







THE INTERNATIONAL ORGANISATION
FOR RESEARCH INFORMATION

Compliance of the
OpenAIRE Guidelines for CRIS Managers v1.1.1
with the FAIR Principles

a report by

OpenAIRE and euroCRIS

authors

Andreas Czerniak^{1,4} , Jan Dvořák^{2,5} ,
Jochen Schirrwagen^{1,4} , and Dragan Ivanović^{3,5} 

¹ Bielefeld University Library, Germany [ROR::02hpadn98](https://ror.org/02hpadn98)

² Institute of Information Studies and Librarianship, Charles University, Czechia

[ROR::024d6js02](https://ror.org/024d6js02)

³ University of Novi Sad, Serbia [ROR::00xa57a59](https://ror.org/00xa57a59)

⁴ OpenAIRE Nexus (Grant ID: 101017452)

⁵ euroCRIS [ROR::04340dw19](https://ror.org/04340dw19)

May, 3. 2023

Abstract

Current research information systems (CRIS), also known as Research Information Management Systems (RIMS), open up the potential to provide contextualized research information and to exchange it with other services.

The Common European Research Information Format (CERIF) XML Exchange Format provides a basis for this. The OpenAIRE Guidelines for CRIS Managers, among others, are based on this format. They enable the aggregation of research information from CRIS systems into OpenAIRE and thus become part of the OpenAIRE Research Graph.

An open question is to what extent these guidelines are aligned with the requirements for discoverability, accessibility, interoperability, and reuse of research information defined in the FAIR Data Principles. The FAIR Data Maturity Model criteria are an important reference point for assessing compliance with these principles. It is novel to apply these criteria to properties and attributes of research information as well. Building on this, the report examines how the *OpenAIRE Guidelines for CRIS managers* relate to the FAIR Data Maturity Model.

Keywords: CRIS. RIM. CERIF. OpenAIRE. euroCRIS. FAIR. RDA.

DOI: 10.5281/zenodo.6627246

Contents

- Introduction 1
 - Adapting the maturity levels 2
- Result of Evaluation 3
 - Resource Type: **Publications** 3
 - Resource Type: **Products** 12
- Conclusions 18

Introduction

The *OpenAIRE Guidelines for CRIS managers* describe the CERIF-XML profile for CRIS managers to be compatible with OpenAIRE [HJD15], [DBRS18a]. The Guidelines provide orientation for CRIS managers to expose their metadata in a way that is compatible with the OpenAIRE infrastructure [Ope22]. By implementing the Guidelines, CRIS managers support the inclusion and therefore the reuse of metadata from their systems within the OpenAIRE infrastructure. For developers of CRIS platforms, the Guidelines provide advice to add supporting functionality for CRIS managers and users. The exchange of information between individual CRIS systems and the OpenAIRE infrastructure serves as an example of point-to-point data exchange between CRIS and aggregating systems. The guidelines are continuously refined and improved as part of a participatory, community-based process. In recent years, one of the goals in updating the Guidelines has been to include elements with descriptions that explicitly support the FAIR principles. For this purpose, the current version of the *OpenAIRE Guidelines for CRIS Managers* version 1.1.1[DBRS18a] was evaluated against the *RDA FAIR Data Maturity Principles: Specification and Guidelines* [Gro20]. This report is also guided by the discussions on FAIRness of research information in the FAIRIO project [HNAK21].

The following section describes the FAIRification process of the OpenAIRE Guidelines for CRIS Managers. In this evaluation process, the specification and sheet (xlsx) provided by RDA, which are based on the resource type *data*, are adapted to the specific requirements in CRIS environments. The resource types that the OpenAIRE guidelines [DBRS18b] focus on are: Publication, Product, Patent, Person, Organization Unit, Project, Funding, Service, Equipment, and Event.

The FAIR principles focus on metadata- and data-elements that are based on the resource type **data sets**. Other resources types are not considered in the FAIR principle context. Different facets are identified for adaptation:

- Identify FAIR indicators which are relevant to CRIS resources types
- Adaption and redefinition of the "Maturity level per indicator"

Adapting the maturity levels

The RDA Specification proposes the indicator maturity levels and are defined as follows:

0. not applicable
1. not being considered yet
2. under consideration or in planning phase
3. in implementation phase
4. fully implemented

The above levels do not fit into the context of a metadata schema and are adapted for the report as follows:

0. not applicable / out of scope of the Guidelines / not supported
3. enabled through the Guidelines, the concrete CRIS needs to implement
4. guaranteed through the Guidelines

The above three metric items are set to each indicator.

Result of Evaluation

This chapter describes the assessment results for each of the FAIR indicators. The FAIR Data maturity indicators are published at <https://doi.org/10.15497/rda00050>. The figures are showing the specifications from RDA FAIR data maturity model sheet and the results in detail. However, the Guidelines are fostering the findability, accessibility, interoperability, and reusability of records. The evaluation was based on the table provided by the RDA working group.

Evaluation

The following sections show the results and briefly describe the RDA FAIR Data Maturity Principle element identifier, a description, and the priority. Each FAIR principle element is expanded to include further references to the *OpenAIRE Guidelines for CRIS Managers*. These references include dedicated fields, the requirements with **M** equal Mandatory or **O** equal Optional, and additionally comments, e.g.

Field(s): *Internal Identifier*^M

Comments: *'Internal Identifier' is not guaranteed which has local scope in a repository*

Resource Type: Publications

Findable

Figure 1 is showing the results for the Findable indicator based on the above mentioned methods. As a result, the Guidelines promote and require the findability of data and metadata.

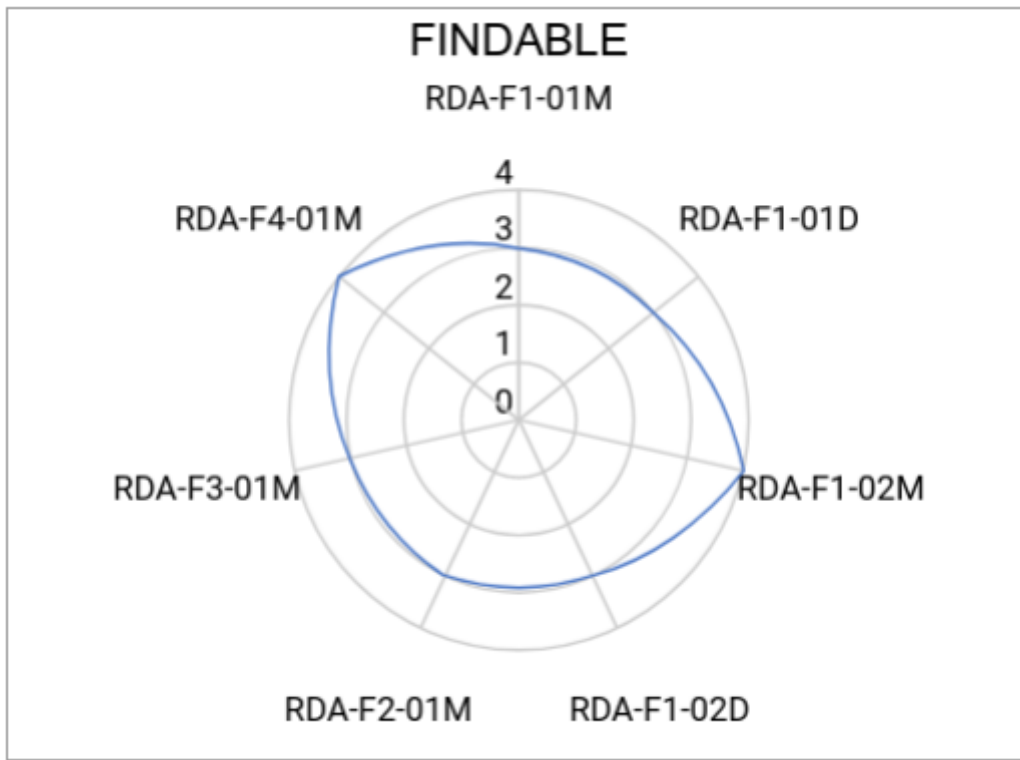


Figure 1: Resource type: **Publications** – Evaluation overview of *Findability*

In detail the results of the evaluation are:

RDA-F1-01M: *Metadata is identified by a persistent identifier* (Essential)

- Field(s): -
- Comments: *not controlled by / in the scope of the Guidelines*

RDA-F1-01D: *Data is identified by a persistent identifier, Essential*

- Field(s): *DOI^M, Handle^M, ISI-Number^M, SCP^M, PMCID^M, ISBN^M (, ISSN^M)*
- Comments: -

RDA-F1-02M: *Metadata is identified by a globally unique identifier* (Essential)

- Field(s): *Internal Identifier^M, OAI-PMH Id^M*
- Comments: *'Internal Identifier' is not guaranteed which has local scope in a repository*

RDA-F1-02D: *Data is identified by a globally unique identifier* (Essential)

- Field(s): *URI^O, URL^O*
- Comments: -

RDA-F2-01M: *Rich metadata is provided to allow discovery* (Essential)

- Field(s): *implicitly*
- Comments: *'Internal Identifier' is not guaranteed which has local scope in a repository*

RDA-F3-01M: *Metadata includes the identifier for the data* (Essential)

- Field(s): *URI^O*
- Comments: -

RDA-F4-01M: *Metadata is offered in such a way that it can be harvested and indexed* (Essential)

- Field(s): *OAI-PMH Id^M*
- Comments: -

Accessible

Figure 2 is showing the results for the Accessible indicator. Some elements especially regarding "Data" indicators are out of scope of the Guidelines.

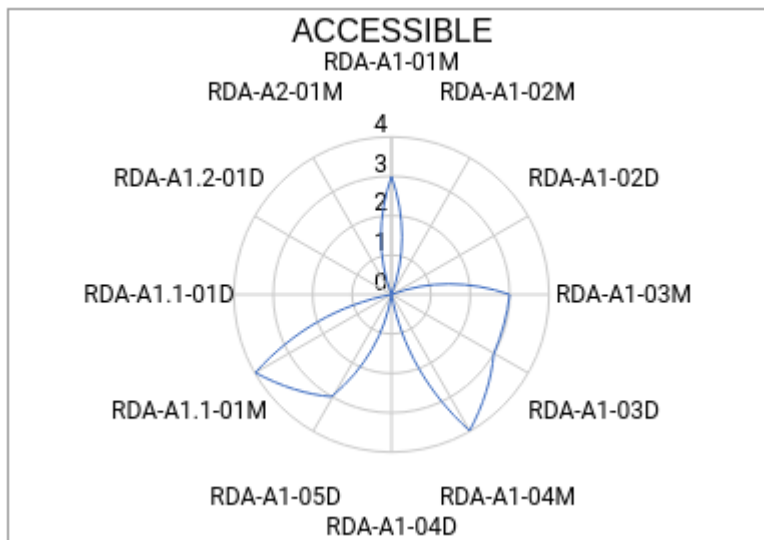


Figure 2: Resource type: **Publications** – Evaluation result of *Accessibility*

RDA-A1-01M: *Metadata contains information to enable the user to get access to the data* (Important)

- Field(s): *DOI^O*
- Comments: *principle is supported but not mandatory*

RDA-A1-02M: *Metadata can be accessed manually (i.e. with human intervention)* (Essential)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: *CRIS UI display and curate metadata records*

RDA-A1-02D: *Data can be accessed manually (i.e. with human intervention)* (Essential)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: *URI can be accessed (manually)*

RDA-A1-03M: *Metadata identifier resolves to a metadata record* (Essential)

- Field(s): *OAI-PMH protocol with OAI-PMH-Identifier^O*
- Comments: -

RDA-A1-03D: *Data identifier resolves to a digital object* (Essential)

- Field(s): *URL^M*
- Comments: -

RDA-A1-04D: *Data is accessible through standardised protocol* (Essential)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: *if references are given*

RDA-A1-05D: *Data can be accessed automatically (i.e. by a computer program)* (Important)

- Field(s): -
- Comments: *enabled by the Guidelines*

RDA-A1.1-01M: *Metadata is accessible through a free access protocol* (Essential)

- Field(s): *Open Archive Interface - Protocol for Metadata Harvesting^M*
- Comments:

-

RDA-A1.1-01D: *Data is accessible through a free access protocol* (Important)

- Field(s): -

- Comments: *not controlled by / in the scope of the Guidelines*

RDA-A1.2-01D: *Data is accessible through an access protocol that supports authentication and authorisation (Useful)*

- Field(s): -
- Comments: *not controlled by / in the scope of the Guidelines*

RDA-A2-01M: *Metadata is guaranteed to remain available after data is no longer available (Essential)*

- Field(s):
not controlled by / in the scope of the Guidelines
- Comments:
OAI-PMH excludes the metadata after the record is gone.

Interoperable

Figure 3 is showing the results for the Interoperable indicator. Some elements, especially regarding "Data" indicators, are out of scope of the Guidelines.

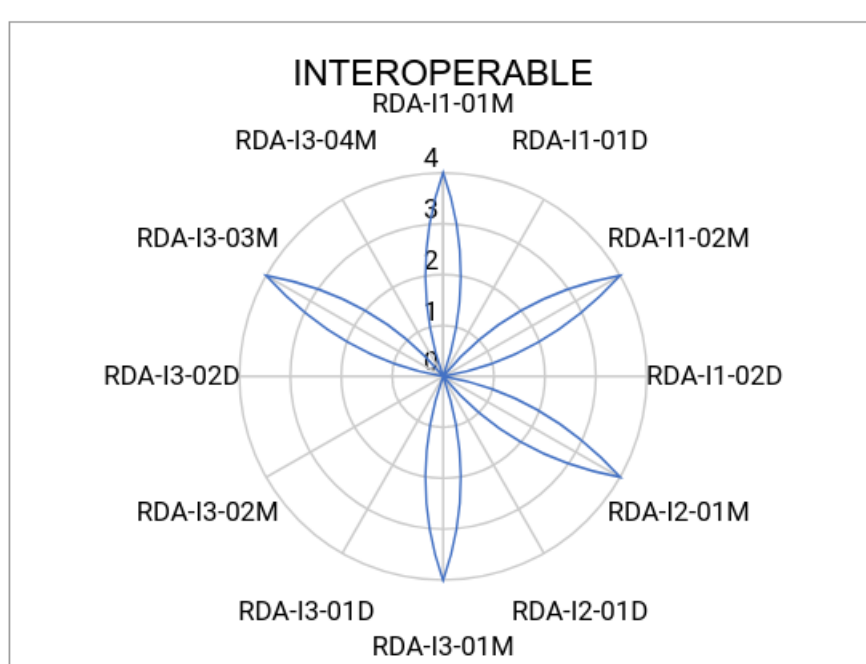


Figure 3: Resource type: **Publications** – Evaluation result of *Interoperability*

RDA-I1-01M: *Metadata uses knowledge representation expressed in standardised format* (Important)

- Field(s): *CERIF XML*^O
- Comments: *Metadata representation in CERIF XML*

RDA-I1-01D: *Data uses knowledge representation expressed in standardised format* (Important)

- Field(s): -
- Comments: *not controlled by / in the scope of the Guidelines*

RDA-I1-02M: *Metadata uses machine-understandable knowledge representation* (Important)

- Field(s): *CERIF XML*^M
- Comments: *Knowledge representation is cross-domain and on the level of CERIF-XML*

RDA-I1-02D: *Data uses machine-understandable knowledge representation* (Important)

- Field(s): -
- Comments: *not controlled by / in the scope of the Guidelines*

RDA-I2-01M: *Metadata uses FAIR-compliant vocabularies* (Important)

- Field(s): *controlled vocabularies*^M
- Comments: *supported for some vocabularies in the Guidelines, e.g. COAR vocabularies, euroCRIS*

RDA-I2-01D: *Data uses FAIR-compliant vocabularies* (Useful)

- Field(s): -
- Comments: *not controlled by / in the scope of the Guidelines*

RDA-I3-01M: *Metadata includes references to other metadata* (Important)

- Field(s): *Internal Identifiers*^M, *References*^O
- Comments: -

RDA-I3-02M: *Metadata includes references to other data* (Useful)

- Field(s): -
- Comments: *not controlled by / in the scope of the Guidelines*

RDA-I3-02D: *Data includes qualified references to other data* (Useful)

- Field(s): -
- Comments: *not controlled by / in the scope of the Guidelines*

RDA-I3-03M: *Metadata includes qualified references to other metadata* (Important)

- Field(s): *Internal Identifiers^M, References^O*
- Comments: -

RDA-I3-04M: *Metadata include qualified references to other data* (Useful)

- Field(s): -
- Comments: *not controlled by / in the scope of the Guidelines*

Reusable

Figure 4 is showing the results for the Reusable indicator.

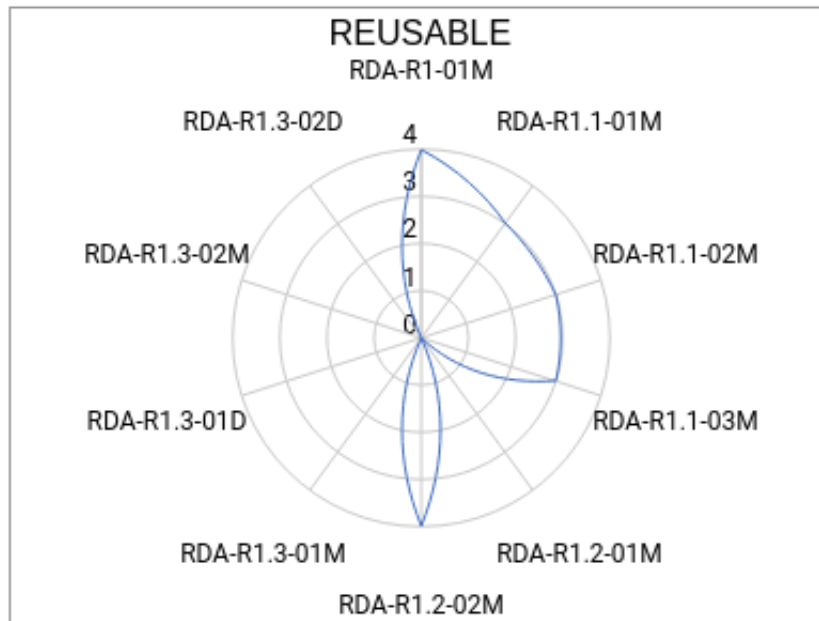


Figure 4: Resource type: **Publications** – Evaluation result of *Reusability*

RDA-R1-01M: *Plurality of accurate and relevant attributes are provided to allow reuse* (Essential)

- Field(s): -
- Comments: *Mandatory, as long as it concerns bibliographic information*

RDA-R1.1-01M: *Metadata includes information about the licence under which the data can be reused* (Essential)

- Field(s): *License*^O
- Comments: -

RDA-R1.1-02M: *Metadata refers to a standard reuse licence* (Important)

- Field(s): *License*^O
- Comments: -

RDA-R1.1-03M: *Metadata refers to a machine-understandable reuse licence* (Important)

- Field(s): *License*^O
- Comments: -

RDA-R1.2-01M: *Metadata includes provenance information according to community-specific standards* (Important)

- Field(s): *not supported*
- Comments: *Improving in the guidelines*

RDA-R1.2-02M: *Metadata includes provenance information according to a cross-community language* (Useful)

- Field(s): *Implementation*^M
- Comments: *Improving the description*

RDA-R1.3-01M: *Metadata complies with a community standard* (Essential)

- Field(s): *supported*
- Comments: *Mandatory and support thru metadata schema format: oai_cerif_openaire*

RDA-R1.3-01D: *Data complies with a community standard* (Essential)

- Field(s): -
- Comments: *not controlled by / in the scope of the Guidelines*

RDA-R1.3-02M: *Metadata is expressed in compliance with a machine-understandable community standard* (Essential)

- Field(s): *supported*
- Comments: *Mandatory, implicitly supported by using CERIF-XML-Schema for the application profile*

RDA-R1.3-02D: *Data is expressed in compliance with a machine-understandable community standard* (Important)

- Field(s): -
- Comments: *not controlled by / in the scope of the Guidelines*

Resource Type: Products

The evaluation for the resource type: Product shows similar results. For clarification, the elements of the RDA FAIR Principles and their results are presented in this section.

Findable

RDA-F1-01M: *Metadata is identified by a persistent identifier* (Essential)

- Field(s): DOI^M
- Comments: -

RDA-F1-01D: *Data is identified by a persistent identifier* (Essential)

- Field(s): DOI^M , $Handle^M$, $ISI-Number^M$, SCP^M , $PMCID^M$, $ISBN^M$ ($,ISSN^M$)
- Comments: -

RDA-F1-02M: *Metadata is identified by a globally unique identifier* (Essential)

- Field(s): *Internal Identifier*, $OAI-PMH Id^M$
- Comments: *'Internal Identifier' is not guaranteed which has local scope in a repository*

RDA-F1-02D: *Data is identified by a globally unique identifier* (Essential)

- Field(s): URI^O , URL^O
- Comments: -

RDA-F2-01M: *Rich metadata is provided to allow discovery* (Essential)

- Field(s): *implicitly*^M
- Comments: *'Internal Identifier' is not guaranteed which has local scope in a repository*

RDA-F3-01M: *Metadata includes the identifier for the data* (Essential)

- Field(s): URI^O
- Comments: -

RDA-F4-01M: *Metadata is offered in such a way that it can be harvested and indexed* (Essential)

- Field(s): $OAI-PMH Id^M$
- Comments: -

Accessible

RDA-A1-01M: *Metadata contains information to enable the user to get access to the data* (Important)

- Field(s): *DOI^O*
- Comments: *principle is supported but not mandatory*

RDA-A1-02M: *Metadata can be accessed manually (i.e. with human intervention)* (Essential)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: *CRIS UI display and curate metadata records*

RDA-A1-02D: *Data can be accessed manually (i.e. with human intervention)* (Essential)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: *URI can be accessed (manually)*

RDA-A1-03M: *Metadata identifier resolves to a metadata record* (Essential)

- Field(s): *OAI-PMH protocol with OAI-PMH-Identifier^M*
- Comments: -

RDA-A1-03D: *Data identifier resolves to a digital object* (Essential)

- Field(s): *URL^M*
- Comments: -

RDA-A1-04D: *Data is accessible through standardised protocol* (Essential)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: *if references are given*

RDA-A1-05D: *Data can be accessed automatically (i.e. by a computer program)* (Important)

- Field(s):
- Comments: *enabled by the Guidelines*

RDA-A1.1-01M: *Metadata is accessible through a free access protocol* (Essential)

- Field(s): *Open Archive Interface - Protocol for Metadata Harvesting^M*
- Comments: -

RDA-A1.1-01D: *Data is accessible through a free access protocol* (Important)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: -

RDA-A1.2-01D: *Data is accessible through an access protocol that supports authentication and authorisation* (Useful)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: -

RDA-A2-01M: *Metadata is guaranteed to remain available after data is no longer available* (Essential)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: *OAI-PMH excludes the metadata after the record is gone.*

Interoperable

RDA-I1-01M: *Metadata uses knowledge representation expressed in standardised format* (Important)

- Field(s): *CERIF XML^O*
- Comments: *Metadata representation in CERIF XML*

RDA-I1-01D: *Data uses knowledge representation expressed in standardised format* (Important)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: -

RDA-I1-02M: *Metadata uses machine-understandable knowledge representation* (Important)

- Field(s): *CERIF XML^M*
- Comments: *Knowledge representation is cross-domain and on the level of CERIF-XML*

RDA-I1-02D: *Data uses machine-understandable knowledge representation* (Important)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: -

RDA-I2-01M: *Metadata uses FAIR-compliant vocabularies* (Important)

- Field(s): *controlled vocabularies^M*

- Comments: *supported for some vocabularies in the Guidelines, e.g. COAR vocabularies, euroCRIS*

RDA-I2-01D: *Data uses FAIR-compliant vocabularies* (Useful)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: -

RDA-I3-01M: *Metadata includes references to other metadata* (Important)

- Field(s): *Internal Identifiers^M, References^O*
- Comments: -

RDA-I3-01D: *Data includes references to other data* (Useful)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: -

RDA-I3-02M: *Metadata includes references to other data* (Useful)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: -

RDA-I3-02D: *Data includes qualified references to other data* (Useful)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: -

RDA-I3-03M: *Metadata includes qualified references to other metadata* (Important)

- Field(s): *Internal Identifiers^M, GeneratedBy^M, PresentedAt^M, Coverage^M*
- Comments: -

RDA-I3-04M: *Metadata include qualified references to other data* (Useful)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: -

Reusable

RDA-R1-01M: *Plurality of accurate and relevant attributes are provided to allow reuse*
(Essential)

- Field(s): -
- Comments: *Mandatory, as long as it concerns bibliographic information*

RDA-R1.1-01M: *Metadata includes information about the licence under which the data can be reused* (Essential)

- Field(s): *License*^O
- Comments: -

RDA-R1.1-02M: *Metadata refers to a standard reuse licence* (Important)

- Field(s): *License*^O
- Comments: -

RDA-R1.1-03M: *Metadata refers to a machine-understandable reuse licence* (Important)

- Field(s): *License*^O
- Comments: -

RDA-R1.2-01M: *Metadata includes provenance information according to community-specific standards* (Important)

- Field(s): *not supported*
- Comments: *Improving in the guidelines*

RDA-R1.2-02M: *Metadata includes provenance information according to a cross-community language* (Useful)

- Field(s): *Implementation*^M
- Comments: *Improving the description*

RDA-R1.3-01M: *Metadata complies with a community standard* (Essential)

- Field(s): *supported*
- Comments: *Mandatory, support thru metadata schema format: oai_cerif_openaire*

RDA-R1.3-01D: *Data complies with a community standard* (Essential)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: -

RDA-R1.3-02M: *Metadata is expressed in compliance with a machine-understandable community standard* (Essential)

- Field(s): *supported*
- Comments:

Mandatory and implicitly supported by using CERIF-XML-Schema for the application profile

RDA-R1.3-02D: *Data is expressed in compliance with a machine-understandable com-*

munity standard (Important)

- Field(s): *not controlled by / in the scope of the Guidelines*
- Comments: -

Conclusions

This report describes the results of the FAIRification of the OpenAIRE Guidelines for CRIS Managers (v1.1.1) and assesses compliance with the FAIR principles in relation to the two main entities of research outputs: Publications and Products. We draw the following conclusions from the results of our assessment.

The Findability of publications is well supported by the Guidelines. Accessibility, Interoperability and Reusability are only in the scope of the Guidelines at the metadata level, and this has been found to be largely sufficient.

A similar picture emerges for Products, although the very broad spectrum of resource types should be taken into account. The metadata description remains at a generic level and does not provide specific characteristics for datasets, software, images and other sub-types. DOIs or other established identifiers – where available – provide access to additional metadata that is usually more specific to the resource type and subject discipline.

This assessment of the Guidelines necessarily remains at a theoretical level. It is the practice of operating a particular CRIS and managing its content that has the greatest impact on the level of FAIR compliance that particular data source can provide.

Good quality of metadata is therefore a key element in achieving FAIR compliance. The use of internationally established identifiers should be emphasised as they both provide access to the resources and form a bridge to additional metadata that complement the basic description of the resource in a CRIS.

We anticipate that further versions of this report will expand its scope to cover a wider range of entities that can be transferred using the OpenAIRE Guidelines for CRIS Managers metadata format and harvesting protocol.

Bibliography

- [DBRS18a] Jan Dvořák, Andrea Bollini, Laurent Rémy, and Jochen Schirrwagen. Openaire guidelines for crisis managers 1.1. <https://doi.org/10.5281/zenodo.2316420>, December 2018.
- [DBRS18b] Jan Dvořák, Andrea Bollini, Laurent Rémy, and Jochen Schirrwagen. Openaire guidelines for crisis managers 1.1 - elements. https://openaire-guidelines-for-cris-managers.readthedocs.io/en/v1.1.1/cris_elements_openaire.html, December 2018.
- [Gro20] FAIR Data Maturity Model Working Group. FAIR Data Maturity Model. Specification and Guidelines. <https://doi.org/10.15497/rda00050>, June 2020.
- [HJD15] Nikos Houssos, Brigitte Joerg, and Jan Dvořák. Openaire guidelines for crisis managers 1.0. <https://doi.org/10.5281/zenodo.17065>, July 2015.
- [HNAK21] Christian Hauschke, Serhii Nazarovets, Franziska Altemeier, and Nataliia Kaliuzhna. Roadmap to fair research information in open infrastructures, Apr 2021.
- [Ope22] OpenAIRE. OpenAIRE Website. <https://www.openaire.eu>, 2022.