



OpenAIRE Guidelines for Literature Repository Managers 3.0

December 2012

Contents

[Introduction](#)

[Aim](#)

[What's new](#)

[Acknowledgements & Contributors](#)

[Editors](#)

[Experts & Reviewers](#)

[Version](#)

[Use of OAI-PMH](#)

[Metadata Format](#)

[OpenAIRE OAI Set](#)

[Set content](#)

[Compatibility of aggregators](#)

[OpenAIRE-specific use of oai_dc](#)

[All elements: Short Description](#)

Introduction

Aim

The OpenAIRE Guidelines for Literature Repository Managers 3.0 provide orientation for repository managers to define and implement their local data management policies to meet the Open Access demands of the European Commission. Furthermore, they will be compatible with the technical requirements of the OpenAIRE infrastructure that is being established to support and monitor the implementation of the FP7 OA pilot.¹

By implementing these Guidelines, repository managers enable authors who deposit publications in their repository to fulfill the EC Open Access requirements.

For developers of repository platforms, the Guidelines provide guidance to add supportive functionalities for authors of EC funded research in future versions.

What's new

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

- 1 The OpenAIRE OAI set has been renamed from `ec_fundedresources` to `openaire`.
- 2 New elements for indicating external identifiers, relations to other publications (references), and relations to research datasets have been defined.

¹ For information on the FP7 OA pilot and OpenAIRE see <http://www.openaire.eu>.

Acknowledgements & Contributors

Editors

- Marnix van Berchum (SURF, the Netherlands)
- Eloy Rodrigues (Universidade do Minho, Portugal)

Experts & Reviewers

- Mikael Karstensen Elbæk (Technical University of Denmark, Denmark)
- Samuele Kaplun (CERN, Switzerland)
- Mathias Loesch (University of Bielefeld, Germany)
- Lars Holm Nielsen (CERN, Switzerland)
- Paolo Manghi (CNR, Italy)
- Natalia Manola (University of Athens, Greece)
- Jochen Schirrwagen (University of Bielefeld, Germany)
- Birgit Schmidt (University of Goettingen, Germany)
- Friedrich Summann (University of Bielefeld, Germany)
- Pedro Príncipe (Universidade do Minho, Portugal)
- Najko Jahn (University of Bielefeld, Germany)
- Najla Rettberg (University of Goettingen, Germany)



This work is licensed under the Creative Commons Attribution-ShareAlike 3.0 Unported License.

Version

1.0	July 2010	Initial document
1.1	November 2010	Correction of names and references; addition of license and version statement
2.0	October 2012	Compatibility for aggregators; extended Namespace for Project Identification
3.0	December 2012	<ul style="list-style-type: none"> the OpenAIRE OAI set has been renamed from <code>ec_fundedresources</code> to <code>openaire</code> new relation elements for indicating external identifiers, references and connections to datasets

Use of OAI-PMH

OpenAIRE uses the OAI-PMH v2.0 protocol² for harvesting publication metadata.

Metadata Format

OpenAIRE expects metadata to be encoded in the *Dublin Core*³ metadata format (metadataPrefix `oai_dc`). For information on how to use the individual DC fields, please refer to the [table at the end](#) of this document.

OpenAIRE OAI Set

For harvesting the records relevant to OpenAIRE, the use of a specific OAI set⁴ at the local repository is mandatory. The set must have the following characteristics:

setName	setSpec*
OpenAIRE	openaire

**A harvester only uses the setSpec request to perform selective harvesting. The letters must be in small caps.*

Set content

The specific content of the `openaire` set is to be determined at the local repository. Publications to be inserted in the OpenAIRE set must conform to **one** of the following criteria:

- They are available in Open Access (full text with no access restrictions)
- They are outcome of a funded research project identified by a projectID (see below).

² <http://www.openarchives.org/OAI/openarchivesprotocol.html>

³ <http://dublincore.org/documents/dces/>

⁴ <http://www.openarchives.org/OAI/openarchivesprotocol.html#Set>

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 guidelines⁵, the provenance information has to be provided in the `about` element of an OAI record, as displayed in the following example.

```
<about>
  <provenance xmlns="http://www.openarchives.org/OAI/2.0/provenance"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/provenance
      http://www.openarchives.org/OAI/2.0/provenance.xsd">
    <originDescription altered="true" harvestDate="2012-09-17T14:58:36Z">
      <baseURL>http://dspace.library.uu.nl:8080/dspace-oai/request</baseURL>
      <identifier>oai:dspace.library.uu.nl:1874/218065</identifier>
      <timestamp>2012-01-19T12:38:56Z</timestamp>
      <metadataNamespace>http://www.loc.gov/mods/v3</metadataNamespace>
    </originDescription>
  </provenance>
</about>
```

Example of provenance information.

⁵ <http://www.openarchives.org/OAI/2.0/guidelines-provenance.htm>

OpenAIRE-specific use of oai_dc

At the record level, OpenAIRE-specific information is encoded by using a special syntax in the DC fields `relation` and `rights` to express the following information:

- **projectID (dc:relation, mandatory)** is needed to connect project information to the publication in the OpenAIRE information space. In the case of EC-funded projects, this is the Grant Agreement number as found in all documentation/correspondence between the EC and the researcher/coordinator.⁶ Further project information (project period, funding area etc.) will be ingested into the OpenAIRE information space by other means.
- **accessRights (dc:rights, mandatory)** will define the type of access to the publication. When the value of accessRights is “embargoedAccess” embargoEndDate will define the date of the embargo period.
- **externalIdentifier (dc:relation, optional)** encodes other identifiers for this publication (e.g., DOI of the publisher version)
- **reference (dc:relation, optional)** encodes links to publications referenced by this publication
- **dataset (dc:relation, optional)** encodes links to research datasets connected to this publication

Element name	projectID
DCMI definition	dc:relation
Usage	Mandatory
Usage instruction	<p>A vocabulary of projects will be exposed by OpenAIRE through OAI-PMH, and available for all repository managers. Values will include the project name and projectID.</p> <p>The projectID equals the Grant Agreement number, and is defined by the namespace⁷</p> <pre>info:eu-repo/grantAgreement/Funder/FundingProgram/ProjectNumber/[Jurisdiction]/[ProjectName]/[ProjectAcronym]/</pre>

⁶ The (generic) namespace can be used to express project information from other funders than the EC.

⁷ See <http://wiki.surffoundation.nl/display/standards/info-eu-repo/#info-eu-repo-GrantAgreementIdentifiers>

	<p>For OpenAIRE compatibility, the first three elements are mandatory and the last three are optional. Repositories may choose to use only the first part of the namespace (as in previous versions of the Guidelines), or use the extended version with six parts, which is recommended.</p> <p>When omitting fields in the extended version, the number of fields must nevertheless be preserved by using “/”. A correct example for omitting the the <code>ProjectName</code> field would therefore look like this: <code>EC/FP7/12345/EU//OpenAIREplus</code></p>
Do not confuse with	--
Examples	<pre><dc:relation> info:eu-repo/grantAgreement/EC/FP7/12345 </dc:relation> Or <dc:relation> info:eu- repo/grantAgreement/EC/FP7/12345/EU//OpenAIRE </dc:relation></pre>

Element name	accessRights
DCMI definition	dc:rights
Usage	Mandatory
Usage instruction	<p>Use values from vocabulary Access Rights at http://wiki.surffoundation.nl/display/standards/info-eu-repo/#info-eu-repo-AccessRights; values are:</p> <pre>info:eu-repo/semantics/closedAccess info:eu-repo/semantics/embargoedAccess info:eu-repo/semantics/restrictedAccess info:eu-repo/semantics/openAccess</pre>
Do not confuse with	--
Examples	<pre><dc:rights> info:eu-repo/semantics/openAccess</pre>

	</dc:rights>
--	--------------

Element name	embargoEndDate
DCMI definition	dc:date
Usage	Recommended
Usage instruction	<p>Recommended when <code>accessRights = info:eu-repo/semantics/embargoedAccess</code></p> <p>The date type is controlled by the name space <code>info:eu-repo/date/embargoEnd/</code>, see http://wiki.surffoundation.nl/display/standards/info-eu-repo/#info-eu-repo-DateTypesandvalue. Encoding of this date should be in the form YYYY-MM-DD (conform ISO 8601).</p>
Do not confuse with	--
Examples	<pre><dc:date> info:eu-repo/date/embargoEnd/2011-05-12 <dc:date></pre>

Element name	externalIdentifier
DCMI definition	dc:relation
Usage	Optional
Usage instruction	<p>List alternative identifiers for this publication (e.g., DOI of publisher's version, PubMed/arXiv ID). The syntax is: <code>openaire/externalIdentifier/scheme/identifier</code> where <code>scheme</code> must be one of the following:</p> <ul style="list-style-type: none"> • <code>doi</code> – Digital Object Identifier • <code>hdl</code> – Handle • <code>ark</code> – Archival Resource Key • <code>purl</code> – Persistent Uniform Resource Locator • <code>issn</code> – International Standard Serial Number

	<ul style="list-style-type: none"> • isbn – International Standard Book Number • wos – Web of Science accession number • pmid – PubMed ID • arxiv – arXiv.org identifier • uri – Uniform Resource Identifier
Do not confuse with	--
Examples	<pre><dc:relation> openaire/externalIdentifier/doi/10.1234/foo </dc:relation></pre>

Element name	reference
DCMI definition	dc:relation
Usage	Optional
Usage instruction	<p>Encode links to publications referenced by this publication. The syntax is: openaire/reference/scheme/identifier where scheme must be one of the following:</p> <ul style="list-style-type: none"> • doi – Digital Object Identifier • hdl – Handle • ark – Archival Resource Key • purl – Persistent Uniform Resource Locator • uri – Uniform Resource Identifier
Do not confuse with	--
Examples	<pre><dc:relation> openaire/reference/doi/10.1234/foo </dc:relation></pre>

Element name	dataset
DCMI definition	dc:relation
Usage	Optional
Usage instruction	Encodes links to research datasets connected with this

	<p>publication. The syntax is: <code>openaire/dataset/scheme/identifier</code> where scheme must be one of the following:</p> <ul style="list-style-type: none"> • <code>doi</code> – Digital Object Identifier • <code>hdl</code> – Handle • <code>ark</code> – Archival Resource Key • <code>purl</code> – Persistent Uniform Resource Locator • <code>uri</code> – Uniform Resource Identifier
Do not confuse with	--
Examples	<pre><dc:relation> openaire/dataset/doi/10.1234/foo </dc:relation></pre>

All elements: Short Description

Legend

- **mandatory (M)** = the element must always be present in the metadata record. An empty element is not allowed.
- **mandatory when applicable (MA)** = when the element can be obtained it must be present in the metadata record
- **recommended (R)** = the use of the element is recommended
- **optional (O)** = it is not important whether the element is used or not

Basic element	Status	Encoding schemes
Title	M	None, free text
Creator	M	None. Recommended practice is “last name, first name(s)”
Subject	MA	Choice of keywords and classifications can be free text (preferably in English) and defined by an URI scheme (preferably <code>info:eu-repo/classification</code>).
Description	MA	None, free text. Recommended practice is to include an abstract in English. “Abstract” is the default interpretation to the value for <code>dc:description</code>

Publisher	R	None
Contributor	O	None. Recommended practice is “last name, first name(s)”
Date	M	Date ISO 8601 W3C-DTF - “Published” is the default interpretation to the value for dc:date; When accessRights = embargoedAccess, the addition of the embargo end date is recommended, defined by the URI scheme <code>info:eu-repo/date/embargoEnd</code> ⁸
Type	M	Publication type and Version type defined by an URI scheme (preferably <code>info:eu-repo/semantics</code>).
Format	R	IANA registered list of Internet Media Types (MIME types) ⁹
Identifier	M	URI scheme, linking to persistent identifier (URN, handle, DOI), full text document or human start page. Identifiers should be unique. ¹⁰
Source	O	Guidelines for Encoding Bibliographic Citation Information in Dublin Core Metadata ¹¹ as in <code>dcterms:bibliographicCitation</code>
Language	R	ISO 639-3
Relation	M	Project identification defined by namespace <code>info:eu-repo/grantAgreement/Funder/FundingProgram/ProjectNumber/[Jurisdiction]/[ProjectName]/[Project Acronym]</code> / ¹² Square brackets indicate optional fields.
Relation	O	External identifiers <code>openaire/externalIdentifier/scheme/XX</code>

⁸ <http://wiki.surffoundation.nl/display/standards/info-eu-repo/#info-eu-repo-DateTypesandvalue>

⁹ <http://www.iana.org/assignments/media-types/>

¹⁰ See <http://www.openarchives.org/OAI/openarchivesprotocol.html#UniqueIdentifier>

¹¹ <http://dublincore.org/documents/dc-citation-guidelines/>

¹² <http://wiki.surf.nl/display/standards/info-eu-repo/#info-eu-repo-GrantAgreementIdentifiers>

		X Publication References openaire/reference/scheme/XXX Dataset References openaire/dataset/scheme/XXX
Coverage	O	“Period” is the default interpretation to the value for dc:coverage. Encoding: DCMI Period ¹³
Rights	M	accessRights mandatory, from vocabulary “AccessRights” ¹⁴
Audience	O	None. “Education level” is the default value for dc:audience.

¹³ <http://dublincore.org/documents/dcmi-period/>

¹⁴ <http://wiki.surffoundation.nl/display/standards/info-eu-repo/#info-eu-repo-AccessRights>