Progress in the implementation of OpenAIRE CERIF-XML Guidelines

Andreas Czerniak
https://orcid.org/0000-0003-3883-4169
Bielefeld University Library, Bielefeld, DE
Agenda

Evolution of OpenAIRE guidelines

Interoperability

Making Research Information visible to the European Commission and European Open Science Cloud
Evolution of OpenAIRE Guidelines
Guidelines for Science Content Providers

RESEARCH LITERATURE
- Thematic Repositories
- Institutional Repositories
- E-journals

RESEARCH SOFTWARE
- Software Repositories
- Catch-all Repositories

RESEARCH DATA
- Data Repositories
- Catch-all Repositories
- Institutional & thematic repositories

OTHER RESEARCH PRODUCTS
- Catch-all Repositories

RESEARCH INFORMATION
- Project, Funder, Organisation Units,... (CRIS)

https://guidelines.openaire.eu
Landscape of Repositories types

- Institutional/publication repositories
- Journals/publishers
- Data repositories
- Software repositories
- Other Products repositories
- CRIS
Repository types and research output

- **Literature**
  - Article
  - Preprint
  - Report
  - ...

- **Datasets**
  - Dataset
  - Collection
  - Clinical Trials
  - ...

- **Software**
  - Research Software
  - ...

- **Other Research Products**
  - Service
  - Workflow
  - Interactive Resource
  - ...

- **Projects**
  - Funder
  - Organization Units
  - Equip.
  - ...

Diagram:
- Institutional/publication repositories
- Journals/publishers
- Data repositories
- Software repositories
- Other Products repositories
- CRIS

euroCRIS series of webinars | 23.-27. November 2020
Research information is heavily interlinked, contextually rich.
The OpenAIRE Guidelines for CRIS Managers

- Researchers
- Institutions
- Funders
- Research facility operators
- Persons
- Organisations
  - incl. their organisational structure

Research actors

Scope

Research process
- Projects
- Funding
- Equipment
- Service
- Communication
- Events

Research results
- Publications
- Products
- Patents
- Research datasets
- Research software

Resource Types vocabulary

CRIS

Protocol: OAI-PMH 2.0
Metadata format: CERIF XML

OpenAIRE
Evolution of the OpenAIRE CRIS Guidelines - and further perspective

- CRIS-CERIF v1.0 (2015)
- updating Guidelines for CRIS (v1.1.2) (2020/21)
- registration of CRISs (2021)

Research Output in perspective of European Open Science Cloud

OpenAIRE Guidelines Objectives

- to improve interoperability of metadata information exchange
- to support FAIR Data Principles
- by defining application profiles to describe different kinds of research products
- by (re-)using existing standards (CERIF, Dublin Core, DataCite)
- by re-using vocabularies and extending them when necessary
Interoperability with OpenAIRE

euroCRIS series of webinars | 23.-27. November 2020
Three CRIS platforms with built-in support

**DSpace-CRIS**
- OpenSource CRIS implementation
- based on DSpace
- extended entities
- since v5.10, v6.3

**Pure**
- a successful commercial CRIS product
- since v4.14

**Omega-PSIR**
- Polish institutions

---

**Warsaw University of Technology**

- in-house built institutional CRISs
- national aggregators
  - NARCIS from DANS/NL
  - VIRTA from CSC/FI
- others to follow
Validation @ PROVIDE dashboard

Validate your datasource

- Literature repository
- Data repository
- CRIS systems
Registration @ PROVIDE dashboard

OpenAIRE PROVIDE

Register
Validator
Validate
Validation History
Notifications

REPOSITORIES

Test

REPOSITORIES

Literature repository
Data repository
Journal

Aggregator
CRIS systems
DRIS: The Directory of Research Information Systems

750+ systems registered
Collecting metadata, links, and full-texts from more than 12K sources worldwide to materialize a graph where entities of the research life cycle are linked to each other.
OpenAIRE Research Graph
the supply chain

Data sources

Aggregation → Deduplication → Enrichment

Different records representing the same entity (results or organization) are merged in one.

Metadata records corresponding to equivalent objects are merged. Pre-print, post-print, published versions are considered equivalent for stats & monitoring purposes.

Full-text mining
Inference of new properties and links

OpenAIRE series of webinars | 23.-27. November 2020
An open metadata research graph of interlinked scientific products, with access rights information, linked to funding information and research communities.
OpenAIRE Research Graph

the supply chain

Data sources

Aggregation

Deduplication

Enrichment

Full-text mining
Inference of new properties and links

Different records representing the same entity (results or organization) are merged in one

Metadata records corresponding to equivalent objects are merged. Pre-print, post-print, published versions are considered equivalent for stats & monitoring purposes

Stats analysis

indexing

OpenAIRE | MONITOR
OpenAIRE | EXPLORE
OpenAIRE | CONNECT
OpenAIRE | DEVELOP

EC Participant Portal

euroCRIS series of webinars | 23.-27. November 2020
community participation

❖ for the Current Research Information Systems [CRIS] Guidelines
https://docs.google.com/document/d/1BkkO3ysr5Pmsqt7kZ118ZxBv1wE9Md3TYvp9vtUUV-c/edit ; https://github.com/openaire/guidelines-cris-managers

❖ Overview of repository platform applications & versions compliant with which version of Guidelines:
https://docs.google.com/spreadsheets/d/1_TQzbMf1jz-YEgatt76IME-MBbt28g-1_N7j_GZyoPc/edit#gid=0

❖ OpenAIRE community calls: https://www.openaire.eu/provide-community-calls
Thank you!

Andreas Czerniak (Bielefeld University Library)

andreas.czerniak@uni-bielefeld.de

Aenne Löhden, Amelie Bäcker, Jochen Schirrwagen, Alessia Bardi, Claudio Atzori, Paolo Manghi
Evolution of the OpenAIRE-Guidelines - and further perspective

- Literature Guidelines v1
- Literature Guidelines v3
- CRIS-CERIF v1.0
- CRIS-CERIF minor update v1.1.1
- new releases of inst.&thematic Guidelines v4 & Content Acquisition Policy
- updating Guidelines for CRIS-/Data-/Software-/Other Research Products
- in perspective of EOSC

2010
2012
2013
2014
2015
2018
2020

https://guidelines.openaire.eu