `IsIdenticalTo` (indicates that A is identical to B, for use when there is a need to register two separate instances of the same resource)
- `IsReviewedBy` (indicates that A is reviewed by B)
- `Reviews` (indicates that A is a review of B)
- `IsDerivedFrom` (indicates B is a source upon which A is based)
- `IsSourceOf` (indicates A is a source upon which B is based)
- `IsRequiredBy` (indicates A is required by B)
- `Requires` (indicates A requires B)
- `IsPublishedIn` (indicates that A is published in B)

Attribute relatedMetadataScheme (O)

The name of the scheme (occurrences: 0-1).

Allowed values, examples, other constraints

Use only with this relation pair: (`HasMetadata`/`IsMetadataFor`).

Attribute schemeURI (O)

The URI of the relatedMetadataScheme (occurrences: 0-1).

Allowed values, examples, other constraints

Use only with this relation pair: (`HasMetadata`/`IsMetadataFor`).

Attribute schemeType (O)

The type of the relatedMetadataScheme, linked with the schemeURI (occurrences: 0-1).

Allowed values, examples, other constraints

Use only with this relation pair: (`HasMetadata`/`IsMetadataFor`).

Examples: `XSD`, `DDT`, `Turtle`

Attribute resourceTypeGeneral (O)

The general type of the related resource (occurrences: 0-1).

Controlled list values

- Audiovisual
- Collection
- DataPaper
- Dataset
- Event
- Image
- InteractiveResource
- Model
- PhysicalObject
- Service
- Software
- Sound
- Text
- Workflow
- Other

3.6.3 Example

```
<datacite:relatedIdentifiers>
   <datacite:relatedIdentifier relatedIdentifierType="URL" relationType="HasPart">
↪http://someUrl</datacite:relatedIdentifier>
</datacite:relatedIdentifiers>
```

3.7 Embargo Period Date (MA)

`datacite:date`

3.7.1 Cardinality

Mandatory if applicable

Occurrence: 2

3.7.2 Definition and Usage Instruction

Dates relevant to describe an embargo period.

A date associated with an event in the life cycle of the resource. Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and follows the `YYYY-MM-DD` format.

Remarks

- introduced as info:eu-repo/date/embargoEnd/[YYYY-MM-DD][40] in previous versions of the OpenAIRE Guidelines
- this version of the application profile adopts the *Date* element in combination with *dateType* attributes from DataCite MetadataKernel v4.1 which replaces the `info:eu-repo/date/EmbargoEnd` syntax.

When *Access Rights (M)* is set to:

```
<datacite:rights uri="http://purl.org/coar/access_right/c_f1cf">embargoed access</
↪<datacite:rights>
```

the start and end date of the embargo period must be provided.

Do Not Confuse With

- *Publication Date (M)* (Use `datacite:date` for **Publication Date** related to the resource.)

Property date (MA, 2)

Use the date of the embargo start as value in one property and the date of the embargo end in the other property.

Attribute dateType (M)

The type of date. Choose from the date type vocabulary the controlled term `Accepted` to indicate the start and the term `Available` to indicate the end of an embargo period.

3.7.3 Example

```
<datacite:dates>
  <datacite:date dateType="Accepted">2011-12-01</datacite:date>
  <datacite:date dateType="Available">2012-12-01</datacite:date>
</datacite:dates>
```

3.8 Language (MA)

`dc:language`

3.8.1 Cardinality

Mandatory if applicable

Occurrence: 0-n

[40] https://wiki.surfnet.nl/display/standards/info-eu-repo/#info-eu-repo-DateTypesandvalue

3.8.2 Definition and Usage Instruction

DCMI Definition

A language of the intellectual content of the resource.

Usage Instruction

A specific resource (an instance of scientific output) is either written in one human language or more. In these cases all used languages are used in the DC element `language`. If a specific resource (an instance of scientific output) is written in one human language and is translated into other human languages, each translation does have its own record.

Recommendation: take values from one of the following lists:

- IETF BCP 47, the IANA Language Subtag Registry[41]
- ISO 639-x, where x can be 1,2 or 3. Best Practice: we use ISO 639-3 and by doing so we follow: http://www.sil.org/iso639-3/

If necessary, repeat this element to indicate multiple languages.

If ISO 639-2 and 639-1 are sufficient for the contents of a repository they can be used alternatively. Since there is a unique mapping this can be done during an aggregation process.

Remarks

- introduced in DRIVER Guidelines v2 element language[42]

Property language (MA, 0-n)

Use the language code as value.

3.8.3 Example

```
<dc:language>eng</dc:language>
<dc:language>deu</dc:language>
<dc:language>nld</dc:language>
<dc:language>nld/dut</dc:language>
<dc:language>dut</dc:language>
<dc:language>nl</dc:language>
```

3.9 Publisher (MA)

`dc:publisher`

3.9.1 Cardinality

Mandatory if applicable

Occurrence: 0-n

[41] http://www.iana.org/assignments/language-subtag-registry
[42] https://wiki.surfnet.nl/display/DRIVERguidelines/Language

3.9.2 Definition and Usage Instruction

DCMI Definition

An entity responsible for making the resource available. Examples of a Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity.

Usage Instruction

The (commercial or non-commercial) publisher of the resource; not the (sub)institution the author is affiliated with. Publisher is used only in the bibliographic / functional sense, not an organisational one. Use only the full name of the given (commercial) publisher, not the name of an organization or institute that is otherwise [in a broader sense] associated with the creator.

With university publications place the name of the faculty and/or research group or research school after the name of the university. In the case of organizations where there is clearly a hierarchy present, list the parts of the hierarchy from largest to smallest, separated by full stops. If it is not clear whether there is a hierarchy present, or unclear which is the larger or smaller portion of the body, give the name as it appears in the eprint.

The use of publisher names from authority lists constructed according to local or national thesaurus files is optional.

Do Not Confuse With

- *Contributor (MA)*
- *Creator (MA)*

In most cases the publisher and the creator are not the same.

Remarks

- introduced in DRIVER Guidelines v2 element publisher[43]

Property publisher (MA, 0-n)

Use the name of the publisher as value.

3.9.3 Example

```
<dc:publisher>
  Loughborough University. Department of Computer Science
</dc:publisher>
<dc:publisher>John Wiley &amp;amp; Sons, Inc. (US)</dc:publisher>
```

3.10 Publication Date (M)

`datacite:date`

3.10.1 Cardinality

Mandatory

Occurrence: 1

[43] https://wiki.surfnet.nl/display/DRIVERguidelines/Publisher

3.10.2 Definition and Usage Instruction

DCMI Definition

A date associated with an event in the life cycle of the resource. Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF][44] and follows the `YYYY-MM-DD` format.

Usage Instruction

The date should be formatted according to the W3C encoding rules for dates and times:

Complete date:

`YYYY-MM-DD` (e.g. 1997-07-16)

where:

- `YYYY` [four-digit year] is ''mandatory"
- `MM` [two-digit month (01=January, etc.)] is ''optional"
- `DD` [two-digit day of month (01 through 31)] is ''optional"

One date field – Date of Publication:

Often repository systems have more then one date fields that serve different purposes. Date of creation, publication, modified, promotion, etc. Preferably in the end-users perspective the most logical and meaningful date will be the date of publication.

No date of publication available:

If no date of publication is available, use any other date available. It is better to use one date than no date at all.

Datestamp additions:

Additions like "Zulu time" should NOT be part of the metadata.

Fuzzy dates:

For fuzzy dates use a logical year that most represents that period, e.g. `1650` instead of `17th century`.

To express more about that temporal period, one can use the `dc:coverage` field. A temporal period can be expressed in a standard way when precisely defined (see *Coverage (R)*) or when "fuzzy" or uncertain by free text expressions. A service provider is able to sort dates based on date standards like W3CDTF. Since there is no standard for fuzzy dates for terms like "Renaissance" or "17th Century", they will simply not appear on date-based query results.

Remarks

- introduced in DRIVER Guidelines v2 element date[45]
- this version of the application profile adopts the *Date* element in combination with *dateType* attribute from DataCite MetadataKernel[46] v4.4.

Do Not Confuse With

- *Embargo Period Date (MA)* (Use `datacite:date` for **Embargo Period Date** related to the resource.)

Property date (M, 1)

Use the publication date as value.

[44] https://www.iso.org/iso-8601-date-and-time-format.html
[45] https://wiki.surfnet.nl/display/DRIVERguidelines/Date
[46] https://schema.datacite.org/meta/kernel-4.4/

Attribute dateType (M)

The type of date. Choose from the date type vocabularyUse the controlled term `Issued` to indicate the date of publication.

3.10.3 Example

```
<datacite:date dateType="Issued">2000-12-25</datacite:date>
```

3.11 Resource Type (M)

`oaire:resourceType`

3.11.1 Cardinality

Mandatory

Occurrence: 1

3.11.2 Definition and Usage Instruction

The type of scientific output the resource is a manifestation of. It describes the genre of the resource.

Usage

The attribute *resourceTypeGeneral* is used to categorize the resource to belong to a main class of research outputs. The attribute *uri* holds an HTTP URI of a resource type concept and indicates the sub-property of *resourceTypeGeneral*. The label of this concept is used as value for the *ResourceType* element.

Do Not Confuse With

- *Format (R)* which describes the media type of this resource.

Remarks

- former versions of the OpenAIRE Guidelines used the info:eu-repo vocabulary for publication types[47].
- adopting *resourceType* element from DataCite MetadataKernel v4.1.
- adding the *uri* attribute for resource type concept URI to this application profile

Property resourceType (M, 1)

Use the label of the resource type term as value. In the below table the preferred english labels are listed, but labels (preferred or alternative) in other languages can be chosen from the COAR Resource Type Vocabulary.

Attribute resourceTypeGeneral (M)

The general type of a resource.

Controlled list values

- `literature`
- `dataset`
- `software`

[47] https://wiki.surfnet.nl/display/standards/info-eu-repo/#info-eu-repo-Publicationtypes

- other research product

Attribute uri (M)

Use terms from the COAR Resource Type Vocabulary[48] (occurrence: 1).

Controlled list values

conceptURI	label
http://purl.org/coar/resource_type/c_1162	annotation
http://purl.org/coar/resource_type/c_0640	journal
http://purl.org/coar/resource_type/c_6501	journal article
http://purl.org/coar/resource_type/c_545b	letter to the editor
http://purl.org/coar/resource_type/c_b239	editorial
http://purl.org/coar/resource_type/c_2df8fbb1	research article
http://purl.org/coar/resource_type/c_dcae04bc	review article
http://purl.org/coar/resource_type/c_beb9	data paper
http://purl.org/coar/resource_type/c_3e5a	contribution to journal
http://purl.org/coar/resource_type/c_ba08	book review
http://purl.org/coar/resource_type/c_3248	book part
http://purl.org/coar/resource_type/c_2f33	book
http://purl.org/coar/resource_type/c_86bc	bibliography
http://purl.org/coar/resource_type/c_816b	preprint
http://purl.org/coar/resource_type/c_8042	working paper
http://purl.org/coar/resource_type/c_71bd	technical documentation
http://purl.org/coar/resource_type/c_18gh	technical report
http://purl.org/coar/resource_type/c_18ws	research report
http://purl.org/coar/resource_type/c_18hj	report to funding agency
http://purl.org/coar/resource_type/c_18op	project deliverable
http://purl.org/coar/resource_type/c_186u	policy report
http://purl.org/coar/resource_type/c_18wq	other type of report
http://purl.org/coar/resource_type/c_18wz	memorandum
http://purl.org/coar/resource_type/c_18ww	internal report
http://purl.org/coar/resource_type/c_efa0	review
http://purl.org/coar/resource_type/c_baaf	research proposal
http://purl.org/coar/resource_type/c_ba1f	report part
http://purl.org/coar/resource_type/c_93fc	report
http://purl.org/coar/resource_type/c_15cd	patent
http://purl.org/coar/resource_type/c_18co	conference poster not in proceedings
http://purl.org/coar/resource_type/c_18cp	conference paper not in proceedings
http://purl.org/coar/resource_type/c_6670	conference poster
http://purl.org/coar/resource_type/c_5794	conference paper
http://purl.org/coar/resource_type/c_c94f	conference object
http://purl.org/coar/resource_type/c_f744	conference proceedings
http://purl.org/coar/resource_type/c_7a1f	bachelor thesis
http://purl.org/coar/resource_type/c_bdcc	master thesis
http://purl.org/coar/resource_type/c_db06	doctoral thesis
http://purl.org/coar/resource_type/c_46ec	thesis
http://purl.org/coar/resource_type/c_0857	letter
http://purl.org/coar/resource_type/c_8544	lecture
http://purl.org/coar/resource_type/c_18cf	text
http://purl.org/coar/resource_type/c_18cw	musical notation
http://purl.org/coar/resource_type/c_18cd	musical composition
http://purl.org/coar/resource_type/c_18cc	sound

Continued on next page

[48] http://vocabularies.coar-repositories.org/documentation/resource_types/

Table 2 – continued from previous page

conceptURI	label
http://purl.org/coar/resource_type/c_12ce	video
http://purl.org/coar/resource_type/c_8a7e	moving image
http://purl.org/coar/resource_type/c_2659	periodical
http://purl.org/coar/resource_type/c_ecc8	still image
http://purl.org/coar/resource_type/c_c513	image
http://purl.org/coar/resource_type/c_12cd	map
http://purl.org/coar/resource_type/c_12cc	cartographic material
http://purl.org/coar/resource_type/c_5ce6	software
http://purl.org/coar/resource_type/c_ddb1	dataset
http://purl.org/coar/resource_type/c_e9a0	interactive resource
http://purl.org/coar/resource_type/c_7ad9	website
http://purl.org/coar/resource_type/c_393c	workflow
http://purl.org/coar/resource_type/c_1843	other

3.11.3 Example

```
<oaire:resourceType resourceTypeGeneral="literature" uri="http://purl.org/coar/
↪resource_type/c_6501">journal article</oaire:resourceType>
```

3.12 Description (MA)

`dc:description`

3.12.1 Cardinality

Mandatory if applicable

Occurrence: 0-n

3.12.2 Definition and Usage Instruction

DCMI Definition

An account of the content of the resource. Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.

Usage Instruction

This element is used for a textual description of the content. When a resource consists of several separate physical object files, do not use `dc:description` to list the URLs of these files.

Remarks

- introduced in DRIVER Guidelines v2 element description[49]

Property description (MA, 0-n)

Use the textual description as value.

[49] https://wiki.surfnet.nl/display/DRIVERguidelines/Description

Attribute lang (O)

The language of the description (occurrence: 0-1).

Use the `xml:lang` attribute to indicate the language of the description.

3.12.3 Example

```
<dc:description>
  Foreword [by] Hazel Anderson; Introduction; The scientific heresy:
  transformation of a society; Consciousness as causal reality [etc]
</dc:description>

<dc:description xml:lang="en-US">
  A number of problems in quantum state and system identification are
  addressed.
</dc:description>
```

3.13 Format (R)

`dc:format`

3.13.1 Cardinality

Recommended

Occurrence: 0-n

3.13.2 Definition and Usage Instruction

DCMI Definition

The physical or digital manifestation of the resource. Typically, Format may include the media-type or dimensions of the resource. Format may be used to determine the software, hardware or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [MIME] defining computer media formats).

Usage Instruction

Based on best practice, the IANA registered list of Internet Media Types (MIME types) is used to select a term from. For the full list see http://www.iana.org/assignments/media-types

If one specific resource (an instance of scientific output) has more than one physical formats (e.g. postscript and pdf) stored as different object files, all formats are mentioned in the DC element format, for example:

- `<dc:format>application/pdf</dc:format>`
- `<dc:format>application/postscript</dc:format>`
- `<dc:format>application/vnd.oasis.opendocument.text</dc:format>`

Do Not Confuse With

- *Resource Type (M)*
- *Resource Identifier (M)*

DC element `format` describes the media type of this resource. `oaire:resourceType` describes the kind of academic output the resource is a representation of. `datacite:identifier` is used to represent the manifestation of the digital resource.

Remarks

- introduced in DRIVER Guidelines v2 element format[50]

Property format (R, 0-n)

Use the media type of the resource as value.

3.13.3 Example

```
<dc:format>video/quicktime</dc:format>
<dc:format>application/pdf</dc:format>
<dc:format>application/xml</dc:format>
<dc:format>application/xhtml+xml</dc:format>
<dc:format>application/html</dc:format>
<dc:format>application/vnd.oasis.opendocument.text</dc:format>
```

3.14 Resource Identifier (M)

`datacite:identifier`

3.14.1 Cardinality

Mandatory

Occurrence: 1

3.14.2 Definition and Usage Instruction

The Identifier is a unique string that identifies a resource.

Usage Instruction

Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Example formal identification systems include the Uniform Resource Identifier (URI), the Uniform Resource Locator (URL), the Digital Object Identifier (DOI), and the `URN:NBN`. Also this can be a direct URL, or a redirection URL, like PURL, HANDLE or other international resolution mechanisms.

The ideal use of this element is to use a direct link or a link to a jump-off page (persistent URL) from `identifier` in the metadata record to the digital resource or a jump-off page.

Smart practice:

- use a stable, sustainable URL

Remarks

- adapted from DataCite MetadataKernel[51] v4.1

Do Not Confuse With

- *Alternate Identifier (R)* (Use `datacite:alternativeIdentifier` to list other identifiers than the primary identifier applied to the same resource.)

[50] https://wiki.surfnet.nl/display/DRIVERguidelines/Format
[51] https://schema.datacite.org/meta/kernel-4.4/

- *Related Identifier (R)* (Use `datacite:relatedIdentifier` to refer to related resources.)
- *File Location (MA)* (Use `oaire:file` to point to the resource being desccribed by this metadata, e.g. the fulltext file.)
- *Source (R)* (Use `dc:source` for bibliographic citation of the originating resource.)

Property identifier (M, 1)

Use the identifier link as value.

Attribute identifierType (M)

The type of the Identifier (occurrences: 1).

Allowed values, examples, other constraints

Controlled list values

- ARK
- DOI
- Handle
- IGSN
- PURL
- URL
- URN

Note: Unlike DataCite, OpenAIRE allows for DOIs and other types of identifiers.

3.14.3 FAIR enabled

Indicator	Priority	Description
RDA-F1-01M	essential	Metadata is identified by a persistent identifier
RDA-F1-02M	essential	Metadata is identified by a global unique id
RDA-A1.1-01D	important	Data is accessible through a free access protocol

3.14.4 Example

In this example the handle redirects to the jump-off page. A jump-off page is a good way to refer to. The end-user has the opportunity to see more information about the object(s) he has found, see the context and enjoy the other services a local repository has to offer:

```
<datacite:identifier identifierType="Handle">http://hdl.handle.net/1234/5628</
↪datacite:identifier>
```

3.15 Access Rights (M)

`datacite:rights`

3.15.1 Cardinality

Mandatory

Occurrence: 1

3.15.2 Definition and Usage Instruction

Access right of the resource.

Information about the right or mode the resource can be accessed. If the metadata describe more than one resource, e.g. fulltext and supplementary material, the access right of the main resource should be provided.

Use terms from the COAR Access Right Vocabulary[52] (occurrence: 1).

conceptURI	label
http://purl.org/coar/access_right/c_abf2	`open access`
http://purl.org/coar/access_right/c_f1cf	`embargoed access`
http://purl.org/coar/access_right/c_16ec	`restricted access`
http://purl.org/coar/access_right/c_14cb	`metadata only access`

Note: Unlike DataCite, OpenAIRE restricts the use of this property to indicate the access right.

Do Not Confuse With

- *License Condition (R)* (Use `oaire:licenseCondition` for license information related to the resource.)

Remarks

- former versions of the OpenAIRE Guidelines used the info:eu-repo-Access-Terms vocabulary[53].

Property accessRights (M, 1)

Use the label of the vocabulary term as value.

Attribute uri (M)

Use the conceptURI of the vocabulary term.

3.15.3 Example

```
<datacite:rights rightsURI="http://purl.org/coar/access_right/c_abf2">open access</
↪datacite:rights>
```

3.16 Source (R)

`dc:source`

[52] http://vocabularies.coar-repositories.org/documentation/access_rights/
[53] https://wiki.surfnet.nl/display/standards/info-eu-repo/#info-eu-repo-AccessRights

3.16.1 Cardinality

Recommended

Occurrence: 0-n

3.16.2 Definition and Usage Instruction

DCMI Definition

A reference to a resource from which the present resource is derived.

Usage Instruction

The present resource may be derived from the *Source* resource in whole or in part. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.

Best practice: Use only when the described resource is the result of digitization of non-digital originals. Otherwise, use *:ref:'dci:relatedIdentifier'*. Optionally metadata about the current location and call number of the digitized publication can be added.

Use: Guidelines for Encoding Bibliographic Citation Information in Dublin Core Metadata (http://dublincore.org/documents/dc-citation-guidelines/).

Do Not Confuse With

- *Resource Identifier (M)*

Remarks

- introduced in DRIVER Guidelines v2 element source[54]

Property source (R, 0-n)

Use source information as value.

3.16.3 Example

```
<dc:source>Ecology Letters (1461023X) vol.4 (2001)</dc:source>
```

3.17 Subject (MA)

`datacite:subject`

3.17.1 Cardinality

Mandatory if applicable

Occurrence: 0-n

[54] https://wiki.surfnet.nl/display/DRIVERguidelines/Source

3.17.2 Definition and Usage Instruction

Subject, keyword, classification code, or key phrase describing the resource (occurrences: 0-n).

Usage Instruction

In the subject property two kinds of values are possible: encode either a keyword or a classification.

In general, choose the most significant and unique words for keywords, avoiding those too general to describe a particular resource.

For keywords/keyphrases that are not controlled by a vocabulary or thesaurus either encode multiple terms with a semi-colon separating each keyword/keyphrase; or repeat the element for each term. There are no requirements regarding the capitalization of keywords though internal (within archive) consistency is recommended.

Where terms are taken from a standard classification schema: encode each term using the additional attributes of the subject property. Encode the complete subject descriptor according to the relevant scheme. Use the capitalisation and punctuation used in the original scheme.

It is recommended to use an URI when using classification schemes or controlled vocabularies especially when codified schemes are used DDC or UDC. Service providers can recognise encoding schemas more easy when the schema is "URI-fied" by an authority namespace.

If no specific classification scheme is used we recommend the Dewey Decimal Classification (DDC). More information about the DDC and the DDC Summaries can be found at https://www.oclc.org/en/dewey/resources.html . Please note that OCLC owns all copyright rights in the Dewey Decimal Classification system. Dewey, Dewey Decimal Classification, DDC, OCLC and WebDewey are registered trademarks of OCLC.

Remarks

- adapted from DataCite MetadataKernel[55] v4.1

Property subject (MA, 0-n)

Use subject name or keyword as value.

Attribute subjectScheme (O)

The name of the subject scheme or classification code or authority if one is used (occurrences: 0-1).

Allowed values, examples, other constraints

Free text.

Attribute schemeURI (O)

The URI of the subject identifier scheme (occurrences: 0-1).

Allowed values, examples, other constraints

Examples:

- http://id.loc.gov/authorities/subjects
- http://dewey.info/

Attribute valueURI (O)

The URI of the subject term.

[55] https://schema.datacite.org/meta/kernel-4.4/

3.17.3 Example

```
<datacite:subjects>
 <datacite:subject>Earth sciences and geology</datacite:subject>
 <datacite:subject subjectScheme="DDC" schemeURI="http://dewey.info/" valueURI="">
 551 Geology, hydrology, meteorology
 </datacite:subject>
</datacite:subjects>
```

3.18 License Condition (R)

`oaire:licenseCondition`

3.18.1 Cardinality

Recommended

Occurrence: 1

3.18.2 Definition and Usage Instruction

DCMI Definition

Information about license rights held in and over the resource.

Usage Instruction

Typically, a rights element will contain a rights management statement for the access or use of the object, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. It is preferred to refer to a rights service where the reuse rights are made clear to the end-user by using a URL. For example the Creative Commons organisation has created URIs for their different licences in the different jurisdictions. This can be applied to create machine readable usage licenses.

Property licenseCondition (R, 1)

Use the name of the license as value.

Attribute uri (MA)

The URL provides the location where the license can be read. With creative common licenses the type of license can be recognized in the URL name itself. A pro for having the license point to an URL in this way, is that this is machine readable.

Attribute startDate (MA)

This attribute indicates the date when the license comes into effect. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and follows the `YYYY-MM-DD` format.

3.18.3 Example

```
<!-- example 1 -->
<oaire:licenseCondition startDate="2019-02-01" uri="http://creativecommons.org/
↪licenses/by-nc/4.0/">Creative Commons Attribution-NonCommercial</
↪oaire:licenseCondition>
```

(continues on next page)

(continued from previous page)

3.19 Coverage (R)

`dc:coverage`

3.19.1 Cardinality

Recommended

Occurrence: 0-n

3.19.2 Definition and Usage Instruction

The extent or scope of the content of the resource. Coverage will typically include temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity).

Usage Instruction

It is recommended to use literal or non-literal values. If necessary, repeat this element to encode multiple periods.

Remarks

- introduced in DRIVER Guidelines v2 element coverage[56]
- to describe spatial location information (a place name or geographic coordinates) use the property *Geo Location (O)*

Property coverage (R, 0-n)

Use temporal period or jurisdiction information as value.

3.19.3 Example

Example Spatial: temporal topic:

```
<dc:coverage>2000-2010</dc:coverage>
```

Example Spatial: BOX:

```
<dc:coverage>
  scheme=historic; content=Ming Dynasty
</dc:coverage>
```

3.20 Size (O)

`datacite:size`

[56] https://wiki.surfnet.nl/display/DRIVERguidelines/Coverage

3.20.1 Cardinality

Optional

Occurrence: 0-n

3.20.2 Definition and Usage Instruction

Unstructured size information about the resource.

Allowed values, examples, other constraints

Free text.

Examples: “15 pages”, “6 MB”

Remarks

- adapted from DataCite MetadataKernel[57] v4.1

Property size (O, 0-n)

Use size information as value. Repeat the property for different size information domains.

3.20.3 Example

```
<datacite:sizes>
        <datacite:size>15 pages</datacite:size>
        <datacite:size>6 MB</datacite:size>
</datacite:sizes>
```

3.21 Geo Location (O)

`datacite:geoLocation`

3.21.1 Cardinality

Optional

Occurrence: 0-n

3.21.2 Definition and Usage Instruction

Spatial region or named place where the data was gathered or about which the data is focused.

Property geoLocation (O, 0-n)

Repeat this property to indicate several different locations.

[57] https://schema.datacite.org/meta/kernel-4.4/

Subproperty geoLocationPoint (O)

A point location in space (occurrences: 0-1).

A point contains a single latitude-longitude pair.

See *Detailed usage instructions*.

pointLongitude (M)

Longitudinal dimension of point (occurrence: 1).

Mandatory if *geoLocationPoint* is used.

pointLatitude (M)

Latitudinal dimension of point (occurrence: 1).

Mandatory if *geoLocationPoint* is used.

Subproperty geoLocationBox (O)

The spatial limits of a place or box (occurrences: 0-1).

Allowed values, examples, other constraints

A box is defined by two geographic points. Left lower corner (normally south west), right upper corner (normally north east). Each point is defined by its longitude and latitude.

See *Detailed usage instructions*.

westBoundLongitude (M)

Western longitudinal dimension of box. Mandatory if *geoLocationBox* is used.

eastBoundLongitude (M)

Eastern longitudinal dimension of box. Mandatory if *geoLocationBox* is used.

southBoundLatitude (M)

Southern latitudinal dimension of box. Mandatory if *geoLocationBox* is used.

northBoundLatitude (M)

Northern latitudinal dimension of box. Mandatory if *geoLocationBox* is used.

Subproperty geoLocationPlace (O)

Description of a geographic location (occurrences: 0-1).

Allowed values, examples, other constraints

Free text. Use to describe a geographic location.

Subproperty geoLocationPolygon (O)

A drawn polygon area, defined by a set of points and lines connecting the points in a closed chain (occurrences: 0-n).

polygonPoint (M)

A point location in a polygon (occurrences: 4-n). Mandatory if *geoLocationPolygon* is used.

pointLongitude (M)

Longitudinal dimension of point (occurrence: 1). Mandatory if *polygonPoint* is used.

pointLatitude (M)

Latitudinal dimension of point (occurrence: 1). Mandatory if *polygonPoint* is used.

inPolygonPoint (O)

For any bound area that is larger than half the earth, define a (random) point inside.

pointLongitude (M)

Longitudinal dimension of point (occurrence: 1). Mandatory if *inPolygonPoint* is used.

pointLatitude (M)

Latitudinal dimension of point (occurrence: 1). Mandatory if *inPolygonPoint* is used.

Detailed usage instructions

Use WGS 84 (World Geodetic System) coordinates. Use only decimal numbers for coordinates. Longitudes are -180 to 180 (0 is Greenwich, negative numbers are west, positive numbers are east), Latitudes are -90 to 90 (0 is the equator; negative numbers are south, positive numbers north).

Remarks

- adapted from DataCite MetadataKernel[58] v4.1

3.21.3 Example

```
<datacite:geoLocations>
  <datacite:geoLocation>
    <datacite:geoLocationPlace>Atlantic Ocean</datacite:geoLocationPlace>
    <datacite:geoLocationPoint>
              <datacite:pointLongitude>31.233</datacite:pointLongitude>
              <datacite:pointLatitude>-67.302</datacite:pointLatitude>
    </datacite:geoLocationPoint>
    <datacite:geoLocationBox>
```

(continues on next page)

[58] https://schema.datacite.org/meta/kernel-4.4/

(continued from previous page)

```
            <datacite:westBoundLongitude>-71.032</datacite:westBoundLongitude>
            <datacite:eastBoundLongitude>-68.211</datacite:eastBoundLongitude>
            <datacite:southBoundLongitude>41.090</datacite:southBoundLongitude>
            <datacite:northBoundLongitude>42.893</datacite:northBoundLongitude>
    </datacite:geoLocationBox>
  </datacite:geoLocation>
</datacite:geoLocations>
```

3.22 Resource Version (R)

`oaire:version`

3.22.1 Cardinality

Recommended

Occurrence: 1

3.22.2 Definition and Usage Instruction

Depening on the resource type this property is used to indicate

- the version number of a dataset or software
- the status in the publication process of journal articles.

Usage

For *software* and *dataset* resources any string will be accepted, but a semantically-versioned tag is recommended. See <https://semver.org> for more information on semantic versioning.

For *preprints* and *articles* in the journal publishing process a controlled term must be used from the "Journal Article Versions (JAV): Recommendations of the NISO/ALPSP JAV Technical Working Group" (JAV[59]). In tis case the property *must* include the attribute 'uri'. The value of the property is the corresponding label of the HTTP URI.

Property version (R, 1)

Use either a version number or the label of the vocabulary term as value.

Attribute uri (MA)

Allowed HTTP URI are from the COAR Version Types Vocabulary[60].

Version (controlled):

[59] https://www.niso.org/publications/niso-rp-8-2008-jav
[60] http://vocabularies.coar-repositories.org/documentation/version_types/

conceptURI	label	comment
http://purl.org/coar/version/c_b1a7d7d4d402bcce	AO	Author's Original
http://purl.org/coar/version/c_71e4c1898caa6e32	SMUR	Submitted Manuscript Under Review
http://purl.org/coar/version/c_ab4af688f83e57aa	AM	Accepted Manuscript
http://purl.org/coar/version/c_fa2ee174bc00049f	P	Proof
http://purl.org/coar/version/c_970fb48d4fbd8a85	VoR	Version of Record
http://purl.org/coar/version/c_e19f295774971610	CVoR	Corrected Version of Record
http://purl.org/coar/version/c_dc82b40f9837b551	EVoR	Enhanced Version of Record
http://purl.org/coar/version/c_be7fb7dd8ff6fe43	NA	Not Applicable (or Unknown)

3.22.3 Example

```
<oaire:version>1.0.3</oaire:version>
```

```
<oaire:version uri="http://purl.org/coar/version/c_be7fb7dd8ff6fe43">AM</
↪oaire:version>
```

3.23 File Location (MA)

`oaire:file`

3.23.1 Cardinality

Recommended

Occurrence: 0-n

3.23.2 Definition and Usage Instruction

An unambiguous reference to the files, e.g. fulltext, the resource is associated with. Repeat the property for each associated file.

Property file (MA, 0-n)

Use the HTTP URI of the file as value.

Attribute accessRightsURI (R)

Use terms from the COAR Access Right Vocabulary[61].

conceptURI	label
http://purl.org/coar/access_right/c_abf2	open access
http://purl.org/coar/access_right/c_f1cf	embargoed access
http://purl.org/coar/access_right/c_16ec	restricted access
http://purl.org/coar/access_right/c_14cb	metadata only access

[61] http://vocabularies.coar-repositories.org/documentation/access_rights/

Attribute mimeType (R)

Specify the file format. It is recommended to select it from the MIME media type which is registered in IANA. For the full list see http://www.iana.org/assignments/media-types

Attribute objectType (R)

Specify the type of object the file represents. Select it from the following controlled list:

- fulltext
- dataset
- software
- other

3.23.3 Example

```
<oaire:file accessRightsURI="http://purl.org/coar/access_right/c_abf2" mimeType=
↪"application/pdf" objectType="fulltext" version="4.0.1">http://link-to-the-
↪fulltext.org</oaire:file>
```

3.24 Citation Title (R)

oaire:citationTitle

3.24.1 Cardinality

Recommended

Occurrence: 0-1

3.24.2 Definition and Usage Instruction

The title name of the container (e.g. journal, book, conference) this work is published in. This property is considered to be part of the bibliographic citation.

Property citationTitle (R, 0-1)

Use the title name as value.

3.24.3 Example

```
<oaire:citationTitle>some Journal Title</oaire:citationTitle>
```

3.25 Citation Volume (R)

oaire:citationVolume

3.25.1 Cardinality

Recommended

Occurrence: 0-1

3.25.2 Definition and Usage Instruction

The volume, typically a number, of the container (e.g. journal). This property is considered to be part of the bibliographic citation.

Property citationVolume (R, 0-1)

Use the volume number as value.

3.25.3 Example

```
<oaire:citationVolume>10</oaire:citationVolume>
```

3.26 Citation Issue (R)

`oaire:citationIssue`

3.26.1 Cardinality

Recommended

Occurrence: 0-1

3.26.2 Definition and Usage Instruction

The issue of the container (e.g. journal). This property is considered to be part of the bibliographic citation.

Property citationIssue (R, 0-1)

Use the issue number as value.

3.26.3 Example

```
<oaire:citationIssue>1</oaire:citationIssue>
```

3.27 Citation Start Page (R)

`oaire:citationStartPage`

3.27.1 Cardinality

Recommended

Occurrence: 0-1

3.27.2 Definition and Usage Instruction

The start page is part of the pagination information of the work published in a container (e.g. journal issue). This property is considered to be part of the bibliographic citation.

Property citationStartPage (R, 0-1)

Use the start page number as value.

3.27.3 Example

```
<oaire:citationStartPage>100</oaire:citationStartPage>
```

3.28 Citation End Page (R)

`oaire:citationEndPage`

3.28.1 Cardinality

Recommended

Occurrence: 0-1

3.28.2 Definition and Usage Instruction

The end page is part of the pagination information of the work published in a container (e.g. journal issue). This property is considered to be part of the bibliographic citation.

Property citationEndPage (R, 0-1)

Use the end page number as value.

3.28.3 Example

```
<oaire:citationEndPage>105</oaire:citationEndPage>
```

3.29 Citation Edition (R)

`oaire:citationEdition`

3.29.1 Cardinality

Recommended

Occurrence: 0-1

3.29.2 Definition and Usage Instruction

The edition the work was published in (e.g. book edition). This property is considered to be part of the bibliographic citation.

Property citationEdition (R, 0-1)

Use the edition number as value.

3.29.3 Example

```
<oaire:citationEdition>2</oaire:citationEdition>
```

3.30 Citation Conference Place (R)

`oaire:citationConferencePlace`

3.30.1 Cardinality

Recommended

Occurrence: 0-1

3.30.2 Definition and Usage Instruction

The place where the conference took place. This property is considered to be part of the bibliographic citation.

Property citationConferencePlace (R, 0-1)

Use the name of the place as value.

3.30.3 Example

```
<oaire:citationConferencePlace>Berlin</oaire:citationConferencePlace>
```

3.31 Citation Conference Date (R)

`oaire:citationConferenceDate`

3.31.1 Cardinality

Recommended

Occurrence: 0-1

3.31.2 Definition and Usage Instruction

The date when the conference took place. This property is considered to be part of the bibliographic citation. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF][62] and follows the `YYYY-MM-DD` format.

Usage Instruction

The date should be formatted according to the W3C encoding rules for dates and times:

Complete date:

`YYYY-MM-DD` (e.g. 1997-07-16)

where:

- `YYYY` [four-digit year]
- `MM` [two-digit month (01=January, etc.)]
- `DD` [two-digit day of month (01 through 31)]

Property citationConferenceDate (R, 0-1)

Use the *single date* or *start date* and *end date* as values following these patterns:

- `YYYY-MM-DD` [*single date*]
- `YYYY-MM-DD - YYYY-MM-DD` [*start date - end date*]

3.31.3 Example

```
<oaire:citationConferenceDate>2013-10-22</oaire:citationConferenceDate>
```

```
<oaire:citationConferenceDate>2013-09-22 - 2013-09-26</
↪oaire:citationConferenceDate>
```

3.32 Audience (O)

`dcterms:audience`

3.32.1 Cardinality

Optional

Occurrence: 0-n

[62] https://www.iso.org/iso-8601-date-and-time-format.html

3.32.2 Definition and Usage Instruction

DCMI Definition

A class of entity for whom the resource is intended or useful.

Usage Instruction

A class of entity may be determined by the creator or the publisher or by a third party. An example of audiences given is derived from the Common Education Data Standards vocabulary[63]. Please note the list is not exhaustive.

- Administrators
- Community Groups
- Counsellors
- Federal Funds Recipients and Applicants
- Librarians
- News Media
- Other
- Parents and Families
- Policymakers
- Researchers
- School Support Staff
- Student Financial Aid Providers
- Students
- Teachers

Remarks

- introduced in DRIVER Guidelines v2 element audience[64]

Property audience (O, 0-n)

Use the class of entity as value.

3.32.3 Example

```
<dcterms:audience>Researchers</dcterms:audience>
<dcterms:audience>Students</dcterms:audience>
```

[63] https://ceds.ed.gov/element/001492
[64] https://wiki.surfnet.nl/display/DRIVERguidelines/Audience

CHAPTER 4

FAIR enabled

The OpenAIRE and DRIVER guidelines, at the beginning of their time in 2006[65], had already put the focus on elements that can be found today in the FAIR principles. Community-based refinement and enhancement of the guidelines over time to include elements with their descriptions that are consistent with motivation of FAIR. The following section describe the FAIRification process of the OpenAIRE Guidelines for institutional & thematic Repository Manager.

4.1 Overview

The FAIR maturity model: specification and guidelines[66] of the Research Data Alliance (RDA)[67] (DOI: 10.15497/rda00050[68]) has the aim to specify the objective indicators for the FAIR assessment. The guidelines are intended to assist evaluators to implement the indicators in the evaluation approach or tool they manage.

The evaluation level analyse is done thru the RDA FAIR Data Maturity specification sheet provided at https://www.rd-alliance.org/system/files/FAIR_evaluation_levels_v0.01.xlsx. The result of this evaluation shows the figure below.

In the case of *Findability* indicators

- RDA-F2-01M: Rich metadata is provided to allow discovery
- RDA-F4-01M: Metadata is offered in such a way that it can be harvested and indexed

are implicit covered.

In such a case for *Accessibility* indicators

- RDA-A1-02M: Metadata can be accessed manually (i.e. with human intervention)
- RDA-A1-04M: Metadata is accessed through standardised protocol

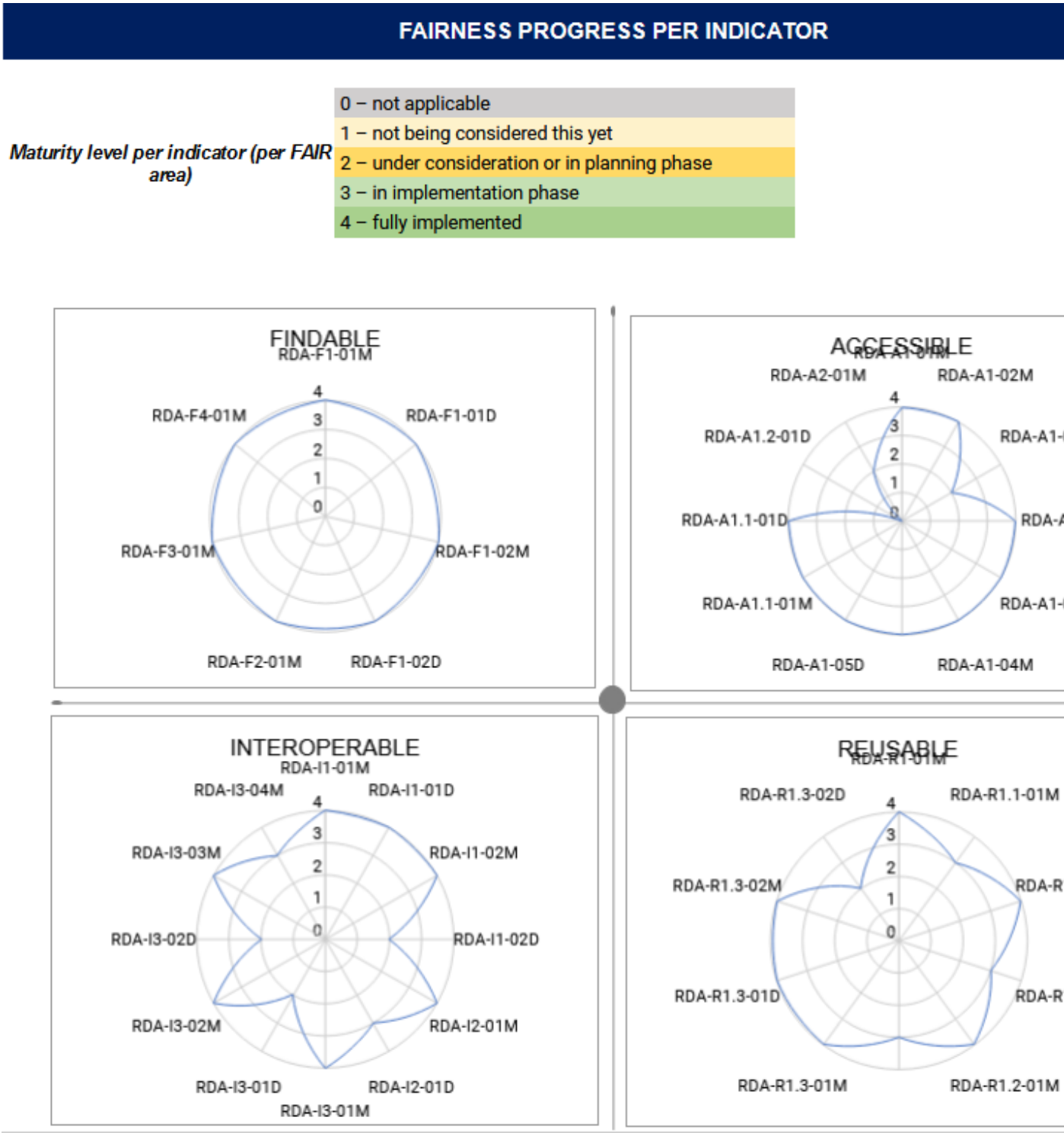
are implemented via the standardized OAI-PMH interface.

And for *Reusability* indicator

- RDA-R1.3-02M: Metadata is expressed in compliance with a machine-understandable community standard

[65] https://www.openaire.eu/history
[66] https://www.rd-alliance.org/group/fair-data-maturity-model-wg/outcomes/fair-data-maturity-model-specification-and-guidelines-0
[67] https://www.rd-alliance.org
[68] https://doi.org/10.15497/rda00050

is implicit covered by the guidelines.



Out of the scope of these guidelines are the indicators:

- RDA-A1-02D: Data can be accessed manually (i.e. with human intervention)
- RDA-A2-01M: Metadata is guaranteed to remain available after data is no longer available
- RDA-I2-01D: Data uses FAIR-compliant vocabularies
- RDA-I3-01D: Data includes references to other data
- RDA-I3-02D: Data includes qualified references to other data

Not directly supported or crosswalks exists are:

- RDA-A1.2-01D: Data is accessible through an access protocol that supports authentication and authorisation
- RDA-I1-02D: Data uses machine-understandable knowledge representation
- RDA-R1.2-01M: Metadata includes provenance information according to community-specific standards
- RDA-R1.2-02M: Metadata includes provenance information according to a cross-community language
- RDA-R1.3-01D: Data complies with a community standard
- RDA-R1.3-02D: Data is expressed in compliance with a machine-understandable community standard