
OpenAIRE Guidelines

Release 3.0

OpenAIRE

January 10, 2017

1	OpenAIRE Guidelines for Literature Repositories	3
1.1	Introduction	3
1.2	Use of OAI-PMH	4
1.3	Use of OAI-DC	5
1.4	Application Profile Overview	6
1.5	Title (M)	7
1.6	Creator (M)	8
1.7	Project Identifier (MA)	9
1.8	Access Level (M)	10
1.9	License Condition (R)	11
1.10	Embargo End Date (MA)	12
1.11	Alternative Identifier (R)	13
1.12	Publication Reference (R)	14
1.13	Dataset Reference (R)	14
1.14	Subject (MA)	15
1.15	Description (MA)	16
1.16	Publisher (MA)	17
1.17	Contributor (R)	18
1.18	Publication Date (M)	19
1.19	Publication Type (M)	20
1.20	Publication Version (R)	22
1.21	Format (R)	22
1.22	Resource Identifier (M)	24
1.23	Source (R)	25
1.24	Language (R)	26
1.25	Relation (O)	27
1.26	Coverage (R)	28
1.27	Audience (R)	29
2	OpenAIRE Guidelines for Data Archives	31
2.1	Introduction	31
2.2	Use of OAI-PMH	32
2.3	Use of DataCite	33
2.4	Application Profile Overview	35
2.5	1. Identifier (M)	37
2.6	2. Creator (M)	37
2.7	3. Title (M)	39
2.8	4. Publisher (M)	39

2.9	5. PublicationYear (M)	40
2.10	6. Subject (R)	40
2.11	7. Contributor (MA/O)	41
2.12	8. Date (M)	44
2.13	9. Language (R)	45
2.14	10. ResourceType (R)	45
2.15	11. AlternateIdentifier (O)	46
2.16	12. RelatedIdentifier (MA)	47
2.17	13. Size (O)	50
2.18	14. Format (O)	50
2.19	15. Version (O)	50
2.20	16. Rights (MA)	51
2.21	17. Description (MA)	52
2.22	18. GeoLocation (O)	53
3	OpenAIRE Guidelines for CRIS Managers based on CERIF-XML	55
4	Participate	57
4.1	Contributing	57
5	OpenAIRE Validator	63
6	Horizon 2020 Open Access Requirements	65
6.1	How the Horizon 2020 Open Access requirements are met	65



Welcome to the OpenAIRE Guidelines. The intention of this is to provide a public space to share OpenAIREs work on interoperability and to engage with the community. The OpenAIRE Guidelines helps repository managers expose publications, datasets and CRIS metadata via the OAI-PMH protocol in order to integrate with OpenAIRE infrastructure.

OpenAIRE Guidelines consists of three guidelines for publication repositories, data archives and CRIS systems respectively:

OpenAIRE Guidelines for Literature Repositories

1.1 Introduction

1.1.1 Aim

The OpenAIRE Guidelines for Literature Repository Managers 3.0 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.

Initially, the requirements of the OpenAIRE infrastructure, and the previous versions of the OpenAIRE Guidelines, were established to support and monitor the implementation of the [FP7 OA pilot](#). However, this version of the Guidelines, according to the expansion of the aims of the OpenAIRE project and infrastructure, has a broader scope. In fact, these guidelines are intended to guide repository manager to expose to the OpenAIRE infrastructure not only EC funded publications, but also other Open Access publications, regardless of their funding.

So, by implementing these Guidelines, repository managers will not only be enabling authors who deposit publications in their repository to fulfill the 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.

According to the ongoing expansion and anticipating the merger of the DRIVER Guidelines into the context of OpenAIRE Guidelines, detailed content from the DRIVER Guidelines 2.0 was adapted and added into this version of the Guidelines. The OpenAIRE Guidelines for Literature Repository Managers 3.0 are now only a part of a set of OpenAIRE Guidelines that also includes the OpenAIRE Guidelines for Data Archive Managers and the OpenAIRE CERIF-XML profile.

1.1.2 What's new

In comparison with previous versions of the Guidelines, this 3.0 version introduces three main changes:

- The OpenAIRE OAI set has been renamed from `ec_fundedresources` to `openaire`.
- New elements for indicating alternative identifiers, relations to other publications (references), and relations to research datasets have been defined.
- The recommendations on how to use the Dublin Core elements have been inherited from the DRIVER Guidelines.

1.1.3 Acknowledgements & Contributors

Editors

- Mathias Loesch (Bielefeld University, Germany)
- Eloy Rodrigues (University of Minho, Portugal)
- Pedro Principe (University of Minho, Portugal)
- Jochen Schirrwagen (Bielefeld University, Germany)

Experts & Reviewers

- Jordan Biserkov (Pensoft, Bulgaria)
- Natasa Bulatovic (Max Planck Digital Library, Germany)
- Paolo Manghi (CNR, Italy)
- Natalia Manola (University of Athens, Greece)
- Najla Rettberg (University of Goettingen, Germany)
- Birgit Schmidt (University of Goettingen, Germany)
- Peter Stanchev (Kettering University, USA)

1.1.4 Versions

- 3.0, April 2013
- 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
 - Compatibility for aggregators; extended Namespace for Project Identification
- 1.1, November 2010 doi:10.5281/zenodo.59206
 - Correction of names and references; addition of license and version statement
- 1.0, July 2010 doi:10.5281/zenodo.59204
 - Initial document

1.2 Use of OAI-PMH

OpenAIRE uses the [OAI-PMH v2.0](#) protocol for harvesting publication metadata.

1.2.1 Metadata Format

OpenAIRE expects metadata to be encoded in the [Dublin Core](#) metadata format (metadataPrefix `oai_dc`). For information on how to use the individual DC fields, please refer to the section “Use of OAI-DC” below.

OpenAIRE OAI Set

For harvesting the records relevant to OpenAIRE, the use of a specific **OAI-Set** at the local repository is *mandatory*. The set must have the following characteristics:

setName	setSpec
OpenAIRE	openaire

Note: A harvester only uses the **setSpec** value to perform selective harvesting. The letters of the setSpec must be in small caps.

Set content

Publications to be inserted in the OpenAIRE set must conform to **at least one** of the following criteria:

- They are available in Open Access (full text with no access restrictions)
- They are the outcome of a funded research project identified by a project identifier (see below) regardless of their access status (see section below on [[Literature Guidelines: Metadata Field Access Level|Application Profile Field Access Level]]).

1.2.2 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 the aggregator has to be encoded for OpenAIRE. In accordance with the [OAI-PMH guidelines](#), the provenance information has to be provided in the `about` element of an OAI record, as displayed in the following example:

```

1 <about>
2   <provenance xmlns="http://www.openarchives.org/OAI/2.0/provenance"
3     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4     xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/provenance
5       http://www.openarchives.org/OAI/2.0/provenance.xsd">
6     <originDescription altered="true" harvestDate="2012-09-17T14:58:36Z">
7       <baseURL>http://dspace.library.uu.nl:8080/dspace-oai/request</baseURL>
8       <identifier>oai:dspace.library.uu.nl:1874/218065</identifier>
9       <datestamp>2012-01-19T12:38:56Z</datestamp>
10      <metadataNamespace>
11        http://www.openarchives.org/OAI/2.0/oai_dc/
12      </metadataNamespace>
13    </originDescription>
14  </provenance>
15 </about>

```

1.3 Use of OAI-DC

1.3.1 OpenAIRE-specific Syntax for OAI-DC

OpenAIRE relies on a specific syntax used in the values of standard Dublin Core metadata fields to identify projects, funders, referenced publications, and datasets. This syntax takes the form of URIs and is defined as the `info:eu-repo` namespace.

This section explains which DC fields to use together with this syntax to correctly mark up the information relevant for OpenAIRE.

Within OpenAIRE the use of the fields is either:

Within OpenAIRE the use of fields are either:

- **Mandatory (M)** = the field must always be present in the metadata record. An empty element is not allowed.
- **Mandatory when applicable (MA)** = when the value of the field can be obtained it must be present in the metadata record.
- **Recommended (R)** = the use of the field is recommended.
- **Optional (O)** = the property may be used to provide complementary information about the resource

A complete example record

```
1 <record>
2   <header>
3     <identifier>oai:repository.example.com:2003292</identifier>
4     <datestamp>2012-11-30T13:40:28Z</datestamp>
5     <setSpec>openaire</setSpec>
6   </header>
7   <metadata>
8     <oai_dc:dc
9       xmlns:dc="http://purl.org/dc/elements/1.1/"
10      xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
11      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
12      xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
13        http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
14       <dc:title>Studies of Unicorn Behaviour</dc:title>
15       <dc:identifier>http://repository.example.org/2003292</dc:identifier>
16       <dc:creator>Jane, Doe</dc:creator>
17       <dc:creator>John, Doe</dc:creator>
18       <dc:description>
19         Lorem ipsum dolor...
20       </dc:description>
21       <dc:subject>info:eu-repo/classification/ddc/590</dc:subject>
22       <dc:subject>Unicorns</dc:subject>
23       <dc:relation>info:eu-repo/grantAgreement/EC/FP7/1234556789/EU//UNICORN</dc:relation>
24       <dc:relation>info:eu-repo/semantics/altIdentifier/eissn/1234-5678</dc:relation>
25       <dc:relation>info:eu-repo/semantics/altIdentifier/pmid/123456789</dc:relation>
26       <dc:relation>info:eu-repo/semantics/altIdentifier/doi/10.1000/182</dc:relation>
27       <dc:relation>info:eu-repo/semantics/reference/doi/10.1234/789.1</dc:relation>
28       <dc:relation>info:eu-repo/semantics/dataset/doi/10.1234/789.1</dc:relation>
29       <dc:rights>info:eu-repo/semantics/openAccess</dc:rights>
30       <dc:rights>http://creativecommons.org/licenses/by-sa/2.0/uk/</dc:rights>
31       <dc:source>Journal Of Unicorn Research</dc:source>
32       <dc:publisher>Unicorn Press</dc:publisher>
33       <dc:date>2013</dc:date>
34       <dc:type>info:eu-repo/semantics/article</dc:type>
35     </oai_dc:dc>
36   </metadata>
37 </record>
```

1.4 Application Profile Overview

This documentation uses the following namespace abbreviation:

- dc: <http://purl.org/dc/elements/1.1/>

OpenAIRE-Field	OAI-DC Element	Refinement by vocabulary
<i>Title (M)</i>	dc:title	
<i>Creator (M)</i>	dc:creator	
<i>Project Identifier (MA)</i>	dc:relation	info:eu-repo/grantAgreement/
<i>Access Level (M)</i>	dc:rights	info:eu-repo/semantics/
<i>License Condition (R)</i>	dc:rights	
<i>Embargo End Date (MA)</i>	dc:date	info:eu-repo/date/embargoEnd/
<i>Alternative Identifier (R)</i>	dc:relation	info:eu-repo/semantics/altIdentifier/
<i>Publication Reference (R)</i>	dc:relation	info:eu-repo/semantics/reference/
<i>Dataset Reference (R)</i>	dc:relation	info:eu-repo/semantics/dataset/
<i>Subject (MA)</i>	dc:subject	
<i>Description (MA)</i>	dc:description	
<i>Publisher (MA)</i>	dc:publisher	
<i>Contributor (R)</i>	dc:contributor	
<i>Publication Date (M)</i>	dc:date	
<i>Publication Type (M)</i>	dc:type	info:eu-repo/semantics/
<i>Publication Version (R)</i>	dc:type	info:eu-repo/semantics/
<i>Format (R)</i>	dc:format	
<i>Resource Identifier (M)</i>	dc:identifier	
<i>Source (R)</i>	dc:source	
<i>Language (R)</i>	dc:language	
<i>Relation (O)</i>	dc:relation	
<i>Coverage (R)</i>	dc:coverage	
<i>Audience (R)</i>	dc:audience	

Application Profile:

1.5 Title (M)

1.5.1 DC Field

dc:title

1.5.2 Usage

Mandatory (M)

1.5.3 Usage Instruction

Preserve the original wording, order and spelling of the resource title. Only capitalize proper nouns. Punctuation need not reflect the usage of the original. Subtitles should be separated from the title by a colon. This instruction would result in Title:Subtitle (i.e. no space). If necessary, repeat this element for multiple titles.

1.5.4 Example

```
<dc:title>Main title:Subtitle</dc:title>
```

1.6 Creator (M)

1.6.1 DC Field

dc:creator

1.6.2 Usage

Mandatory

1.6.3 Usage Instruction

Examples of a Creator include a person, an organization, or a service. If necessary, repeat this element for multiple authors.

Use inverted name, so the syntax will be the following: “surname”, “initials” (“first name”) “prefix”. For example Jan Hubert de Smit becomes

```
<dc:creator>Smit, J.H. (John) de</dc:creator>
```

Within the scope of unqualified DC it is recommended to use a standardised writing style for names, use the writing style used by the publisher when this is available. When that is not available use the encoding of the APA bibliographic writing style as in a reference list when applicable. (outside the scope of Unqualified DC more precise and granular formatting methods are available.)

When initials and first name are both available use this formatting:

```
<dc:creator>Janssen, J. (John)</dc:creator>
```

Generational suffixes (Jr., Sr., etc.) should follow the surname. When in doubt, give the name as it appears, and do not invert. Omit titles (like “Dr”). For example: “Dr. John H. de Smit Jr.” becomes

```
<dc:creator>Smit Jr., J.H. (John) de</dc:creator>
```

In the case of an organization name which clearly includes an organizational hierarchy, list the parts of the hierarchy from largest to smallest, separated by full stops.

For example:

```
<dc:creator>Utrecht University. Department of Computer Sciences</dc:creator>
```

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 resource. Only encode organisations in this element to indicate corporate authorship, not to indicate the affiliation of an individual.

The inclusion of personal and corporate name headings from authority lists constructed according to local or national thesaurus files is optional.

It is recommended to encode thesauri with an URI, for service providers to recognise the thesaurus schema. For example:

```
<dc:creator>urn:NationalOrgThesaurus:nl/234</dc:creator>
```

In cases of lesser responsibility, other than authorship, use dc:contributor. If the nature of the responsibility is ambiguous, recommended best practice is to use dc:publisher for organizations, and dc:creator for individuals.

1.6.4 Do Not Confuse With

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

1.6.5 Since

DRIVER Guidelines v2

1.6.6 Example

```

1 <dc:creator>Evans, R.J.</dc:creator>
2 <dc:creator>Walker Jnr., John</dc:creator>
3 <dc:creator>
4   International Human Genome Sequencing Consortium
5 </dc:creator>
6 <dc:creator>
7   Loughborough University. Department of Computer Science
8 </dc:creator>

```

1.7 Project Identifier (MA)

1.7.1 DC Field

dc:relation

1.7.2 Usage

Mandatory when applicable

1.7.3 Usage Instruction

An authoritative [list of projects](#) is exposed by OpenAIRE through OAI-PMH, and available for all repository managers. Values include the project name and project ID. The projectID equals the Grant Agreement identifier, and is defined by the [info:eu-repo namespace](#) term grantAgreement.

The syntax is:

```

info:eu-repo/grantAgreement/Funder/FundingProgram/ProjectID/
[Jurisdiction]/[ProjectName]/[ProjectAcronym]

```

where:

- Funder refers to the funding organization (e.g., EC for European Commission, WT for Wellcome Trust)
- FundingProgramme refers to a specific programme (e.g., FP7 for Framework Programme Seven, H2020 for Horizon 2020)
- ProjectID refers to a unique identifier in the scope of the funder (and maybe the programme), e.g. a grant agreement number.

- Jurisdiction refers to the authority granted to a formally constituted legal body (e.g. EU for European Union)
- ProjectName contains the full name of the project
- ProjectAcronym contains the project's acronym.

For OpenAIRE compatibility, the elements in square brackets are optional. The three-part namespace is *mandatory when applicable* (`info:eu-repo/grantAgreement/Funder/FundingProgram/ProjectID`), while the six-parts namespace is *recommended*.

When omitting fields in the extended version, the number of fields must be preserved by using /. A correct example for omitting the the ProjectName field would therefore look like this: `EC/FP7/12345/EU//OpenAIREplus`

If any of the field values contains a forward slash (/), it needs to be escaped using URL encoding (%2F). For instance, `My/Project` would be represented as `My%2FProject`.

1.7.4 Since

OpenAIRE Guidelines v1

1.7.5 Example

An example utilizing all six fields:

```
1 <dc:relation>
2   info:eu-repo/grantAgreement/EC/FP7/244909/EU/Making Capabilities Work/WorkAble
3 </dc:relation>
```

An example for omitting the the ProjectName field (note that the number of slashes is correctly preserved):

```
1 <dc:relation>
2   info:eu-repo/grantAgreement/EC/FP7/283595/EU//OpenAIREplus
3 </dc:relation>
```

A correct example using the old three-part form (discouraged):

```
1 <dc:relation>
2   info:eu-repo/grantAgreement/EC/FP7/244909
3 </dc:relation>
```

1.8 Access Level (M)

1.8.1 DC Field

`dc:rights`

1.8.2 Usage

Mandatory

1.8.3 Usage Instruction

Use terms from the [info:eu-repo-Access-Terms](#) vocabulary . The values are:

- `info:eu-repo/semantics/closedAccess`
- `info:eu-repo/semantics/embargoedAccess`
- `info:eu-repo/semantics/restrictedAccess`
- `info:eu-repo/semantics/openAccess`

1.8.4 Example

```
<dc:rights>info:eu-repo/semantics/openAccess</dc:rights>
```

1.9 License Condition (R)

1.9.1 DC Field

`dc:rights`

1.9.2 Usage

Recommended

1.9.3 DCMI Definition

Information about rights held in and over the resource.

1.9.4 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.

1.9.5 Since

OpenAIRE Guidelines v3

1.9.6 Example

Generic usage examples:

```
1 <dc:rights>(c) University of Bath, 2003</dc:rights>
2 <dc:rights>(c) Andrew Smith, 2003</dc:rights>
```

Using Creative Commons right services, makes the usage rights much more clear to the end user. More information see “Use of Intellectual Property Rights”. In this case Andrew Smith referring to <http://creativecommons.org/licenses/by-sa/2.0/uk/>

```
1 <!-- example 1 -->
2 <dc:rights>
3   http://creativecommons.org/licenses/by-sa/2.0/uk/
4 </dc:rights>
```

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.

```
1 <!-- example 2 -->
2 <dc:rights>cc-by-sa, Andrew Smith</dc:rights>
```

The string cc-by-sa provides the licence type in a rough sense. The name is the person or party where the rights apply to.

```
1 <!-- example 3 -->
2 <dc:rights>cc-by-sa, info:eu-repo/dai/nl/344568</dc:rights>
```

or:

```
1 <dc:rights>cc-by-nc-sa, urn:isni:234562-2</dc:rights>
```

Also a Digital Author Identifier (DAI) or International Standard Name Identifier (ISNI) can be used to globally uniquely identify persons and organisations and relate these names with the appropriate rights.

1.10 Embargo End Date (MA)

1.10.1 DC Field

dc:date

1.10.2 Usage

Mandatory when applicable

1.10.3 Usage Instruction

When *Access Level (M)* is set to

- info:eu-repo/semantics/embargoedAccess

the end date of the embargo period must be provided. The corresponding term is defined by

- info:eu-repo/date/embargoEnd/<YYYY-MM-DD>

Encoding of this date should be in the form YYYY-MM-DD conforming to ISO 8601.

1.10.4 Since

OpenAIRE Guidelines v1

1.10.5 Example

```
1 <dc:date>info:eu-repo/date/embargoEnd/2015-12-31</dc:date>
```

1.11 Alternative Identifier (R)

1.11.1 DC Field

```
dc:relation
```

1.11.2 Usage

Recommended

1.11.3 Usage Instruction

List alternative identifiers for this publication that are not the primary identifier (repository splash page), e.g., the DOI of publisher's version, the PubMed/arXiv ID. The term is defined by `info:eu-repo/semantics/altIdentifier/<scheme>/<identifier>` where `<scheme>` must be one of the following:

- `ark` – Archival Resource Key
- `arxiv` – arXiv.org identifier
- `doi` – Digital Object Identifier
- `hdl` – Handle
- `isbn` – International Standard Book Number
- `piissn` – International Standard Serial Number (print version)
- `eissn` – International Standard Serial Number (electronic Version)
- `pmid` – PubMed ID
- `purl` – Persistent Uniform Resource Locator
- `urn` – Uniform Resource Name
- `wos` – Web of Science accession number

1.11.4 Since

OpenAIRE Guidelines v3

1.11.5 Example

```
1 <dc:relation>
2   info:eu-repo/semantics/altIdentifier/doi/10.1234/789.1
3 </dc:relation>
```

1.12 Publication Reference (R)

1.12.1 DC Field

dc:relation

1.12.2 Usage

Recommended

1.12.3 Usage Instruction

Encode links to publications referenced by this publication. The syntax is: `info:eu-repo/semantics/reference/<scheme>/<identifier>` where `<scheme>` must be one of the following:

- `ark` – Archival Resource Key
- `arxiv` – arXiv.org identifier
- `doi` – Digital Object Identifier
- `hdl` – Handle
- `isbn` – International Standard Book Number
- `issn` – International Standard Serial Number
- `pmid` – PubMed ID
- `purl` – Persistent Uniform Resource Locator
- `url` – Uniform Resource Locator
- `urn` – Uniform Resource Name
- `wos` – Web of Science accession number

1.12.4 Since

OpenAIRE Guidelines v3

1.12.5 Example

```
1 <dc:relation>
2   info:eu-repo/semantics/reference/doi/10.1234/789.1
3 </dc:relation>
```

1.13 Dataset Reference (R)

1.13.1 DC Field

dc:relation

1.13.2 Usage

Recommended

1.13.3 Usage Instruction

Encodes links to research datasets connected with this publication. The syntax is: `info:eu-repo/semantics/dataset/<scheme>/<identifier>` where `<scheme>` must be one of the following:

- `ark` – Archival Resource Key
- `doi` – Digital Object Identifier
- `hdl` – Handle
- `purl` – Persistent Uniform Resource Locator
- `url` – Uniform Resource Locator
- `urn` – Uniform Resource Name

1.13.4 Since

OpenAIRE Guidelines v3

1.13.5 Example

```
1 <dc:relation>
2   info:eu-repo/semantics/dataset/doi/10.1234/789.1
3 </dc:relation>
```

1.14 Subject (MA)

1.14.1 DC Field

`dc:subject`

1.14.2 Usage

Mandatory when applicable (MA)

1.14.3 Usage Instruction

In the DC subject element two kinds of values are possible: encode either a keyword or a classification. When both are available use separate occurrences of this element.

Use the first occurrence of the DC element `subject` for a human readable keyword.

In general, choose the most significant and unique words for keywords, avoiding those too general to describe a particular resource. If the subject of the resource is a person or an organization, use the same form of the name as you would if the person or organization were an author, but do not repeat the name in the `dc:creator` element.

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 in a separate element. 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 easily when the schema is “URI-fied” by an authority namespace. When the classification scheme is codified, use a human readable text of the code, preferably in English, directly below the codified element. For example:

```
1 <dc:subject>info:eu-repo/classification/ddc/641</dc:subject>
2 <dc:subject>Anatomy</dc:subject>
```

If no specific classification scheme is used we recommend the Dewey Decimal Classification (DDC). The first 1000 terms is called the Dewey Decimal Classification Summary and can be downloaded at <http://www.oclc.org/dewey/resources/summaries/> if one agrees with the following terms and conditions: <http://www.oclc.org/research/researchworks/ddc/terms.htm>

1.14.4 Do Not Confuse With

- *Publication Type (M)*

1.14.5 Since

DRIVER Guidelines v2

1.14.6 Example

```
1 <dc:subject>
2   polar oceanography; boundary current; masstransport; water masses;
3   halocline; mesoscaleeddies
4 </dc:subject>
5 <dc:subject>Germany--History--1933-1945</dc:subject>
6 <dc:subject>info:eu-repo/classification/ddc/641</dc:subject>
7 <dc:subject>Anatomy</dc:subject>
```

1.15 Description (MA)

1.15.1 DC Field

dc:description

1.15.2 Usage

Mandatory when applicable

1.15.3 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.

1.15.4 Since

DRIVER Guidelines v2

1.15.5 Example

```

1 <dc:description>
2   Foreword [by] Hazel Anderson; Introduction; The scientific heresy:
3   transformation of a society; Consciousness as causal reality [etc]
4 </dc:description>
5
6 <dc:description>
7   A number of problems in quantum state and system identification are
8   addressed.
9 </dc:description>

```

1.16 Publisher (MA)

1.16.1 DC Field

`dc:publisher`

1.16.2 Usage

Mandatory when applicable

1.16.3 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.

1.16.4 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.

1.16.5 Do Not Confuse With

- *Contributor (R)*
- *Creator (M)*

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

1.16.6 Since

DRIVER Guidelines v2

1.16.7 Example

```
1 <dc:publisher>
2   Loughborough University. Department of Computer Science
3 </dc:publisher>
4 <dc:publisher>John Wiley & Sons, Inc. (US)</dc:publisher>
```

1.17 Contributor (R)

1.17.1 DC Field

dc:contributor

1.17.2 Usage

Recommended

1.17.3 DCMI Definition

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

1.17.4 Usage Instruction

Examples of contributors are: a supervisor, editor, technician or data collector.

Personal names should be listed as: see instructions under Creator. A “promotor”, i.e. a professor supervising a student’s work for a doctor’s degree - is considered a contributor of a dissertation in his or her role as promotor/examiner. In less-rich Unqualified DC it is difficult to express all roles in different contexts. In the PhD thesis as a document, the key figures are the author and the supervisor. In the overall PhD process other roles are involved, such as committee members and the Master of Ceremonies, but in Unqualified these roles have to be sacrificed.

In the case of organizations : see instructions under Creator The inclusion of personal and corporate name headings from authority lists constructed according to local or national thesaurus files is optional.

1.17.5 Do Not Confuse With

- *Publisher (MA)*
- *Creator (M)*

The DC element `contributor` describes the scientist(s) that has/have made contributions to the given scientific output, not as a primary creator or (commercial) publisher.

1.17.6 Since

DRIVER Guidelines v2

1.17.7 Example

```

1 <dc:contributor>Evans, R. J.</dc:contributor>
2 <dc:contributor>
3   International Human Genome Sequencing Consortium
4 </dc:contributor>

```

1.18 Publication Date (M)

1.18.1 DC Field

`dc:date`

1.18.2 Usage

Mandatory

1.18.3 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] and follows the YYYY-MM-DD format.

1.18.4 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 than one date fields that serve different purposes. Date of creation, publication, modified, promotion, etc. Unqualified DC is unable to express all these dates, and for the end-user perspective it is confusing to receive more dates from the service provider. The service provider should make a choice what date- field to pick. Preferably in the end-users perspective the most logical and meaningful date will be the date of publication. To reduce the ambiguity of having a number of date fields without qualifiers, we recommend to reduce the number of fields and present the most meaningful date to the service provider. In most cases this is the date of the publication. In other cases this is the date of promotion of a PhD degree.

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.

1.18.5 Since

DRIVER Guidelines v2

1.18.6 Example

```
1 <dc:date>2000-12-25</dc:date>
2 <dc:date>1978-02</dc:date>
3 <dc:date>1650</dc:date>
```

1.19 Publication Type (M)

1.19.1 DC Field

`dc:type`

1.19.2 Usage

Publication type is used for the following purposes:

- **Mandatory (M):** Publication type (controlled): to indicate the type of publication based on the `info:eu-repo-publication-type-terms` vocabulary.
- **Optional (O):** Publication type (free): to indicate the type of publication based on a local repository vocabulary.

1.19.3 DCMI Definition

The type of scientific output the resource is a manifestation of. In the DC element type the kind of dissemination, or the intellectual and/or content type of the resource is described. It is used to explain to the user what kind of resource he is looking at. Is it a book or an article. Was it written for internal or external use, etc.

1.19.4 Usage Instruction

Publication types (controlled):

The first occurrence of the DC Element type is mandatory and should be used for the type indication of the scientific output based on the `info:eu-repo` publication type vocabulary:

- `info:eu-repo/semantics/article`
- `info:eu-repo/semantics/bachelorThesis`
- `info:eu-repo/semantics/masterThesis`
- `info:eu-repo/semantics/doctoralThesis`
- `info:eu-repo/semantics/book`
- `info:eu-repo/semantics/bookPart`
- `info:eu-repo/semantics/review`
- `info:eu-repo/semantics/conferenceObject`
- `info:eu-repo/semantics/lecture`
- `info:eu-repo/semantics/workingPaper`
- `info:eu-repo/semantics/preprint`
- `info:eu-repo/semantics/report`
- `info:eu-repo/semantics/annotation`
- `info:eu-repo/semantics/contributionToPeriodical`
- `info:eu-repo/semantics/patent`
- `info:eu-repo/semantics/other`

Publication types (free text):

The second occurrence of the DC Element type is optional and should be used for the subtype indication of the scientific output.

1.19.5 Do Not Confuse With

- *Format (R)*

DC element *type* describes the kind of academic output the resource is a representation of. DC element ‘format’ describes the media type of this resource.

1.19.6 Since

DRIVER Guidelines v2

1.19.7 Example

```
<dc:type>info:eu-repo/semantics/article</dc:type>
```

1.20 Publication Version (R)

1.20.1 DC Field

dc:type

1.20.2 Usage

Recommended: Version (controlled): to indicate the status in the publication process.

1.20.3 DCMI Definition

1.20.4 Usage Instruction

Version (controlled):

be used for the version of the scientific output based on the DRIVER-version `info:eu-repo` version terms. vocabulary. Use exact text as shown in the list below. For more information about the version model see <http://www.lse.ac.uk/library/versions/>.

- `info:eu-repo/semantics/draft`
- `info:eu-repo/semantics/submittedVersion`
- `info:eu-repo/semantics/acceptedVersion`
- `info:eu-repo/semantics/publishedVersion`
- `info:eu-repo/semantics/updatedVersion`

1.20.5 Since

DRIVER Guidelines v2

1.20.6 Example

```
<dc:type>info:eu-repo/semantics/publishedVersion</dc:type>
```

1.21 Format (R)

1.21.1 DC Field

dc:format

1.21.2 Usage

Recommended

1.21.3 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).

1.21.4 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>`

1.21.5 Do Not Confuse With

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

DC element `format` describes the media type of this resource. DC element `type` describes the kind of academic output the resource is a representation of. `dc:identifier` is used to represent manifestations of digital resources.

1.21.6 Since

DRIVER Guidelines v2

1.21.7 Example

```

1 <dc:format>video/quicktime</dc:format>
2 <dc:format>application/pdf</dc:format>
3 <dc:format>application/xml</dc:format>
4 <dc:format>application/xhtml+xml</dc:format>
5 <dc:format>application/html</dc:format>
6 <dc:format>application/vnd.oasis.opendocument.text</dc:format>

```

1.22 Resource Identifier (M)

1.22.1 DC Field

`dc:identifier`

1.22.2 Usage

Mandatory

1.22.3 DCMI Definition

An unambiguous reference to the resource within a given context.

1.22.4 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.

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

Smart practice:

- use stable URL's
- provide every identifier one can find about the publication.
 - (URL,DOI,URN:NBN,ISBN,ISSN,etc.)
- place the “most appropriate” identifier in the form of a URL at the top of the list of Identifiers. In almost all cases this is the one that will be used by a service provider to let an end-user refer to. This can be a link to a jump-off page or a direct link to the file. Also this can be a direct URL, or a redirection URL, like PURL, HANDLE or other international resolution mechanisms.

1.22.5 Do Not Confuse With

- *Relation (O)* (Use `dc:relation` to refer from one version of the resource to another.)
- *Source (R)* (Use `dc:source` for bibliographic citation of the originating resource.)

1.22.6 Since

DRIVER Guidelines v2

1.22.7 Example

In this example the identifiers are sorted so that URLs are given first. The first URL will be considered as “most appropriate” and will be used in e.g. OpenAIRE to let an end-user redirect to. In this case 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:

```

1 <dc:identifier>http://hdl.handle.net/1234/5628 </dc:identifier>
2 <dc:identifier>http://arno.unimaas.nl/show.cgi?fid=5628 </dc:identifier>
3 <dc:identifier>http://n2t.info/urn:nbn:nl:ui:14123456789</dc:identifier>
4 <dc:identifier>urn:nbn:nl:ui:13-123456789</dc:identifier>
5 <dc:identifier>urn:isbn:123456789</dc:identifier>
6 <dc:identifier>info:doi:10-123456789</dc:identifier>

```

1.23 Source (R)

1.23.1 DC Field

dc:source

1.23.2 Usage

Recommended

1.23.3 DCMI Definition

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

1.23.4 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 Relation. 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/>).

1.23.5 Do Not Confuse With

- *Relation (O)*
- *Resource Identifier (M)*

1.23.6 Since

DRIVER Guidelines v2

1.23.7 Example

```
1 <dc:source>Ecology Letters (1461023X) vol.4 (2001) </dc:source>  
2 <dc:source>ISSN: 0928-0987</dc:source>
```

1.24 Language (R)

1.24.1 DC Field

dc:language

1.24.2 Usage

Recommended

1.24.3 DCMI Definition

A language of the intellectual content of the resource.

1.24.4 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..

Recommended: 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.

1.24.5 Do Not Confuse With

- Country codes (ISO 3166-1)

1.24.6 Since

DRIVER Guidelines v2

1.24.7 Example

```

1 <dc:language>eng</dc:language>
2 <dc:language>deu</dc:language>
3 <dc:language>nld</dc:language>
4 <dc:language>nld/dut</dc:language>
5 <dc:language>dut</dc:language>
6 <dc:language>nl</dc:language>

```

1.25 Relation (O)

1.25.1 DC Field

dc:relation

1.25.2 Usage

Optional

1.25.3 DCMI Definition

The reference to a related resource.

1.25.4 Usage Instruction

For OpenAIRE-specific use, see sections on *Project Identifier (MA)*, *Alternative Identifier (R)*, *Publication Reference (R)*, and *Dataset Reference (R)*.

Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system. The DC element `relation` can be used to indicate different kinds of relations between several metadata records. If relations between metadata records are made visible by using metadata the following holds for the distinction between versions (author version and publisher version, preprint, postprint, etc.):

- A metadata record is self-contained
- Different manifestations of one and the same resource (an instance of scientific output that can be described with exactly the same bibliographic metadata, except for the DC element `format`) are linked to one single metadata record using `dc:relation`.

Changes in the metadata other than the DC element `format` leads to creating a new metadata record of this new instance of scientific output which meets all requirements formulated in this document and has a value in the DC element `relation`.

1.25.5 Do Not Confuse With

- *Resource Identifier (M)*
- *Source (R)*

1.25.6 Since

DRIVER Guidelines v2

1.25.7 Example

Generic use:

```
1 <dc:relation>http://hdl.handle.net/10</dc:relation>
```

The value of `dc:relation` is the identifier of the other document.

Linking two documents:

```
1 <!-- Document A -->
2 <dc:type>info:eu-repo/semantics/submittedVersion</dc:type>
3 <dc:identifier> http://hdl.handle.net/10</dc:identifier>
4 <dc:relation>http://hdl.handle.net/20</dc:relation>
```

```
1 <!-- Document B -->
2 <dc:type>info:eu-repo/semantics/acceptedVersion</dc:type>
3 <dc:identifier> http://hdl.handle.net/20</dc:identifier>
4 <dc:relation>http://hdl.handle.net/10</dc:relation>
```

1.26 Coverage (R)

1.26.1 DC Field

`dc:coverage`

1.26.2 Usage

Recommended

1.26.3 DCMI Definition

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

1.26.4 Usage Instruction

Recommended best practice is to select the value from a controlled vocabulary (for example, the Getty Thesaurus of Geographic Names or TGN) and that, where appropriate, named places or time periods be used in preference to numeric identifiers as, for example, sets of coordinates or date ranges. If necessary, repeat this element to encode multiple locations or periods.

1.26.5 Since

DRIVER Guidelines v2

1.26.6 Example

Example Spatial: ISO 3166:

```
1 <dc:coverage>NL</dc:coverage>
```

Example Spatial: BOX:

```
1 <dc:coverage>
2   name=Western Australia; northlimit=-13.5; southlimit=-35.5;
3   westlimit=112.5; eastlimit=129
4 </dc:coverage>
```

Note: The syntax used here is provisional, and is currently under review as part of the DCMI work on recommending coordinated syntax recommendations for HTML, XML, and RDF. These recommendations and minor editorial changes in this document can be expected to take place in the near future. Point <http://dublincore.org/documents/dcmi-point/>

1.27 Audience (R)

1.27.1 DC Field

dc:audience

1.27.2 Usage

Recommended

1.27.3 DCMI Definition

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

1.27.4 Usage Instruction

A class of entity may be determined by the creator or the publisher or by a third party. On the U.S. Department of Education, Metadata Reference site, an example is given of audiences: <http://www.ed.gov/admin/reference/index.jsp> :

- 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

1.27.5 Since

DRIVER Guidelines v2

1.27.6 Example

```
1 <dc:audience>Researchers</dc:audience>  
2 <dc:audience>Students</dc:audience>
```

OpenAIRE Guidelines for Data Archives

2.1 Introduction

2.1.1 Aim

OpenAIRE gathers together research output related to European funding streams, with the aim of supporting open science and tracking research impact. This content consists of open access publications and related contextual information such as datasets, supplementary material, and research/project information.

Two other sets of OpenAIRE guidelines exist for repository managers:

- *OpenAIRE Guidelines for Literature Repositories*
- *OpenAIRE Guidelines for CRIS Managers based on CERIF-XML*

The OpenAIRE Guidelines for Data Archive Managers 2.0 will provide instruction for data archive managers to expose their metadata in a way that is compatible with the OpenAIRE infrastructure. The metadata from data archives should be included in the OpenAIRE information space, and exposed when data are related to an open access publication e.g. a dataset cited by an article. To access a pdf version the guidelines see here: <http://dx.doi.org/10.5281/zenodo.6918>

By implementing the OpenAIRE Guidelines, data archive managers are facilitating the creation of enhanced publications and building the stepping-stones for a linked data infrastructure for research.

Exposure and visibility of content from a range of European repositories will be significantly increased when a common and interoperable approach is taken as well as adhering to existing guidelines. OpenAIRE is happy to assist in adherence to these guidelines. This compatibility will lead to future interoperability between research infrastructures, and structured metadata is of benefit to individual data repositories and the scholarly community at large.

2.1.2 Acknowledgements & Contributors

Editors

- Mikael K. Elbæk (Technical University of Denmark)
- Lars Holm Nielsen (CERN, Switzerland)

Experts & Reviewers

- Samuele Kaplun (CERN, Switzerland)
- Paolo Manghi (CNR, Italy)
- Natalia Manola (University of Athens, Greece)
- Jochen Schirrwagen (University of Bielefeld, Germany)

- Birgit Schmidt (University of Goettingen, Germany)
- Mathias Lösch (University of Bielefeld, Germany)
- Frauke Ziedorn (DataCite, Germany)
- Pedro Principe (University of Minho, Portugal)
- Eloy Rodrigues (University of Minho, Portugal)
- Najla Rettberg (University of Goettingen, Germany)

2.1.3 Versions

- 2.0 April 2014
 - Updated to DataCite Metadata Schema v3.0
- 1.0 December 2012
 - Initial document

2.2 Use of OAI-PMH

OpenAIRE uses the OAI-PMH v2.0 protocol for harvesting dataset metadata.

2.2.1 Metadata Format

OpenAIRE expects metadata to be encoded in the DataCite metadata format (metadataPrefix oai_datacite). For information on how to use the individual DataCite properties, please refer to [Application Profile Overview](#).

2.2.2 OpenAIRE OAI Set

A specific OAI-Set at the local repository should be configured to select records relevant to OpenAIRE for harvesting. The following characteristics of the set are *recommended*:

setName	setSpec
OpenAIRE	openaire_data

Note: A harvester only uses the **setSpec** value to perform selective harvesting. The letters of the setSpec must be in small caps.

2.2.3 Set content

The specific content of the `openaire_data` set is to be determined at the local repository. OpenAIRE will **harvest all datasets** from the `openaire_data` set, but will **only expose datasets fulfilling at least one** of the following criteria in the OpenAIRE portal:

- The dataset is outcome of a funded research project identified by a project identifier (see section “Funding information” above).
- The dataset is linked with a publication in the OpenAIRE information space (see section “Related publications and datasets information” above).

Both criteria above may be automatically inferred by the OpenAIRE infrastructure.

Specifically a data repository may insert a dataset without funding information or link to a related publication in the `openaire_data` set. If later a publication harvested from a literature repository links to the dataset, the dataset will be exposed in the OpenAIRE portal. If either the literature repository or the data repository explicitly links the publication and dataset (see section "Related publications and datasets information"), the dataset will normally be exposed in the OpenAIRE portal **within 1-2 days after harvesting** (repositories are harvested on average once a week). If the OpenAIRE infrastructure automatically has to infer the link between the publication and the dataset, it may take **up to a month after harvesting** before the dataset is exposed in the OpenAIRE portal.

2.3 Use of DataCite

OpenAIRE has adopted the DataCite Metadata Schema as the basis for harvesting and importing metadata about datasets from data archives.

The core mission of DataCite is to build and maintain a sustainable framework that makes it possible to cite data through the use of persistent identifiers, DOIs.

OpenAIRE shares the goal of the DataCite Metadata Schema - to provide a domain agnostic metadata schema and provide interoperability through a small number of properties - making interoperability possible in the simplest manner possible and as a result keep the technical barriers for implementation as low as possible.

We would like to thank DataCite for their kind support and help making these OpenAIRE Guidelines for Data Archive Managers possible.

2.3.1 What's different

In this document you will find the needed properties to make your data archive compatible with the OpenAIRE infrastructure. OpenAIRE has adopted the DataCite Metadata Schema v3.1 with some minor adjustments.

- OpenAIRE accepts other persistent identifier schemes and not only a DOI in *1.1 identifierType (M)*.
- OpenAIRE recommends exporting links to related publications and datasets (i.e. *12. RelatedIdentifier (MA)* property and attributes are *mandatory when applicable*).
- OpenAIRE recommends using the *7. Contributor (MA/O)* property to relate a dataset to funding information.
- DataCite optional properties *8. Date (M)* and *17. Description (MA)* are *mandatory* in OpenAIRE.
- OpenAIRE enforces an encoding scheme on DataCite property *16. Rights (MA)*.

2.3.2 Access rights and license information

OpenAIRE uses the access rights to enable a better user experience by declaring the access rights clear and explicit in the portal. Access rights are specified using the *16. Rights (MA)* property. Please see example of encoding scheme in the section above.

An example:

```

1 <rightsList>
2   <rights rightsURI="info:eu-repo/semantics/openAccess" />
3 </rightsList>
```

OpenAIRE further recommends including license information if available:

```

1 <rightsList>
2   <rights rightsURI="info:eu-repo/semantics/openAccess" />
3   <rights rightsURI="http://creativecommons.org/licenses/by/4.0/">
4     Creative Commons Attribution 4.0 International
5   </rights>
6 </rightsList>

```

2.3.3 Funding information

OpenAIRE have a specific use of the *7. Contributor (MA/O)* property to identify links to funding streams.

One of OpenAIRE's main goals is to link research output to (EC) research funding. The following application of the contributor property allows unique and persistent identification of the funder who has funded wholly or partly the dataset described.

The following properties are mandatory when applicable to provide funding information: *7. Contributor (MA/O)*, *7.1 contributorType (MA/O)*, *7.2 contributorName (MA/O)*, *7.3 nameIdentifier (MA/O)* and *7.3.1 nameIdentifierScheme (MA/O)*.

An example for linking a research output to the OpenAIREplus FP7 project:

```

1 <contributor contributorType="Funder">
2   <contributorName>European Commission</contributorName>
3   <nameIdentifier nameIdentifierScheme="info">
4     info:eu-repo/grantAgreement/EC/FP7/282896
5   </nameIdentifier>
6 </contributor>

```

2.3.4 Related publications and datasets information

OpenAIRE **harvests all datasets** from a data repository, but **exposes** only certain datasets in the OpenAIRE portal. See the *Set content* for specific details of which datasets are exposed.

For example, datasets related to publication will be exposed in the OpenAIRE portal. The link between the dataset and publication may be explicit defined, as described in this section, or automatically inferred by the OpenAIRE infrastructure.

If the link is explicit defined, the dataset will be exposed in the OpenAIRE portal **within 1-2 days after harvesting** (a repository is harvested once a week on average). If the link is automatically inferred by the OpenAIRE infrastructure it may take **up to a month after harvesting** before the dataset is exposed in the OpenAIRE portal. It is thus *mandatory when applicable* to provide links to related publications and datasets when these links are available in the repository, and thereby ensure faster exposure of the dataset in the OpenAIRE portal.

Related identifiers

DataCite Metadata Schema allows linking publications and datasets by use of persistent identifiers to uniquely identify the resource being described (A) typically a dataset but not limited to that, and the related resource (B) in the case of OpenAIRE typically a publication or a dataset.

Related publications/datasets must have the following properties: *12. RelatedIdentifier (MA)*, *12.1 relatedIdentifierType (M)* and *12.2 relationType (M)*.

An example:

```

1 <relatedIdentifier
2   relatedIdentifierType="DOI"
3   relationType="IsCitedBy">
4 10.1234/bar
5 </relatedIdentifier>

```

2.3.5 Embargo date information

For OpenAIRE two main types of dates are relevant. When the data were made available, published or uploaded to a formal database, this is the date the data were ‘‘Issued’’.

Sometimes data may be embargoed for a period; this information should be managed by the data provider and expressed by exporting an ‘‘Available’’ date to indicate the end of an embargo period and an ‘‘Accepted’’ date to indicate the start of an embargo period.

An example:

```

1 <dates>
2   <date dateType="Issued">2005-04-05</date>
3 </dates>

```

An embargo example:

```

1 <dates>
2   <date dateType="Accepted">2011-12-01</date>
3   <date dateType="Available">2012-12-01</date>
4 </dates>

```

2.4 Application Profile Overview

OpenAIRE builds on the DataCite Metadata Schema v3.1 by making some of the otherwise optional DataCite properties mandatory, as well as enforcing specific encoding schemes on the values of some DataCite properties. The table below details the differences from the DataCite Metadata Schema.

Within OpenAIRE the use of fields are either:

- **Mandatory (M)** = the field must always be present in the metadata record. An empty element is not allowed.
- **Mandatory when applicable (MA)** = when the value of the field can be obtained it must be present in the metadata record.
- **Recommended (R)** = the use of the field is recommended.
- **Optional (O)** = the property may be used to provide complementary information about the resource

The recommended status is made primarily to encourage users to input certain properties when creating a metadata record to enhance services.

Important: The table below details only differences from the DataCite Metadata Schema v3.1. Please refer to the DataCite Metadata Schema v3.1 for definition of properties and their encoding schemes: <http://schema.datacite.org/meta/kernel-3/index.html>

Property	Comment
<i>1. Identifier (M)</i>	
<i>1.1 identifierType (M)</i>	Unlike DataCite, OpenAIRE allows for DOIs and other types of identifiers.

Property	Comment
2. Creator (M)	
2.1 creatorName (M)	
2.2 nameIdentifier (R)	OpenAIRE recommends including a nameIdentifier such as an ORCID or a ISNI if available.
2.2.1 nameIdentifierScheme (R)	
2.2.2 schemeURI (R)	
2.3 affiliation (R)	
3. Title (M)	
3. titleType (O)	
4. Publisher (M)	
5. PublicationYear (M)	
6. Subject (R)	
6.1 subjectScheme (O)	
6.2 schemeURI (O)	
7. Contributor (MA/O)	OpenAIRE uses this property and sub-properties to allow unique and persistent identification of contributors.
7.1 contributorType (MA/O)	
7.2 contributorName (MA/O)	
7.3 nameIdentifier (MA/O)	
7.3.1 nameIdentifierScheme (MA/O)	
7.3.2 schemeURI (O)	
7.4 affiliation (O)	
8. Date (M)	Mandatory property in OpenAIRE instead of <i>recommended</i> in DataCite. See <i>Embargo date info</i> .
8.1 dateType (M)	
9. Language (R)	
10. ResourceType (R)	
10.1 resourceTypeGeneral (R)	
11. AlternateIdentifier (O)	
11.1 alternateIdentifierType (O)	
12. RelatedIdentifier (MA)	Mandatory when applicable property in OpenAIRE instead of <i>recommended</i> in DataCite. See <i>Embargo date info</i> .
12.1 relatedIdentifierType (M)	
12.2 relationType (M)	
12.3 relatedMetadataScheme (O)	
12.1 schemeURI (O)	
12.1 schemeType (O)	
13. Size (O)	
14. Format (O)	
15. Version (O)	
16. Rights (MA)	Mandatory when applicable property in OpenAIRE instead of <i>recommended</i> in DataCite. See <i>Embargo date info</i> .
16.1 rightsURI (MA)	
17. Description (MA)	Mandatory when applicable property in OpenAIRE instead of <i>recommended</i> in DataCite.
17.1 descriptionType (MA)	
18. GeoLocation (O)	
18.1 geoLocationPoint (O)	
18.2 geoLocationBox (O)	
18.3 geoLocationPlace (O)	

Application Profile:

2.5 1. Identifier (M)

The Identifier is a unique string that identifies a resource (occurrences: 1).

Allowed values, examples, other constraints

Format should be “10.1234/foo”

2.5.1 1.1 identifierType (M)

The type of the Identifier (occurrences: 1).

Allowed values, examples, other constraints

Controlled list values (DataCite)

- DOI

Controlled list values (OpenAIRE)

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

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

2.5.2 Example

```

1 <identifier identifierType="DOI">
2   10.1594/WDCC/CCSRNIES_SRES_B2
3 </identifier>
```

2.6 2. Creator (M)

The main researchers involved in producing the data, or the authors of the publication, in priority order (occurrences: 1-n).

Allowed values, examples, other constraints

May be a corporate/institutional or personal name. Note: DataCite infrastructure supports up to between 8000 - 10000 names. For name lists above that size, consider attribution via linking to the related metadata.

2.6.1 2.1 creatorName (M)

The name of the creator (occurrences: 1).

Allowed values, examples, other constraints

Examples: Smith, John; Miller, Elizabeth

The personal name format should be: family, given. Non-roman names may be transliterated according to the [ALA-LC](#) schemes.

2.6.2 2.2 nameIdentifier (R)

Uniquely identifies an individual or legal entity, according to various schemes (occurrences: 0-1).

Allowed values, examples, other constraints

The format is dependent upon scheme.

Note: OpenAIRE recommends including a nameIdentifier such as an ORCID or a ISNI if available.

2.2.1 nameIdentifierScheme (R)

The name of the name identifier scheme (occurrences: 1).

Allowed values, examples, other constraints

If nameIdentifier is used, nameIdentifierScheme is mandatory. Examples: [ORCID](#), [ISNI](#),

2.2.2 schemeURI (R)

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

Allowed values, examples, other constraints

Examples:

- <http://www.isni.org>
- <http://orcid.org>

2.6.3 2.3 affiliation (R)

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

Allowed values, examples, other constraints

Free text.

2.6.4 Example

```
1 <creators>
2   <creator>
3     <creatorName>Miller, John</creatorName>
4   </creator>
5   <creator>
6     <creatorName>Smith, Jane</creatorName>
7     <affiliation>OpenAIRE</affiliation>
8     <nameIdentifier nameIdentifierScheme="ISNI"
9       schemeURI="http://www.isni.org">
10    1422 4586 3573 0476
```

```

11 </nameIdentifier>
12 </creator>
13 </creators>

```

2.7 3. Title (M)

A name or title by which a resource is known (occurrences: 1-n).

Allowed values, examples, other constraints

Free text.

2.7.1 3. titleType (O)

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

Allowed values, examples, other constraints

Controlled List Values:

- Alternative Title
- Subtitle
- TranslatedTitle

Example

```

1 <titles>
2 <title>
3   National Institute for Environmental Studies and Center
4   for Climate System Research Japan
5 </title>
6 <title titleType="Subtitle">A survey</title>
7 </titles>

```

2.8 4. Publisher (M)

The name of the entity that holds, archives, publishes prints, distributes, releases, issues, or produces the resource. This property will be used to formulate the citation, so consider the prominence of the role (occurrences: 1).

Allowed values, examples, other constraints

Examples: World Data Center for Climate (WDCC); GeoForschungsZentrum Potsdam (GFZ); Geological Institute, University of Tokyo

2.8.1 Example

```

1 <publisher>World Data Center for Climate (WDCC)</publisher>

```

2.9 5. PublicationYear (M)

The year when the data was or will be made publicly available. (occurrences: 1)

Allowed values, examples, other constraints

Year in the form: YYYY

If an embargo period has been in effect, use the date when the embargo period ends.

In the case of datasets, “publish” is understood to mean making the data available on a specific date to the community of researchers.

If there is no standard publication year value, use the date that would be preferred from a citation perspective.

Note: See 8. *Date (M)* and 8.1 *dateType (M)* for the recommended application of embargo period in OpenAIRE.

2.9.1 Additional guidance

The year when the data was or will be made publicly available. In the case of datasets, “publish” is understood to mean making the data available on a specific date to the community of researchers.

- If that date cannot be determined, use the date of registration.
- If an embargo period has been in effect, use the date when the embargo period ends.
- If there is no standard publication year value, use the date that would be preferred from a citation perspective.

In the case of a digitized version of a physical object

If the DOI is being used to identify a digitised version of an original item, the recommended approach is to supply the 5. *PublicationYear (M)* for the digital version and not the original object.

The 3. *Title (M)* field may be used to convey the approximate or known date of the original object. Other metadata properties available for additional date information about the object include: 6. *Subject (R)* and 17. *Description (MA)*. However, only 3. *Title (M)* will be part of the citation.

If the DOI is being used to identify the original object and the publication date of this is unknown or not standard then we recommend the following:

- Use the date that would be preferred from a citation perspective.

2.9.2 Example

```
<publicationYear>2004</publicationYear>
```

2.10 6. Subject (R)

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

Allowed values, examples, other constraints

Free text.

2.10.1 6.1 subjectScheme (O)

The name of the subject scheme or classification code or authority if one is used (occurrences: 01).

Allowed values, examples, other constraints

Free text.

2.10.2 6.2 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/>

2.10.3 Example

```

1 <subjects>
2 <subject>Earth sciences and geology</subject>
3 <subject subjectScheme="DDC" schemeURI="http://dewey.info/">
4 551 Geology, hydrology, meteorology
5 </subject>
6 </subjects>

```

2.11 7. Contributor (MA/O)

The institution or person responsible for collecting, managing, distributing, or otherwise contributing to the development of the resource (occurrences: 0-n).

Allowed values, examples, other constraints

Note: DataCite infrastructure supports up to between 8000 - 10000 names. For name lists above that size, consider attribution via linking to the related metadata.

Note: OpenAIRE uses this property to allow unique and persistent identification of the funder who has funded wholly or partly the dataset described. This does not exclude also using this property for additional contributors as defined by DataCite Metadata Schema v3.1.

The property and sub-properties are only *mandatory when applicable (MA)* for describing funding information. Further use is *optional (O)*.

Please refer to the section *Funding information* for a complete example of the use of *7. Contributor (MA/O)*, *7.1 contributorType (MA/O)*, *7.2 contributorName (MA/O)*, *7.3 nameIdentifier (MA/O)* and *7.3.1 nameIdentifierScheme (MA/O)* to provide funding information.

2.11.1 7.1 contributorType (MA/O)

The type of contributor of the resource (occurrences: 1).

Allowed values, examples, other constraints

If 7. *Contributor (MA/O)* is used, then 7.1 *contributorType (MA/O)* is mandatory.

Controlled List Values:

- ContactPerson
- DataCollector
- DataCurator
- DataManager
- Distributor
- Editor
- Funder
- HostingInstitution
- Producer
- ProjectLeader
- ProjectManager
- ProjectMember
- RegistrationAgency
- RegistrationAuthority
- RelatedPerson
- Researcher
- ResearchGroup
- RightsHolder
- Sponsor
- Supervisor
- WorkPackageLeader
- Other

2.11.2 7.2 contributorName (MA/O)

The name of the contributor (occurrences: 1).

Allowed values, examples, other constraints

If 7. *Contributor (MA/O)* is used, then 7.2 *contributorName (MA/O)* is mandatory.

Examples: Patel, Emily; Doe, John

The personal name format may be: family, given. Non-roman names should be transliterated according to the [ALA-LC](#) schemes.

Note: Applicable only when contributorType is “Funder”:

Name of the funding entity. Example for European Commission funded research use `European Commission`, or for Wellcome Trust funded research use `Wellcome Trust`. Specifically do not use the project acronym.

2.11.3 7.3 nameIdentifier (MA/O)

Uniquely identifies an individual or legal entity, according to various schemes (occurrences: 0-1).

Allowed values, examples, other constraints

The format is dependent upon scheme.

Note: Applicable only when contributorType is “Funder”:

An authoritative list of projects is exposed by OpenAIRE through OAI-PMH, and available for all repository managers. Values include the project name and project ID. The projectID equals the Grant Agreement identifier, and is defined by the `info:eu-repo namespace` term `grantAgreement`.

The syntax is:

```
info:eu-repo/grantAgreement/Funder/FundingProgram/ProjectID/
[Jurisdiction]/[ProjectName]/[ProjectAcronym]
```

where:

- Funder refers to the funding organization (e.g., EC for European Commission, WT for Wellcome Trust)
- FundingProgramme refers to a specific programme (e.g., FP7 for Framework Programme Seven, H2020 for Horizon 2020)
- ProjectID refers to a unique identifier in the scope of the funder (and maybe the programme), e.g. a grant agreement number.
- Jurisdiction refers to the authority granted to a formally constituted legal body (e.g. EU for European Union)
- ProjectName contains the full name of the project
- ProjectAcronym contains the project’s acronym.

For OpenAIRE compatibility, the elements in square brackets are optional. The three-part namespace is *mandatory when applicable* (`info:eu-repo/grantAgreement/Funder/FundingProgram/ProjectID`), while the six-parts namespace is *recommended*.

When omitting fields in the extended version, the number of fields must be preserved by using /. A correct example for omitting the the `ProjectName` field would therefore look like this: `EC/FP7/12345/EU//OpenAIREplus`

If any of the field values contains a forward slash (/), it needs to be escaped using URL encoding (%2F). For instance, `My/Project` would be represented as `My%2FProject`.

2.11.4 7.3.1 nameIdentifierScheme (MA/O)

The name of the name identifier scheme (occurrences: 1).

Allowed values, examples, other constraints

If *7.3 nameIdentifier (MA/O)* is used, *7.3.1 nameIdentifierScheme (MA/O)* is mandatory.

Examples: ORCID, ISNI, FundRef.

Note: Applicable only when contributorType is “Funder”:

Controlled List Values:

- info
-

2.11.5 7.3.2 schemeURI (O)

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

Allowed values, examples, other constraints

Examples:

- <http://www.isni.org>
- <http://orcid.org>
- <http://www.crossref.org/fundref/>

2.11.6 7.4 affiliation (O)

The organizational or institutional affiliation of the contributor (occurrences: 0-n).

Allowed values, examples, other constraints

Free text.

2.11.7 Example

```
1 <contributor contributorType="Funder">
2   <contributorName>European Commission</contributorName>
3     <nameIdentifier nameIdentifierScheme="info">
4       info:eu-repo/grantAgreement/EC/FP7/282896
5     </nameIdentifier>
6 </contributor>
```

2.12 8. Date (M)

Different dates relevant to the work (occurrences: 0-n).

Allowed values, examples, other constraints

YYYY, YYYY-MM-DD, YYYY-MM-DDThh:mm:ssTZD or any other format or level of granularity described in W3CDTF. Use RKMS-ISO8601 standard for depicting date ranges. Example: 2004-03-02/2005-06-02

Note: *Mandatory* property in OpenAIRE instead of *recommended* in DataCite.

2.12.1 8.1 dateType (M)

The type of date (occurrences: 1).

Allowed values, examples, other constraints

If Date is used, dateType is mandatory.

Controlled List Values:

- Accepted
- Available
- Copyrighted

- Collected
- Created
- Issued
- Submitted
- Updated
- Valid

Note: Use “Issued” for the date the resource is published or distributed. To indicate the end of an embargo period, use “Available”. To indicate the start of an embargo period, use “Accepted”.

DataCite v3.1 further recommends use of “Created” and “Submitted”.

Further dateTypes may be specified. Please refer to DataCite Metadata Schema v3.1 for details.

2.12.2 Example

An example:

```
1 <dates>
2   <date dateType="Issued">2005-04-05</date>
3 </dates>
```

An embargo example:

```
1 <dates>
2   <date dateType="Accepted">2011-12-01</date>
3   <date dateType="Available">2012-12-01</date>
4 </dates>
```

2.13 9. Language (R)

The primary language of the resource (occurrences: 0-1).

Allowed values, examples, other constraints

Allowed values are taken from IETF BCP 47, ISO 639-1 language codes. Examples: en, de, fr.

2.13.1 Example

```
1 <language>eng</language>
```

2.14 10. ResourceType (R)

A description of the resource (occurrences: 0-1).

Allowed values, examples, other constraints

The format is open, but the preferred format is a single term of some detail so that a pair can be formed with the sub-property.

Text formats can be free-text OR terms from the [CASRAI Publications resource type list](#).

Examples:

Dataset/Census Data, where Dataset is resourceTypeGeneral value and Census Data is Resource-Type value. Text/Conference Abstract, where Text is ResourceTypeGeneral value and Conference Abstract is resourceType value aligned with CASRAI Publications term.

2.14.1 10.1 resourceTypeGeneral (R)

The general type of a resource (occurrences: 1).

Allowed values, examples, other constraints

If ResourceType is used, resourceTypeGeneral is mandatory.

Controlled List Values:

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

Combine “Text” with free-text or terms from the [CASRAI Publications resource type list](#) found [here](#).

2.14.2 Example

```
<resourceType resourceTypeGeneral="Image">Animation</resourceType>
```

2.15 11. AlternateIdentifier (O)

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) (occurrences: 0-n).

Allowed values, examples, other constraints

2.15.1 11.1 alternateIdentifierType (O)

The type of the AlternateIdentifier (occurrences: 1).

Allowed values, examples, other constraints

Free text.

If AlternateIdentifier is used, alternateIdentifierType is mandatory. For the above example, the alternateIdentifierType would be A local accession number.

2.15.2 Example

```

1 <alternateIdentifiers>
2   <alternateIdentifier alternateIdentifierType="ISBN">
3     937-0-1234-56789-X
4   </alternateIdentifier>
5 </alternateIdentifiers>

```

2.16 12. RelatedIdentifier (MA)

Identifiers of related resources. These must be globally unique identifiers (occurrences: 0-n).

Allowed values, examples, other constraints

Free text.

Use this property to indicate subsets of properties, as appropriate.

Note: *Mandatory when applicable* property in OpenAIRE instead of *recommended* in DataCite.

Please refer to *Related publications and datasets information* for specific details on how to link datasets and publications.

2.16.1 12.1 relatedIdentifierType (M)

The type of the RelatedIdentifier (occurrences: 1).

Allowed values, examples, other constraints

If 12. *RelatedIdentifier (MA)* is used, 12.1 *relatedIdentifierType (M)* is mandatory.

Controlled List Values:

- ARK
- arXiv
- bibcode
- DOI
- EAN13
- EISSN
- Handle
- ISBN

- ISSN
- ISTC
- LISSN
- LSID
- PMID
- PURL
- UPC
- URL
- URN

2.16.2 12.2 relationType (M)

Description of the relationship of the resource being registered (A) and the related resource (B) (occurrences: 1).

Allowed values, examples, other constraints

If *12. RelatedIdentifier (MA)* is used, *12.2 relationType (M)* is mandatory.

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)
- *HasMetadata* (indicates resource A has additional metadata B)
- *IsMetadataFor* (indicates additional metadata A for a resource 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)

Note: `Cites` and `IsCitedBy` is specifically for when a publication/dataset directly cites another publication/dataset in its references, whereas `References` and `IsReferencedBy` is for when a dataset/publication is used as a source of information without a direct citation.

OpenAIRE encourages including minimum one of above listed relation types, but allows usage of all DataCite's relation types.

2.16.3 12.3 relatedMetadataScheme (O)

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

Allowed values, examples, other constraints

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

2.16.4 12.1 schemeURI (O)

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

Allowed values, examples, other constraints

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

2.16.5 12.1 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

2.16.6 Example

An example:

```

1 <relatedIdentifier
2   relatedIdentifierType="DOI"
3   relationType="IsCitedBy">
4   10.1234/bar
5 </relatedIdentifier>

```

2.17 13. Size (O)

Unstructured size information about the resource (occurrences: 0-n).

Allowed values, examples, other constraints

Free text.

Examples: “15 pages”, “6 MB”

2.17.1 Example

```
1 <sizes>
2   <size>15 pages</size>
3   <size>6 MB</size>
4 </sizes>
```

2.18 14. Format (O)

Technical format of the resource. (occurrences: 0-n).

Allowed values, examples, other constraints

Free text.

Use file extension or MIME type where possible, e.g., PDF, XML, MPG or application/pdf, text/xml, video/mpeg.

2.18.1 Example

```
1 <formats>
2   <format>PDF</format>
3   <format>application/pdf</format>
4 </formats>
```

2.19 15. Version (O)

The version number of the resource (occurrences: 0-1).

Allowed values, examples, other constraints

Suggested practice: track major_version.minor_version.

Register a new identifier for a major version change. Individual stewards need to determine which are major vs. minor versions (based on the work of the Earth Science Information Partners (ESIP). For more guidance, see [here](#).

May be used in conjunction with properties *11. AlternateIdentifier (O)* and *12. RelatedIdentifier (MA)* to indicate various information updates.

May be used in conjunction with property *17. Description (MA)* to indicate the nature and file/record range of version.

2.19.1 Example

```
<version>1.0</version>
```

2.20 16. Rights (MA)

Any rights information for this resource (occurrences: 0-n).

Allowed values, examples, other constraints

Free text.

Provide a rights management statement for the resource or reference a service providing such information. Include embargo information if applicable.

Use the complete title of a license and include version information if applicable.

Example: Creative Commons Attribution 3.0 Germany License

Note: *Mandatory when applicable* property in OpenAIRE instead of optional in DataCite.

OpenAIRE uses this property to explicit declare the access right of the resource via *16.1 rightsURI (MA)*. This does not exclude also using *16. Rights (MA)* property for additional rights statements as defined by DataCite Metadata Schema v3.1. In particular OpenAIRE also recommends including license information if available.

Please refer to the section *Access rights and license information* for a full example of how to declare access right and license information.

2.20.1 16.1 rightsURI (MA)

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

Allowed values, examples, other constraints

Example:

<http://creativecommons.org/licenses/by/3.0/de/deed.en>

Note: *Mandatory when applicable* property in OpenAIRE instead of optional in DataCite.

Use terms from the *info:eu-repo-Access-Terms* vocabulary. The values are:

- `info:eu-repo/semantics/closedAccess`
- `info:eu-repo/semantics/embargoedAccess`
- `info:eu-repo/semantics/restrictedAccess`
- `info:eu-repo/semantics/openAccess`

This property may also be used to explicitly declare the license for the resource.

Please refer to the section *Access rights and license information* for a full example of how to declare access right and license information.

2.20.2 Example

An example:

```
1 <rightsList>
2   <rights rightsURI="info:eu-repo/semantics/openAccess" />
3 </rightsList>
```

OpenAIRE further recommends including license information if available:

```
1 <rightsList>
2   <rights rightsURI="info:eu-repo/semantics/openAccess" />
3   <rights rightsURI="http://creativecommons.org/licenses/by/4.0/">
4     Creative Commons Attribution 4.0 International
5   </rights>
6 </rightsList>
```

2.21 17. Description (MA)

All additional information that does not fit in any of the other categories. May be used for technical information (occurrences: 0-n).

Allowed values, examples, other constraints

The format is open.

It is a best practice to supply a description.

Note: *Mandatory when applicable* property in OpenAIRE instead of optional in DataCite.

2.21.1 17.1 descriptionType (MA)

The type of the Description. (occurrences: 1).

Allowed values, examples, other constraints

If Description is used, descriptionType is mandatory.

Controlled List Values:

- Abstract
- Methods
- SeriesInformation
- TableOfContents
- Other

Note: Use of "Abstract" is *mandatory when applicable* in OpenAIRE instead of *recommended* in DataCite.

2.21.2 Example

```

1 <descriptions>
2   <description descriptionType="Abstract">
3     This is an abstract
4   </description>
5   <description descriptionType="Other">
6     This is e.g. a note.
7   </description>
8 </descriptions>

```

2.22 18. GeoLocation (O)

Spatial region or named place where the data was gathered or about which the data is focused (occurrences: 0-n).

Allowed values, examples, other constraints

Repeat this property to indicate several different locations.

2.22.1 18.1 geoLocationPoint (O)

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

Allowed values, examples, other constraints

A point contains a single latitude-longitude pair, separated by whitespace.

See *Detailed usage instructions*.

2.22.2 18.2 geoLocationBox (O)

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

Allowed values, examples, other constraints

A box contains two white space separated latitude-longitude pairs, with each pair separated by whitespace. The first pair is the lower corner (normally south west), the second is the upper corner (normally north east).

See *Detailed usage instructions*.

2.22.3 18.3 geoLocationPlace (O)

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

Allowed values, examples, other constraints

Free text. Use to describe a geographic location.

2.22.4 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).

2.22.5 Example

```
1 <geoLocations>
2   <geoLocation>
3     <geoLocationPoint>31.233 -67.302</geoLocationPoint>
4     <geoLocationBox>41.090 -71.032 42.893 -68.211</geoLocationBox>
5     <geoLocationPlace>Atlantic Ocean</geoLocationPlace>
6   </geoLocation>
7 </geoLocations>
```

OpenAIRE Guidelines for CRIS Managers based on CERIF-XML

The OpenAIRE Guidelines for CRIS Managers version 1.0, together with the accompanying material, is available at <http://doi.org/10.5281/zenodo.17065> .

The guidelines specifically provide guidance on how to specify:

- Access right
- Funding information
- Related publication and datasets.

Participate

You are invited to participate by commenting or editing the content. See our guide for how to get started:

4.1 Contributing

The OpenAIRE guidelines can be viewed online on <http://guidelines.openaire.eu>.

4.1.1 Create a GitHub account

You will need a GitHub account to contribute to the OpenAIRE guidelines. It is free and easy - just go to <https://github.com/join> and follow the instructions.

Once you've got your GitHub account, head on over to the OpenAIRE Guidelines on GitHub <https://github.com/openaire/guidelines>

GitHub Explore Features Enterprise Blog Sign up Sign in

Join GitHub

The best way to design, build, and ship software.

Step 1: Set up a personal account | **Step 2:** Choose your plan | **Step 3:** Go to your dashboard

Create your personal account

Username

This will be your username — you can enter your organization's username next.

Email Address

You will occasionally receive account related emails. We promise not to share your email with anyone.

Password

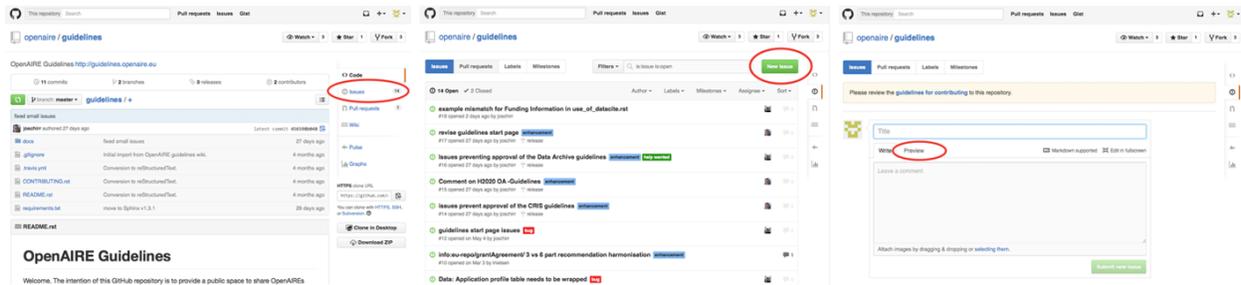
You'll love GitHub

- Unlimited collaborators
- Unlimited public repositories
- Great communication
- Friction-less development
- Open source community

4.1.2 Participate in the discussion

- Want to participate in our discussion? Join the ongoing conversation on our list of issues: <https://github.com/openaire/guidelines/issues>
- Have a question, feedback or problem? Simply go open a new issue on our GitHub repository: <https://github.com/openaire/guidelines/issues/new>.

You can also get to the “New issue” from the frontpage of our GitHub repository:



Styles

The description of issues can be styled using a syntax called Markdown. See more on how to create headings, bold and bullets on <https://guides.github.com/features/mastering-markdown/>. You can even include images in the text, simply by dragging and dropping the image onto the issue description text area.

4.1.3 Editing the guidelines

Found a typo or want to revise major parts of the guidelines? Editing the guidelines can be done directly on GitHub. If this is the first time you are using GitHub, follow our guide here: [Editing the guidelines](#)

Editing the guidelines

Found a typo or want to revise major parts of the guidelines? Editing the guidelines can be done directly on GitHub.

1. Go to <http://guidelines.openaire.eu> and find the page you want to edit.
2. Click “Edit on GitHub” in the top right corner. This will take you directly to the source file on GitHub.

OpenAIRE Guidelines

Search docs

OpenAIRE Guidelines for Literature Repositories

OpenAIRE Guidelines for Data Archives

- Introduction
- Use of OAI-PMH
- Use of DataCite
- Application Profile Overview**
- 1. Identifier (M)
- 2. Creator (M)
- 3. Title (M)
- 4. Publisher (M)
- 5. PublicationYear (M)
- 6. Subject (R)
- 7. Contributor (MA/O)
- 8. Date (M)
- 9. Language (R)
- 10. ResourceType (R)
- 11. AlternateIdentifier (O)
- 12. RelatedIdentifier (MA)

Read the Docs v: latest

Docs » OpenAIRE Guidelines for Data Archives » Application Profile Overview [Edit on GitHub](#)

Application Profile Overview

OpenAIRE builds on the DataCite Metadata Schema v3.1 by making some of the otherwise optional DataCite properties mandatory, as well as enforcing specific encoding schemes on the values of some DataCite properties. The table below details the differences from the DataCite Metadata Schema.

Within OpenAIRE the use of fields are either:

- **Mandatory (M)** = the field must always be present in the metadata record. An empty element is not allowed.
- **Mandatory when applicable (MA)** = when the value of the field can be obtained it must be present in the metadata record.
- **Recommended (R)** = the use of the field is recommended.
- **Optional (O)** = the property may be used to provide complementary information about the resource

The recommended status is made primarily to encourage users to input certain properties when creating a metadata record to enhance services.

Important

The table below details only differences from the DataCite Metadata Schema v3.1. Please refer to the DataCite Metadata Schema v3.1 for definition of properties and their encoding schemes: <http://schema.datacite.org/meta/kernel-3/index.html>

3. Click the edit icon. If it's grayed out, it's likely because you are not signed into your GitHub account.

This repository Search Pull requests Issues Gist

openaire / guidelines Watch 3 Star 1 Fork 3

branch: master guidelines / docs / data / application_profile.rst

Inielson on Mar 3 Conversion to reStructuredText. 1 contributor

22 lines (13 sloc) | 0.959 kB Raw Blame History

Application Profile Overview

OpenAIRE builds on the DataCite Metadata Schema [\[datacitever\]](#) by making some of the otherwise optional DataCite properties mandatory, as well as enforcing specific encoding schemes on the values of some DataCite properties. The table below details the differences from the DataCite Metadata Schema.

The recommended status is made primarily to encourage users to input certain properties when creating a metadata record to enhance services.

Important

The table below details only differences from the DataCite Metadata Schema [\[datacitever\]](#). Please refer to the DataCite Metadata Schema [\[datacitelink\]](#) for definition of properties and their encoding schemes: [\[datacitelink\]](#)

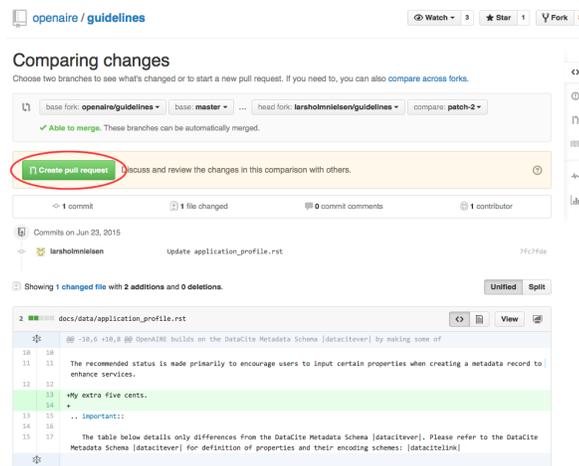
4. Edit the file. Note that the editor supports full-screen mode and allows you to preview your changes.

The screenshot shows the GitHub web editor interface for the file 'application_profile.rst' in the 'guidelines / docs / data' directory. At the top, there are navigation links and repository statistics (Watch: 3, Star: 1, Fork: 3). A blue notification box states: 'You're editing a file in a project you don't have write access to. Submitting a change to this file will write it to a new branch in your fork larsholmielsen/guidelines, so you can send a pull request.' Below this, the breadcrumb path 'guidelines / docs / data / application_profile.rst' is shown. The editor toolbar includes 'Edit file', 'Preview changes' (circled in red), 'Spaces' (set to 3), 'Soft wrap', and a 'Full-screen mode' icon (circled in red). The code editor shows the file content with line numbers 1-23. A red arrow points from the 'Preview changes' button to the 'Preview changes' text in the code editor. Another red arrow points from the 'Full-screen mode' icon to the text 'Full-screen mode'. A third red arrow points from the 'Propose file changes' button to the 'Propose file change' form below the editor. The 'Propose file change' form has a text input with 'Update application_profile.rst', a larger text area for an optional description, and two buttons: 'Propose file change' (green) and 'Cancel' (grey).

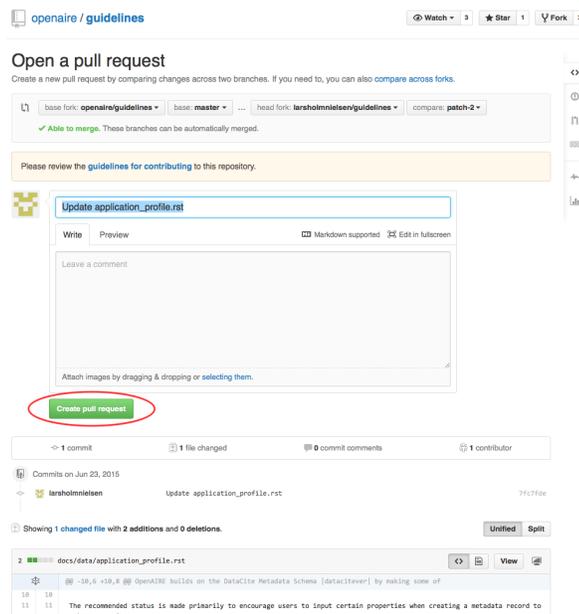
5. Once your done editing the file, write a small commit message to propose the file change and finally click “Propose file change”. This will allow you to create what is called a “pull request”. See next section.

Submitting a pull request

1. After clicking “Propose file change”, you can review your changes compared against the latest master version. If you’re happy with the changes, go a head and click “Create pull request”.

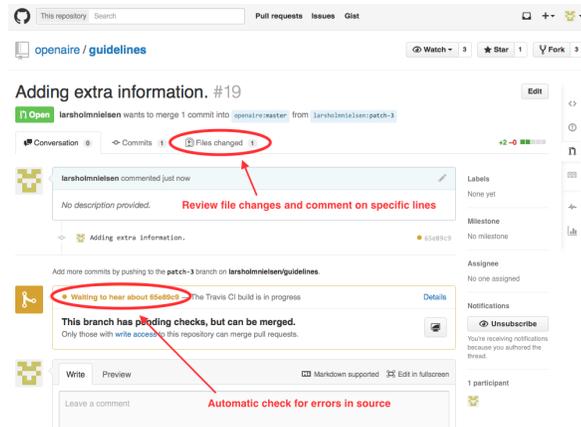


2. Enter a title for your pull request and describe the changes made if necessary and finally click, “Create pull request”.



Discussing the pull request

1. Everyone can see your proposed changes, and the changes can be discussed. Every pull request is also automatically checked for errors in the documentation source files by an external service called TravisCI.



2. The tab “Files changed”, others can see the changes to the files, and comment on individual lines as well.

Merging the pull request

Once the changes have been approved, one of the editors will merge the changes into the master version. After the merge, the changes are automatically propagated to online version on <http://guidelines.openaire.eu>.

4.1.4 Get notified

Afraid you missed an update? GitHub allows you to easily keep updated with the latest changes using their notifications:

- <https://github.com/openaire/guidelines/notifications>

OpenAIRE Validator

The OpenAIRE Validator service <http://validator.openaire.eu> allows to test your repository's compatibility with the OpenAIRE Guidelines.

If validation succeeds the data source can be registered for regular aggregation and indexing in OpenAIRE. OpenAIRE allows for registration of institutional and thematic repositories registered in OpenDOAR, research data repositories registered in re3data, individual e-Journals, CRIS, aggregators and publishers.

Horizon 2020 Open Access Requirements

The European Commission has published [Guidelines on Open Access to Scientific Publications and Research Data](#) (version 1.0, 11-Dec-2013). By following the *OpenAIRE Guidelines for Literature Repositories* it is ensured that specific requirements on bibliographic information about Open Access publications are met. These requirements are summarized here:

6.1 How the Horizon 2020 Open Access requirements are met

The European Commission has published [Guidelines on Open Access to Scientific Publications and Research Data](#) (version 1.0, 11-Dec-2013).

In addition to basic bibliographic information about deposited publications extra information is expected in the following fields:

6.1.1 Property ‘EU funding acknowledgement’

DC Field

`dc:contributor`

Usage Instruction

Values for this field must include one of the following applicable terms:

- ["European Union (EU)" and "Horizon 2020"]
- ["Euratom" and "Euratom research and training programme 2014–2018"]

6.1.2 Property ‘Peer Reviewed’

DC Field

`dc:type`

Usage Instruction

`info:eu-repo/semantics/publishedVersion`

6.1.3 Property ‘Embargo Period’

If the research outcome is published under an embargo please specify the embargo end date and the embargoed access mode.

DC Field

dc:date

Usage Instruction

info:eu-repo/date/embargoEnd/<YYYY-MM-DD>

DC Field

dc:rights

Usage Instruction

info:eu-repo/semantics/embargoedAccess

6.1.4 Property ‘Project Information’

DC Field

dc:relation

Usage Instruction

info:eu-repo/grantAgreement/Funder/FundingProgram/ProjectID/[Jurisdiction]/[ProjectName]/[ID]

Example

Project information for the OpenAIRE2020 project would be encoded as:

<dc:relation>info:eu-repo/grantAgreement/EC/H2020/643410/EU/OpenAIRE2020/OpenAIRE2020/

6.1.5 Property ‘Publication Date’

DC Field

dc:date

Usage Instruction

6.1.6 Property ‘Persistent Identifier for the Research Outcome’

DC Field

dc:identifier

dc:relation

Usage Instruction

6.1.7 Property ‘License’

DC Field

dc:rights

Usage Instruction

It is recommended to provide a URL which provides the location where the license can be read, e.g.:

<http://opendatacommons.org/licenses/by/1.0>

6.1.8 Property ‘Persistent Identifiers for Authors / Contributors’

DC Field

dc:creator

dc:contributor

Usage Instruction

Lastname, Firstname; id_orcid 0000-0000-0000-0000

6.1.9 Property ‘Reference to related research outcome’

DC Field

dc:relation

Usage Instruction

info:eu-repo/semantics/dataset/<scheme>/<id>