Discovery Metadata

[From a UK RepositoryNet+ Project Perspective]
RepNet Worklines for Repository Enhancement

- RepNet Service Catalogue
- RepNet Service Desk
- IR Managers
- HEI Finance Systems
- HEI CRIS Systems

- Supporting activities
- Consider for inclusion in RepNet service catalogue

- KPL Report on Search
- IRUS-UK
- Project Reporting
- Benchmark Publisher Website Delivery

- OARR

- ORI
  - OpenDOAR
  - ROAR

- R-J Broker
- RoMEO
- JULIET

- NAMES

- Metadata Quality

- CERIF/OAI-PMH
- RIOXX
- ISNI/ORCID
- CRIS/Sword EndPoints
- Keepln (Eprints)
- PDF->XML
- DOI retrieval – Cross-Ref
- Fund-Ref

- Gap analysis
The Complex CRIS/IR Landscape in the UK

RepNet map for CRIS/IR implementation at UK HEIs
This is a map of the available Research Information Management (RIM) infrastructure across UK HEIs. Five use cases for CRIS/IR implementation are featured on the map as identified by RepNet in collaboration with external stakeholders including IR managers and Research Office managers. Colour codes are as follows:

- IR-only: green
- IR+CRIS: blue
- CRIS-only: yellow
- IR+Symplectic: red
- IR+RMS: violet
- No data: white

If your institution is white-marked, please drop us a line with your specific RIM system configuration to support@repositorynet.ac.uk or fill in the RepNet survey at http://ukcor.org/events/future-events/teesside-university-friday-9th-november-2012/repnet_survey/. Thanks!

Unlisted · 2 Collaborators · 977 views
Created on Feb 5 · By pcastrom1 · Updated Apr 4
Rate this map · Write a comment · KML · 🌍

UK RepositoryNet+
The Complex CRIS/IR Landscape in the UK: Use Cases

• IR-only (plus IR-as-CRIS)

• IR + RMS

• IR + Symplectic

• IR + bespoke CRIS

• IR + CERIF-based (commercial) CRIS
  - IR linked to CRIS (item exchange)
  - IR detached from CRIS (no exchange)

• CRIS-only (plus CRIS-as-IR)

Source: ‘RepNet use cases for CRIS/IR implementation in UK HEIs’
The UKRepNet Metadata Enhancement Strand

• The “Minimum Sufficient Metadata” Concept

• Specific Requirements for Institutional Repositories:
  - Visibility (Open Access, OAI-PMH)
  - Simplicity (aiming to have direct deposit from researchers)
  - Needs to Cover Preservation Issues

• Significant Changes in the UK Open Access Landscape
  - Finch Report
  - RCUK Open Access Policy

• New requirements for 'enhanced metadata':
  - Funder Information
  - Author Identifiers

• RIOXX

RIOXX: Developing Repository Metadata Guidelines

Requirements from research funders, users and those seeking to exploit repository content in various ways have led to a conclusion that there would be tangible benefits if UK repositories were to expose a more consistent set of metadata. Building on the RIO Extension Report, there were clear use cases from stakeholders to develop an application profile for interoperability; to allow consistency and quality in the sharing of metadata between institutions, funders, and publishers.
Adoption of CERIF in Higher Education Institutions in the UK: A Landscape Study

1 Executive summary

However despite the widespread use of CERIF as an underlying standard, many institutions are not engaging with CERIF directly. Staff find CERIF complex and rely on external expertise from CRIS vendors and UK user groups; only institutions involved in JISC projects are properly engaging with CERIF. However many staff would like to engage more, citing lack of time as the main factor. There is increasing recognition that local understanding of RIM processes and how local data maps to CERIF allows more efficient and cost-effective CRIS implementation and ongoing development.
Simple Dublin Core
The Simple Dublin Core Metadata Element Set (DCMES) consists of 15 metadata elements:
1. Title
2. Creator
3. Subject
4. Description
5. Publisher
6. Contributor
7. Date
8. Type
9. Format
10. Identifier
11. Source
12. Language
13. Relation
14. Coverage
15. Rights

Qualified DublinCore

Full metadata record

DC Field
dc.contributor.author Adams, Sam
dc.contributor.author de Castro, Pablo
dc.contributor.author Echenique, Pablo
dc.contributor.author Estrada, Jorge
dc.contributor.author Hanwell, Marcus D
dc.contributor.author Murray-Rust, Peter
dc.contributor.author Sherwood, Paul
dc.contributor.author Thomas, Jens
dc.contributor.author Townsend, Joe A
dc.date.accessioned 2011-11-01T12:13:00Z
dc.date.available 2011-11-01T12:13:00Z
dc.date.issued 2011-10-14
dc.identifier http://dx.doi.org/10.1186/1758-2946-3-38
dc.identifier.citation Journal of Cheminformatics 3(1) : 38 (2011)
dc.identifier.uri http://hdl.handle.net/10261/41947
dc.description.abstract Abstract Computational Quantum Chemistry has been used to study small to medium sized molecules. Many thou...
Deep DC-based metadata customisation

Specific Requirements for Content Aggregation: Guidelines for Metadata Harmonisation

DRIVER Guidelines 2.0

Guidelines for content providers - Exposing textual resources with OAI-PMH
Guidelines as an Answer to Content Aggregation Needs

OpenAIRE Guidelines 2.0
Guidelines for content providers of the OpenAIRE information space
October 2012

OpenAIRE compliant repositories

Sort by: Name  Filter:  
Show all - Hide all

AMS Acta
Show details  No statistics available

Academic Bibliography and Institutional Archive of Ghent University
Show details  No statistics available

Access to Research and Communications Annals
Show details  No statistics available

Antropológicas
Show details  No statistics available

ArchiMer - Institutional Archive of Ifremer
Show details  No statistics available

Archivio Aperto di Ateneo
Show details  No statistics available

Archivio Giuliano Marini
Show details  No statistics available

euroCRIS and OpenAIRE work together to enable data exchange with CRIS
20 February 2013
Pending Errors -- CSIC Institutional Repository
[example, dated 2009]

e-ciencia Aggregation: Potential Issues
- Protected PDF: 404 records
- No Text Extraction (no OCR?): 16 records

Mandatory DRIVER Guidelines:
- Missing Agreed Arc Type: 7779 records
- Non-Textual Document Type: 8 records
- Full-Text: No Link from Splash-page: 73 records
- Full-Text: Error on Full-Text Download: 49 records
- Undefined Author: 34 records
- Undefined Title: 2 records

Recommended DRIVER Guidelines:
- Undefined Subject: 1427 records
- Missing Description: 41 records
- Wrong Language Format (Arc): 147 records
- Missing Agreed Document Type: 3548 records
- Missing Format: 3 records
- Missing Rights: 296 records
- Wrong Author Name Format (Last Name, First Name): 20 records

Note. - A UKRepNet-driven Aggregation is being developed for the UK Repository Network that could deliver similar metadata validation functionality
<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>dc.contributor.author</td>
<td>Rajon, Etienne</td>
</tr>
<tr>
<td>dc.contributor.author</td>
<td>Plotkin, Joshua B.</td>
</tr>
<tr>
<td>dc.date.accessioned</td>
<td>2013-09-06T18:28:00Z</td>
</tr>
<tr>
<td>dc.date.available</td>
<td>2013-09-06T18:28:00Z</td>
</tr>
<tr>
<td>dc.date.issued</td>
<td>2013-08-28</td>
</tr>
<tr>
<td>dc.identifier</td>
<td>doi:10.5061/dryad.1f217</td>
</tr>
<tr>
<td></td>
<td>underlying quantitative traits. Proceedings of the Royal Society B</td>
</tr>
<tr>
<td></td>
<td>280(1769): 20131552</td>
</tr>
<tr>
<td>dc.identifier.uri</td>
<td><a href="http://hdl.handle.net/10255/dryad.52207">http://hdl.handle.net/10255/dryad.52207</a></td>
</tr>
<tr>
<td>dc.description</td>
<td>In the classic view introduced by R. A. Fisher, a quantitative trait</td>
</tr>
<tr>
<td></td>
<td>is encoded by many loci with small, additive effects. Recent advances</td>
</tr>
</tbody>
</table>
METS, MODS and PREMIS! Oh My!

(and a little MIX and other schema too)

Integrating Digital Library Standards for **Interoperability and Preservation**

Thomas Habing, thabing@uiuc.edu
Grainger Engineering Library Information Center
University of Illinois at Urbana-Champaign

http://www.loc.gov/standards/mods/presentations/habing-ala07/index.html
New Metadata Requirements: Funder Info

What’s new
In comparison with previous versions of the Guidelines, this 2.0 version introduces two main changes:

1. Support for *aggregators to become OpenAIRE compatible* in order to expose their metadata to the OpenAIRE infrastructure.

2. **Extended namespace for project identification**. The extended namespace is intended to support a generic way of expressing project information, allowing its use not only for EC/FP projects, but ideally for any funder and project (national or international). Although the use of the extended namespace is recommended, repositories and journals that are already OpenAIRE compatible will remain compatible with no additional work, by

rioxx
the RIOXX metadata profile and guidelines

Approaches to handling funders and projectIDs in a RIOXX record

The most important requirement for RIOXX is to provide a way to collect basic information about how research outputs (especially papers) have been funded. That is to say, we need to collect two types of information for a given paper:

- the funder
- the project or grant ID under which this research output was funded.
# Preliminary Approaches to Funder Info Coding: Enlighten

**Enlighten: Research and APC funding workflows at the University of Glasgow**


<table>
<thead>
<tr>
<th>Item Type:</th>
<th>Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status:</td>
<td>Published</td>
</tr>
<tr>
<td>Refereed:</td>
<td>Yes</td>
</tr>
<tr>
<td>Glasgow Author(s):</td>
<td>Padgett, Prof Miles and Dienerowitz, Dr Maria</td>
</tr>
<tr>
<td>Authors:</td>
<td>Dienerowitz, M., Gibson, G., Dienerowitz, F., and Padgett, M.</td>
</tr>
<tr>
<td>College/School:</td>
<td>College of Science and Engineering &gt; School of Physics and Astronomy</td>
</tr>
<tr>
<td>Journal Name:</td>
<td>Journal of Optics</td>
</tr>
<tr>
<td>ISSN:</td>
<td>2040-8978</td>
</tr>
<tr>
<td>ISSN (Online):</td>
<td>2040-8986</td>
</tr>
<tr>
<td>Copyright Holders:</td>
<td>Copyright © 2012 European Optical Society</td>
</tr>
<tr>
<td>First Published:</td>
<td>First published in Journal of Optics 14(4):045003</td>
</tr>
<tr>
<td>Publisher Policy:</td>
<td>Reproduced in accordance with the copyright policy of the publisher</td>
</tr>
</tbody>
</table>

University Staff: Request a correction | Enlighten Editors: Update this record

## Funder and Project Information

<table>
<thead>
<tr>
<th>Project Award Code</th>
<th>Project Name</th>
<th>Principal Investigator</th>
<th>Funder's Name</th>
<th>Funder Ref</th>
<th>Lead Dept</th>
</tr>
</thead>
<tbody>
<tr>
<td>48257</td>
<td>Multi-object, high-throughput, spectro-microscopy: Life Sciences Interface</td>
<td>Miles Padgett</td>
<td>Engineering &amp; Physical Sciences Research Council (EPSRC)</td>
<td>EP/H007636/1</td>
<td>Physics and Astronomy</td>
</tr>
</tbody>
</table>
New fields have been added to Enlighten which comply with RCUK’s requirements for providing funder and data set references. These have been added in the document upload section of the Enlighten deposit workflow and will assist us in reporting the number of papers which are green or gold open access. (See Figure 7.)

<table>
<thead>
<tr>
<th>License:</th>
<th>UNSPECIFIED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embargo expiry date:</td>
<td>Year: [Blank] Month: Unspecified Day: ?</td>
</tr>
<tr>
<td>Open Access Option:</td>
<td>UNSPECIFIED Green Gold Other</td>
</tr>
<tr>
<td>Funder Acknowledgement:</td>
<td>UNSPECIFIED Yes No</td>
</tr>
<tr>
<td>Research Materials Acknowledgement:</td>
<td>UNSPECIFIED Yes No</td>
</tr>
</tbody>
</table>
Public Library of Science’s Cameron Neylon on the state of Open Access: Where are we, what still needs to be done?

“The single most important task today is putting in place robust and transparent mechanisms to report on policy compliance, pricing, and monitor the growth of access.

“This may seem rather prosaic but we have wildly different estimates of the proportion and quantity of OA. Much of the fragmentation in today’s debate is caused by people building arguments on contradictory data. And it has been too easy for institutions and funders to announce mandates without systems to monitor their success, let alone enforce them.”

DATA-DRIVEN INFORMATION WORKFLOW

**THE DASHBOARD** (OA flavour-agnostic)

<table>
<thead>
<tr>
<th>Research article 1</th>
<th>Accepted</th>
<th>Published</th>
<th>Embargoed</th>
<th>Available OA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research article 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research article 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research article 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research article 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ITEM EMBARGOED
IR MANAGER DEALS W/ AUTHOR(S) FOR FULL-TEXT APPENDING

FULL-TEXT PAPER AVAILABLE OA
Thanks!

Pablo de Castro
UK RepositoryNet+ Project
EDINA National Data Centre
v1pdeca@staffmail.ed.ac.uk