

OpenAIRE Integration



Available in DSpace-CRIS 5.10.x since 16th November 2019 see <https://github.com/4Science/DSpace/commit/622d79d627f6321e735412209908d40602c6d0a3>

DSpace-CRIS provides a strong out-of-box integration with OpenAIRE.

There are three are of integration:

- OpenAIRE Guidelines compliance for Literature Repository, DataArchive and CRIS System
- OpenAIRE Project API
- OpenAIRE's Usage Statistic service

OpenAIRE Guidelines



OAI-PMH peculiarity in DSpace-CRIS

Please be aware that DSpace-CRIS introduces some extra feature to the OAI-PMH Server as described in the latest paragraph of this section that are used to implement the OpenAIRE guidelines

Literature Repository

DSpace-CRIS has the same level of compatibility than the basic DSpace with the Literature Repository Guidelines v3

DataArchive

DSpace-CRIS ships with a custom OAI context implementing the [OpenAIRE guidelines for DataArchive](#) already updated to implement the **DataCite metadata version 4**

The default context is available at **`$(dspace.url)/oai/openairedata`**

xoai.xml

```
<Context baseurl="openairedata" name="OpenAIRE Data Repository Context">
  <!-- OpenAIRE filter -->
  <Filter ref="openaireDataFilter"/>
  <!-- Just an alias, in fact it returns all items within the driver context -->
  <Set ref="openaireDataSet"/>
  <!-- Metadata Formats -->
  <Format ref="oaidc"/>
  <!-- set the reference at the new format called datacite41_openaire -->
  <Format ref="datacite41_openaire"/>
  <Description>
    This contexts complies with Data Archive OpenAIRE rules.
  </Description>
</Context>
```



Please note that such endpoint supports the exposure of the ORCID identifiers for the authors!

CRIS System

DSpace-CRIS ships with a custom OAI context implementing the [OpenAIRE guidelines for CRIS Manager v1.1.1](#)

The default context is available at **`$(dspace.url)/oai/openairecris`**

xoai.xml

```
<Context baseurl="openairecris" name="OpenAIRE Context for CRIS System">
  <!-- Date format, field prefixes, etc are ensured by the transformer -->
  <Transformer ref="openairecrisTransformer"/>
  <Filter ref="openAireCRISFilter"/>
  <!-- Just an alias, in fact it returns all items within the driver context -->
  <Set ref="openairePublicationsCrisSet"/>
  <Set ref="openaireProductsCrisSet"/>
  <Set ref="openairePatentsCrisSet"/>
  <Set ref="openairePersonsCrisSet"/>
  <Set ref="openaireOrgunitsCrisSet"/>
  <Set ref="openaireProjectsCrisSet"/>
  <Set ref="openaireEventsCrisSet"/>
  <Set ref="openaireEquipmentsCrisSet"/>
  <!-- Metadata Formats -->
  <Format ref="oaidc"/>
  <Format ref="oai_cerif"/>
  <Description>
    This contexts complies with OpenAIRE rules for CRIS system v 1.1.1
  </Description>
</Context>
```

The mapping is configured in the `dspace/config/crosswalks/oai/metadataFormats/oai_cerif.xsl` file



Please note that such endpoint supports the exposure of the ORCID identifiers for the authors as well the exposure of the richer set of information collected in DSpace-CRIS for each entity

example of metadata for a publication

```
<oai_cerif:Publication
  xmlns:doc="http://www.lyncode.com/xoai"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:oai_cerif="https://www.openaire.eu/cerif-profile/1.1/"
  id="123456789/7">
  <oai_cerif:Type
    xmlns:oai_cerif="https://www.openaire.eu/cerif-profile/vocab/COAR_Publication_Types"
  >http://purl.org/coar/resource_type/c_18wq</oai_cerif:Type>
  <oai_cerif:Title>The Data Model of the OpenAIRE Scientific Communication e-Infrastructure</oai_cerif:
Title>
  <oai_cerif:PublicationDate>2012-11-30</oai_cerif:PublicationDate>
  <oai_cerif:Volume>343</oai_cerif:Volume>
  <oai_cerif:Issue>168</oai_cerif:Issue>
  <oai_cerif:StartPage>168</oai_cerif:StartPage>
  <oai_cerif:EndPage>180</oai_cerif:EndPage>
  <oai_cerif:DOI>10.1007/978-3-642-35233-1_18</oai_cerif:DOI>
  <oai_cerif:Handle>123456789/7</oai_cerif:Handle>
  <oai_cerif:URN>http://hdl.handle.net/123456789/7</oai_cerif:URN>
  <oai_cerif:Authors>
    <oai_cerif:Author>
      <oai_cerif:DisplayName>Manghi, Paolo</oai_cerif:DisplayName>
      <oai_cerif:Person
        id="d1d9582d-1a8b-4e1b-882f-1b421f5a95d3">
        <oai_cerif:PersonName>
          <oai_cerif:FirstNames>Paolo</oai_cerif:FirstNames>
          <oai_cerif:FamilyNames>Manghi</oai_cerif:FamilyNames>
        </oai_cerif:PersonName>
      </oai_cerif:Person>
    </oai_cerif:Author>
    <oai_cerif:Author>
      <oai_cerif:DisplayName>Houssos, Nikos</oai_cerif:DisplayName>
      <oai_cerif:Person
        id="797a27d6-f2a1-4cfl-8b2d-c25f79de742a">
        <oai_cerif:PersonName>
          <oai_cerif:FirstNames>Nikos</oai_cerif:FirstNames>
          <oai_cerif:FamilyNames>Houssos</oai_cerif:FamilyNames>
```

```

</oai_cerif:PersonName>
<oai_cerif:ORCID>0000-0002-3748-8359</oai_cerif:ORCID>
<oai_cerif:ResearcherID>XXXXXXXXX-ID</oai_cerif:ResearcherID>
<oai_cerif:ScopusAuthorID>SCP-XXXXXX-ID</oai_cerif:ScopusAuthorID>
<oai_cerif:Affiliation startDate="2006-09-01">
  <oai_cerif:OrgUnit
    id="8a17cae0-242a-4154-8bd6-6ea16ae1fe32">
    <oai_cerif:Name>National Hellenic Research Foundation | EIE<
/oai_cerif:Name>
    <oai_cerif:PartOf>
      <oai_cerif:OrgUnit
        id="ae388eb5-21c6-44b9-b1fb-a897d2f09d8d">
        <oai_cerif:Name>National Documentation Centre |
EKT</oai_cerif:Name>
      </oai_cerif:OrgUnit>
    </oai_cerif:PartOf>
  </oai_cerif:OrgUnit>
</oai_cerif:Affiliation>
</oai_cerif:Person>
</oai_cerif:Author>
<oai_cerif:Author>
  <oai_cerif:DisplayName>Mikulicic, Marko</oai_cerif:DisplayName>
  <oai_cerif:Person
    id="ed964c75-25f4-4106-a559-ea4015601e48">
    <oai_cerif:PersonName>
      <oai_cerif:FirstNames>Marko</oai_cerif:FirstNames>
      <oai_cerif:FamilyNames>Mikulicic</oai_cerif:FamilyNames>
    </oai_cerif:PersonName>
  </oai_cerif:Person>
</oai_cerif:Author>
<oai_cerif:Author>
  <oai_cerif:DisplayName>Jörg, Brigitte</oai_cerif:DisplayName>
  <oai_cerif:Person
    id="02b50a4a-2749-40a3-b56f-3b3d02c69c16">
    <oai_cerif:PersonName>
      <oai_cerif:FirstNames>Brigitte</oai_cerif:FirstNames>
      <oai_cerif:FamilyNames>Jörg</oai_cerif:FamilyNames>
    </oai_cerif:PersonName>
  </oai_cerif:Person>
</oai_cerif:Author>
</oai_cerif:Authors>
<oai_cerif:Keyword>open access</oai_cerif:Keyword>
<oai_cerif:Keyword>infrastructure</oai_cerif:Keyword>
<oai_cerif:Keyword>data model</oai_cerif:Keyword>
<oai_cerif:Keyword>CERIF</oai_cerif:Keyword>
<oai_cerif:Keyword>DataCite</oai_cerif:Keyword>
<oai_cerif:Abstract>The OpenAIREplus project aims to further develop and operate the OpenAIRE e-
infrastructure, in order to provide a central entry point to Open Access and
non-Open Access publications and datasets funded by the
European Commission and National agencies. The infrastructure provides the services to populate, curate, and
enrich an Information Space
by collecting metadata descriptions relative to organizations,
data sources, projects, funding programmes, persons, publications, and datasets. Stakeholders in the research
process and
scientific communication, such as researchers, funding
agencies, organizations involved in projects, project coordinators, can here find the information to improve
their research and statistics
to measure the impact of Open Access and funding schemes over
research. In this paper, we introduce the functional requirements to be satisfied and describe the OpenAIREplus
data model entities
and relationships required to represent information capable of
meeting them.</oai_cerif:Abstract>
  <oai_cerif:OriginatesFrom>
    <oai_cerif:DisplayName>2nd-Generation Open Access Infrastructure for
Research in Europe</oai_cerif:DisplayName>
    <oai_cerif:Project
      id="e9ed438e-c7f7-4a18-95e5-3f635ea65fee">
      <oai_cerif:Title>2nd-Generation Open Access Infrastructure for Research in Europe<
/oai_cerif:Title>
    <oai_cerif:Funded>
      <oai_cerif:By>

```

```

        <oai_cerif:DisplayName>EC</oai_cerif:DisplayName>
        <oai_cerif:OrgUnit />
    </oai_cerif:By>
    <oai_cerif:As>
        <oai_cerif:Funding>
            <oai_cerif:Identifier
                type="https://w3id.org/cerif/vocab
/IdentifierTypes#ProjectReference">283595</oai_cerif:Identifier>
            <oai_cerif:PartOf>
                <oai_cerif:Type
                    xmlns="https://www.openaire.eu/cerif-profile
/vocab/OpenAIRE_Funding_Types">https://www.openaire.eu/cerif-profile/vocab
/OpenAIRE_Funding_Types#FundingProgramme</oai_cerif:Type>
                <oai_cerif:Name>FP7</oai_cerif:Name>
            </oai_cerif:PartOf>
        </oai_cerif:Funding>
    </oai_cerif:As>
</oai_cerif:Funded>
</oai_cerif:Project>
</oai_cerif:OriginatesFrom>
    <oai_cerif:Access>http://purl.org/coar/access_right/c_abf2</oai_cerif:Access>
</oai_cerif:Publication>

```

OAI-PMH peculiarity in DSpace-CRIS

DSpace-CRIS extends the functionalities of the OAI-PMH Server on three area:

1. the index script identifies object to include in the OAI core (or to update) querying the search SOLR core. This mean that **DSpace-CRIS must be indexed (discovery) before to run the XOAI script**. It is possible to configure a custom query to limit the amount of objects to index in addition to the standard capability to limit the OAI objects to expose over a specific OAI Context.
2. the XOAI document embed information from the 1st level linked objects, i.e. data about person that have authored a publication or organisations that have funded a project
3. nested metadata are stored in the XOAI document ensuring the possibility to read them as a whole group so to associate for instance the affiliation used in a publication with the specific author

The XOAI script has an additional option `-t` that allow to specify which type of entity is indexed: item, rp, project, ou, other, all.

The `item.compile` of the XOAI document embed the information of the linked object where available (i.e. when the authority points to a CRIS object) just after the metadata inside an `<element name="authority" />` tag keeping the right sequence

item.compile of the XOAI document

```
<metadata
xmlns="http://www.lyncode.com/xoai"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.lyncode.com/xoai
http://www.lyncode.com/xsd/xoai.xsd">
  <element name="dc">
    <element name="contributor">
      <element name="author">
        <element name="en_US">
          <field name="value">Cadili, Francesco</field>
          <field name="authority">rp00003</field>
          <field name="confidence">600</field>
          <element name="authority">
            <element name="crisrp">
              <element name="fullName">
                <element name="none">
                  <element name="none">
                    <field name="value">Cadili, Francesco</field>
                  </element>
                </element>
              </element>
            </element>
          <field name="dept">
            <element name="none">
              <element name="none">
                <field name="value">Department of ...</field>
                <field name="authority">ou00001</field>
                <field name="confidence">600</field>
              </element>
            </element>
          </element>
          ...
        </element>
        <field name="value">Unknown, Author</field>
        <field name="value">Another Internal, Author</field>
        <field name="authority">rp00005</field>
        <field name="confidence">600</field>
        <element name="authority">
          .....
        </element>
      </element>
    </element>
  </element>
</metadata>
```

OpenAIRE Project API

A choiceAuthority is provided out of box in DSpace-CRIS to query the OpenAIRE project database via API to link a publication, dataset or patent with a project.

The openAIRE Project authority is configured out-of-box for the dc.relation metadata

dspace.cfg

```
##### Authority Control Settings #####
plugin.named.org.dspace.content.authority.ChoiceAuthority = \
  org.dspace.app.cris.integration.ORCIDAuthority = RPAuthority,\
  org.dspace.content.authority.ItemAuthority = PublicationAuthority,\
  org.dspace.content.authority.ItemAuthority = DataSetAuthority,\
  org.dspace.app.cris.integration.OpenAIREProjectAuthority = ProjectAuthority,\...
# dc.relation, Publication / Product / Patent
choices.plugin.dc.relation = ProjectAuthority
choices.presentation.dc.relation = suggest
authority.controlled.dc.relation = true
```

Once the publication is approved the retrieved project is imported locally in a CRIS Project mapping the OpenAIRE information to the key attribute of the CRIS project entity.

cris-plugin.xml

```
<bean id="org.dspace.app.cris.integration.authority.AuthoritiesFillerConfig" class="org.dspace.app.cris.
integration.authority.AuthoritiesFillerConfig">
    <property name="fillers">
        <map>
            <entry key="orcid">
                <bean class="org.dspace.app.cris.integration.authority.
ORCIDImportFiller" parent="fullitemMetadataConfiguration" />
            </entry>
            <entry key="openAireProject">
                <ref bean="openAIREProjectImportFiller" />
            </entry>
            ...
        </map>
    </property>
</bean>

...

<bean id="openAIREProjectImportFiller" class="org.dspace.app.cris.integration.authority.
OpenAIREProjectImportFiller" parent="fullitemMetadataConfiguration">
    <property name="searchService" ref="org.dspace.discovery.SearchService" />
    <property name="openAIREProjectService" ref="org.dspace.authority.openaireproject.
OpenAIREProjectService" />
    <property name="customMapping">
        <bean class="java.util.HashMap" />
    </property>
</bean>
```

OpenAIRE's Usage Statistic service

The OpenAIRE plugin for DSpace has been integrated out-of-box in DSpace-CRIS and some compatibility issues have been solved so that no extra installation is required.

To send usage statistics to OpenAIRE is sufficient to configure the relevant parameter in the build.properties

```
### OpenAIRE Piwik Tracker ###
oapiwik.piwik.siteID =
oapiwik.piwik.tokenAuth =
oapiwik.piwik.enabled =
```