The DRIS+ Project – Enhancing the euroCRIS Directory of Research Information Systems

Pablo de Castro
DRIS+ coordinator
University of Strathclyde Glasgow
https://orcid.org/0000-0001-6300-1033
pablo.de-castro@strath.ac.uk
euroCRIS Secretary

Jan Dvořák
DRIS+ technical lead
Charles University Prague
https://orcid.org/0000-0001-8985-152X
Jan.Dvorak@ff.cuni.cz
euroCRIS CERIF Task Group leader
DRIS+: Enhancing the euroCRIS Directory of Research Information Systems (DRIS)

The DRIS+ proposal aims to enhance the euroCRIS Directory of Research Information Systems (DRIS) and to make it automatically searchable via a dedicated API.

This improvement is a follow-up action to the euroCRIS-led 3-month METIS2OpenAIRE project that was awarded funding by OpenAIRE in early 2018. This project allowed the first institutional CRIS (METIS at Radboud University in Nijmegen, the Netherlands) to expose and test its metadata feed against the CERIF-XML Guidelines and to undergo the first test validation against a minimally sufficient validator developed by euroCRIS.

As the number of test-harvested CRIS system increases and the opportunities grow for expanding this OpenAIRE data provider role across vendors and solutions, the enhanced DRIS is seen as a key element to streamline the process for metadata harvesting from the OpenAIRE portal.

**Tender priority topics addressed:** The DRIS+ proposal addresses Challenge no 3 within the OpenAIRE Call for Innovators, i.e. “Expanding the OpenAIRE Service Portfolio”. The proposed work will provide a missing piece in the set of features required to make the CRIS validation and harvesting into the OpenAIRE content aggregation a reality.

https://www.openaire.eu/open-call-winner-phase-1-eurocris
Using DRIS API: Use Case no 1 – OpenAIRE

Identification of candidate data provider and its URL for metadata exposure (OAI-PMH) → API → Metadata harvesting and aggregation from data provider

OpenDOAR (literature repositories) → API → re3data (data repositories) → euroCRIS DRIS (CRIS systems)

Validate your datasource

Run compatibility test against the OpenAIRE literature guidelines.
Run compatibility test against the OpenAIRE Guidelines for Data Archives.
Run compatibility test against the OpenAIRE Guidelines for CRIS Managers based on CERIF-XML.
euroCRIS Directory of Research Information Systems (DRIS)

- 894 entries as of today (up from 800 on Dec 23rd, 2020)
- Sitting in the same Dspace-CRIS platform that hosts the euroCRIS repository
- Platform enhancement through geolocation features (Phase 1) and badges (Phase 3)
- Work-in-progress: the DRIS keeps growing

https://dspacecris.eurocris.org/cris/explore/dris
euroCRIS DRIS: geolocation feature
euroCRIS DRIS keeps growing (thanks to RIM community input)

- The DRIS keeps steadily growing as stated in the DRIS+ project goals
- DRIS dissemination at euroCRIS membership meetings has switched to online webinars
- Data collection from institutions and vendors: ongoing effort
- End-goal is to promote system interoperability via the implementation of the OpenAIRE CERIF-XML Guidelines for CRIS Managers

http://hdl.handle.net/11366/1252
“Minimally sufficient metadata” for DRIS entries

The Directory of Research Information Systems (DRIS) is an euroCRIS initiative to map the available research information management infrastructure in Europe (and beyond). As of mid-May 2021, it includes close to 900 CRIS systems for institutions, research funders and other bodies in multiple European countries and beyond.

In order to make the addition of new entries as simple as possible, a minimally sufficient metadata set has been adopted for their description in the DRIS. This also aims to keep the maintenance effort as low as possible by avoiding fields prone to frequent changes such as contact person or number of records kept in the CRIS. This basic description contains just six fields:

- **Type**: type of CRIS, i.e. institutional, funder, national, regional, aggregation, subject-specific, other
- **Organisation**: body operating the CRIS (usually a university or research centre, or a funder)
- **Name**: name of the CRIS (where available)
- **Platform**: software solution underpinning the CRIS. It may be a commercial platform (Pure, Converis, Symplectic Elements, SoleCRIS, etc.), an open source solution (DSpaceCRIS, Haplo) or it may have been developed in-house
- **Status**: status of the CRIS, usually operational or under construction
- **URL**: URL for the open CRIS portal where available. If the CRIS is closed, any webpage where information on the system is provided, see for instance the entry for METIS at Radboud Uni

https://eurocris.org/services/dris
Online DRIS submission form

DRIS: Directory of Research Information Systems

Type
CRIS type. Possible values: institutional, funder, regional, national, international, aggregation, other

Organisation
Organisation served by the CRIS or running the CRIS

Name
Name of the CRIS. If there’s also an acronym, it can be added in brackets

Platform
CRIS solution. Possible values are in-house-built: Pure, Convera, Symplectic Elements, DSpace-CRIS, Haslo, Worktribe, etc. Details may be added to complete the description for an in-house-built platform, e.g. “In-house-built (Java/Web services)”

Status
- Operational
- Under construction

URL
URL for an open CRIS portal, e.g. https://www.researchportal.be/en. If the CRIS has no open portal, then a URL may be used that provides info on the system itself such as https://researchsupport.admin.ox.ac.uk/reporting/symplectic

CAPTCHA
This question is for testing whether or not you are a human visitor and to prevent automated spam submissions.

Submit

https://eurocris.org/dris/dris-form
Other interesting DRIS stats

<table>
<thead>
<tr>
<th>Country</th>
<th>Discover</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>301</td>
<td>IRINS</td>
</tr>
<tr>
<td>Norway</td>
<td>132</td>
<td>Pure</td>
</tr>
<tr>
<td>Italy</td>
<td>76</td>
<td>CRISlin</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>71</td>
<td>In-house-built</td>
</tr>
<tr>
<td>United States</td>
<td>42</td>
<td>IRIS</td>
</tr>
<tr>
<td>Germany</td>
<td>38</td>
<td>Omega-PSIR</td>
</tr>
<tr>
<td>Poland</td>
<td>33</td>
<td>DSpace-CRIS</td>
</tr>
<tr>
<td>Spain</td>
<td>23</td>
<td>Symplectic Elements</td>
</tr>
<tr>
<td>Netherlands</td>
<td>18</td>
<td>Converis</td>
</tr>
<tr>
<td>Finland</td>
<td>16</td>
<td>VIVO</td>
</tr>
</tbody>
</table>

next >
DRIS+ project goals (live demo to follow)

1. to add geolocation features to the DRIS entries while keeping the directory growing

2. to expand the underlying DRIS data model in order to add fields like the OAI-PMH endpoint URL*

3. to implement badges for highlighting OpenAIRE-compliant CRIS systems (also searchable via the API) and CRIS systems hosting research datasets

4. to add a basic authentication layer for users of the DRIS API (eg OpenAIRE) to go through as a safety mechanism

5. to test the retrieval of OAI-PMH endpoint URLs by the real-life OpenAIRE validation workflow for a first case study, https://dspacecris.eurocris.org/cris/dris/dris01840 (Polish Platform for Medical Research, PPM) – ongoing work, currently setting up the OAI-PMH endpoint URL

* URL for the endpoint through which a CRIS system will expose its research info metadata
Adding OAI-PMH endpoint URLs to DRIS records
Displaying OAI-PMH endpoint URL & OpenAIRE badge

Internal view: OAI-PMH endpoint URLs are only visible in DRIS records when specific Dspace-CRIS users are logged in.
Displaying OAI-PMH endpoint URL & OpenAIRE badge

External view: OAI-PMH endpoint URLs are not visible in DRIS records if user is not logged in

https://dspacecris.eurocris.org/cris/dris/dris01170
‘Marking’ research datasets on DRIS records
‘Marking’ research datasets on DRIS records: RDM+ badge

Aim: to record CRIS systems holding both research datasets and metadata for research datasets hosted elsewhere
Using DRIS API: Use Case no 2 – CRIS as EOSC providers

• The EOSC Association relies on data (and literature) repositories as default data providers for datasets

• However, CRIS systems often contain datasets too (or at least *metadata* on datasets)

• CRIS systems are rarely registered in data repository directories like re3data – but they are listed in the DRIS

• If we managed to highlight those DRIS entries that contain either datasets or their associated metadata, that would be a very valuable info on additional data providers for the EOSC

• Now where do we get the info on *which CRIS systems are hosting data*?
Retrieving info on dataset hosting by a CRIS: PURE@Strathclyde

- Where the information is available on the CRIS portal, it can directly be taken from there
- Where it’s not available (or where there’s no portal) it will need to be sourced from the CRIS managers: pending dissemination work
Basic authentication layer on the DRIS API

Aim: to ‘protect’ CRIS systems from openly sharing potentially sensitive information – OAI-PMH endpoint URLs can only be retrieved via the DRIS API by authenticated users (such as OpenAIRE)
Using DRIS API: Use Case no 3 – National RIM Offices

National Research Information Management Office

Directory of Research Information Systems (DRIS) (euroCRIS)

Search the DRIS

for driscountry:(Germany)

Results 1-10 of 34 (Search time: 0.0 seconds).
Results/Page 10 | Sorted by Relevance
Test Drive the API

vade mecum
The basic example

List all the CRISs in the DRIS: https://api.eurocris.org/dris/entries

- The total number given
- First 100 records listed
- Stand-alone records, context @included

- Iteration through the Link HTTP header (rel=next)
Examples of filtering

• List all operational CRISs in the DRIS:  
  https://api.eurocris.org/dris/entries?status.id=classcerif00935 or  
  https://api.eurocris.org/dris/entries?status.label=operational

• List all operational CRISs in Germany, Austria and Switzerland:  
  https://api.eurocris.org/dris/entries?status.label=operational&country.code.alpha2=DE,AT,CH
Further examples of filtering

• List all operational CRISs that cover datasets:
  https://api.eurocris.org/dris/entries?status.label=operational&coverage.label=Dataset

• List all CRISs with the RDM+ badge:
  https://api.eurocris.org/dris/entries?rdm=true

• List all CRISs with the OpenAIRE badge:
  https://api.eurocris.org/dris/entries?openaire=true

• List all operational CRIS that are built on the DSpace-CRIS platform:
  https://api.eurocris.org/dris/entries?status.label=operational&cris-platform.label=DSpace-CRIS
A toolset for testing the DRIS API

https://github.com/jdvorak001/DRIS-API-testing

Installation and running

1. Checkout this project
2. Run the Python environment (pkexec visw) and do the rest of this procedure from within there:
3. Run ./download.py to retrieve the whole contents of the API into a raw data directory.
4. Run ./process.py to process the retrieved contents from the raw data directory into individual, pretty-printed entries in a data directory.
5. Run ./download.py to re-retrieve the individual records from the data directory (overwriting the original files). If your DRIS_CREDENTIALS environment variables contain a username:password pair, it is used for basic HTTP authentication with the API endpoint.
6. Run ./list-openapiCred.jsonURLs.sh to summarize the different values of the openapiCredEndpointURL field in JSON files in the data directory.

After having run step 5 with authentication, please ensure the data with the openapiCredEndpointURL fields filled in do not accidentally get committed in the git repo.
Thanks!

Pablo de Castro
DRIS+ coordinator
University of Strathclyde Glasgow
https://orcid.org/0000-0001-6300-1033
pablo.de-castro@strath.ac.uk
euroCRIS Secretary

Jan Dvořák
DRIS+ technical lead
Charles University Prague
https://orcid.org/0000-0001-8985-152X
Jan.Dvorak@ff.cuni.cz
euroCRIS CERIF Task Group leader