FAIR Data Maturity Model Specification and Guidelines

DOI: 10.15497/rda00050

Co-chairs: Edit Herczog, Keith Russell, Shelley Stall

Published: 25th June 2020

Abstract: Findability, Accessibility, Interoperability and Reusability – the FAIR principles – intend to define a minimal set of related but independent and separable guiding principles and practices that enable both machines and humans to find, access, interoperate and re-use data and metadata. The FAIR principles were defined in 2016 in an article by Mark Wilkinson et. al1. FORCE112 and GO FAIR3 provide further information on the principles. The principles have to be considered as inspiring concepts but not strict rules. This means that they may lead to diverse interpretations and ambiguity.
To remedy the proliferation of FAIRness measurements based on different interpretations of the principles, the RDA Working Group "FAIR data maturity model" established in January 2019 aims to develop a common set of core assessment criteria for FAIRness, as an RDA Recommendation. In the course of 2019 and the first half of 2020, the WG established a set of indicators and maturity levels for those indicators.
As a result of the work, a first set of guidelines and a checklist related to the implementation of the indicators were produced, with the objective to further align the guidelines for evaluating FAIRness with the needs of the community.

Keywords: RDA; Recommendations; FAIR data; FAIR indicators.

Language: English

License: Attribution 4.0 International (CC BY 4.0)

Citation and Download: FAIR Data Maturity Model Working Group (2020): FAIR Data Maturity Model. Specification and Guidelines. DOI: 10.15497/rda00050

FAIR Data Maturity Model

Specification and Guidelines

2020

Proposed RDA Recommendation

Produced by: FAIR Data Maturity Model WG, 2019-2020

https://www.rd-alliance.org/groups/fair-data-maturity-model-wg

Document metadata

Identifier	DOI: 10.15497/rda00050
Title	FAIR Data Maturity Model: specification and guidelines
Description	The FAIR Data Maturity Model defines a set of indicators, their priorities and evaluation methods for the evaluation of the FAIR principles to be used as a common approach across assessment methodologies.
Date issued	2020-06-01
Version	Endorsed; final version
Contributors	RDA FAIR Data Maturity Model Working Group
Licence	Creative Commons Attribution 4.0 International (CC BY 4.0)

History of the discussions in the Working Group that led to this document can be viewed in the comments made in the Google document, the Public Review and GitHub repository.

Table of contents

1 Introduction 4
1.1 Context 4
1.2 Objective 4
1.3 Use of this document 4
2 Framework 6
2.1 Indicators 6
2.2 Priorities 7
2.3 Evaluation methods 7
3 Implementation 9
3.1 Flexibility of the framework 9
3.2 Indicators 9
3.3 Priorities 10
3.4 Evaluation methods 10
4 Indicators 11
4.1 List of indicators 11
4.2 Indicators for Findable 13
4.3 Indicators for Accessible 16
4.4 Indicators for Interoperable 21
4.5 Indicators for Reusable 26
5 Priorities 30
5.1 Essential indicators 31
5.2 Important indicators 32
5.3 Useful indicators 33
6 Evaluation methods 34
6.1 Measuring progress 34
6.2 Measuring pass-or-fail 35
6.3 Combined approach 37
7 Future maintenance 38
8 Glossary 39
9 Working Group materials 45

CC-BY 4.0 Creative Commons

Table of figures

Figure 1 Distribution of priorities 30
Figure 2 Five maturity levels 34
Figure 3 'Measuring progress' visualisation 35
Figure 4 Five compliance levels 36
Figure 5 'Measuring pass or fail' visualisation 36

List of tables

Table 1 FAIR data maturity model indicators 11
Table 2 Distribution of priorities per FAIR area 30
Table 3 Essential indicators 31
Table 4 Important indicators 32
Table 5 Useful indicators 33
Table 6 Glossary 39
Table 7 Working group presentations and meeting reports 45

CC-BY 4.0 Creative Commons

1 Introduction

1.1 Context

Findability, Accessibility, Interoperability and Reusability – the FAIR principles – intend to define a minimal set of related but independent and separable guiding principles and practices that enable both machines and humans to find, access, interoperate and re-use data and metadata. The FAIR principles were defined in 2016 in an article by Mark Wilkinson et. al[1]. FORCE11[2] and GO FAIR[3] provide further information on the principles. The principles have to be considered as inspiring concepts but not strict rules. This means that they may lead to diverse interpretations and ambiguity.

To remedy the proliferation of FAIRness measurements based on different interpretations of the principles, the RDA Working Group "FAIR data maturity model" established in January 2019 aims to develop a common set of core assessment criteria for FAIRness, as an RDA Recommendation. In the course of 2019 and the first half of 2020, the WG established a set of indicators and maturity levels for those indicators.

As a result of the work, a first set of guidelines and a checklist related to the implementation of the indicators were produced, with the objective to further align the guidelines for evaluating FAIRness with the needs of the community.

1.2 Objective

This document specifies indicators for assessing adherence to the FAIR principles. These indicators are designed for re-use in evaluation approaches and are accompanied by guidelines for their use. The guidelines are intended to assist evaluators to implement the indicators in the evaluation approach or tool they manage.

The exact way to evaluate data based on the core criteria is up to the owners of the evaluation approaches, taking into account the requirements of their community. The objective is to make sure that the indicators, the maturity levels and the prioritisation are understood in the same way. The maturity model is not meant as a "how to", but instead as a way to normalise assessment.

No part of this document is to be considered 'normative'; it intends to provide guidelines to inform assessment approaches but leaves the way it is implemented to the evaluator.

1.3 Use of this document

The FAIR data maturity model guidelines primarily address owners of (FAIR) assessment methodologies, including questionnaires and automated tools, as listed for example in FAIRassist[4].

Nevertheless, this document is not only restricted to these stakeholders. It may also be used by researchers, data service owners, funders and infrastructures in different scientific and research disciplines,

[1] Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). DOI: 10.1038/sdata.2016.18
[2] FORCE11. The FAIR data principles. https://www.force11.org/group/fairgroup/fairprinciples
[3] GO FAIR. FAIR Principles. https://www.go-fair.org/fair-principles/
[4] https://fairassist.org

CC-BY 4.0 Creative Commons

industry and the public sector, who are active and/or interested in the FAIR data principles and in particular in assessment criteria and methodologies for evaluating their real-life uptake and implementation level.

This document provides definitions and examples for every indicator in order to avoid confusion or ambiguity, and aims to provide a clear outline of the framework (i.e. indicators with their maturity levels and priorities) linking the indicators to the principles, and suggesting the way the indicators may be evaluated.

All terms which could be subject to different interpretation have been defined in a glossary, included in this document. Moreover, on the first mention of a term, a hyperlink to its definition in the glossary is provided.

The remainder of this document is structured as follows:

- Section 2 “Framework” describes the three main components of the FAIR data maturity model, namely (1) the indicators, (2) the priorities and (3) the evaluation method.
- Section 3 “Implementation” outlines some considerations as for the different interpretations depending on one’s perception according to the three main components of the FAIR data maturity model.
- Section 4 “Indicators” lays out the indicators per FAIR principle. Each indicator is defined in the most neutral and explicit fashion. Additionally, some context is provided to help assess the indicators.
- Section 5 “Priorities” presents the indicators and their respective priorities.
- Section 6 “Evaluation methods” gives details on possible evaluation methods.
- Section 7 “Future maintenance” indicates the way future maintenance is going to be done.
- Section 8 “Glossary” contains a list of terms that are used in this document and their definitions.
- Section 9 “Working Group materials” provides links to presentations and reports of the meetings of the Working Group.

CC-BY 4.0 Creative Commons

2 Framework

The framework of the FAIR Data Maturity Model consists of three elements:

- Indicators, i.e. the individual aspects of FAIRness that are evaluated
- Priorities, i.e. the relative importance of the indicators
- Evaluation method, i.e. the way that the results of the evaluation of the indicators can be given a value

2.1 Indicators

The indicators that are used in the FAIR data maturity model are derived from the FAIR principles and aim to formulate measurable aspects of each principle that can be used by evaluation approaches. They were developed in a process of decomposition of the text of the FAIR principles and explanations provided in the original article[5], a later clarifying article[6] and by GO-FAIR[7]. The FAIR principles are taken as they are; in other words, the indicators do not aim to extend or modify the principles but only cover aspects that are mentioned in the principles themselves or in the additional clarifications.

An indicator aims to measure the state or level of a digital resource with regard to a specific FAIR principle (e.g. F1, A2). The indicators developed as part of the RDA FAIR data maturity model working group have the sole purpose of answering the question 'What needs to be measured to assess the FAIRness of a digital object' and not 'How to measure the FAIRness of a digital object'. Nevertheless, the guidelines for each of the indicators in section 3 provide examples of how an indicator could be evaluated or where information to be used for the evaluation may be found.

The indicators can be used not only in the context of data in the conventional sense but also in the context of data-related algorithms, tools, workflows, protocols and other data-related services that are produced or managed, in as far as they are made available as digital objects. Nevertheless, it is important to specify that although the indicators may have value in the contexts cited above, the working group did not include the validation of these scenarios. Assessing data FAIRness is the raison d'être of the indicators, however, it is encouraged to reuse or adapt the indicators to evaluate other digital objects.

The approach for developing the indicators was to create an indicator for each aspect that could be distinguished in the description of the principle; for example, where the principle talks about a *persistent* and *globally unique* identifier, two indicators are defined, one to evaluate persistence and one to evaluate global uniqueness. Also, separate indicators are defined for metadata and for data, wherever a principle refers to '(meta)data' and the evaluation of the aspect for metadata is different from the evaluation for data.

The indicators presented in this document were developed by the RDA FAIR Data Maturity Working Group between January 2019 and June 2020. Further details are given in section 3.

[5] Mark D. Wilkinson et. al. The FAIR Guiding Principles for scientific data management and stewardship. 2016. DOI: 10.1038/sdata.2016.18
[6] Annika Jacobsen et. al. FAIR Principles: Interpretations and Implementation Considerations. 2020. DOI: 10.1162/dint_r_00024
[7] GO-FAIR. https://www.go-fair.org/fair-principles/

CC-BY 4.0 Creative Commons

2.2 Priorities

In addition to a list of indicators that can be used as a common set of aspects that can be tested to evaluate FAIRness, the Working Group also assigned priorities to these indicators. This is based on the understanding that some of the indicators are more important than others. Looking at this from the perspective of a data provider or publisher, the most important aspects would contribute the most toward improvement in terms of effort versus benefits.

Three levels of importance are defined:

- **Essential**: such an indicator addresses an aspect that is of the utmost importance to achieve FAIRness under most circumstances, or, conversely, FAIRness would be practically impossible to achieve if the indicator were not satisfied.
- **Important**: such an indicator addresses an aspect that might not be of the utmost importance under specific circumstances, but its satisfaction, if at all possible, would substantially increase FAIRness.
- **Useful**: such an indicator addresses an aspect that is nice-to-have but is not necessarily indispensable.

Further details of the priorities of the indicators are given in section 5.

2.3 Evaluation methods

The indicators defined in the FAIR Data Maturity Model can be used in several ways to evaluate data objects and their metadata.

First of all, the indicators are primarily intended to be used as a grounding set for evaluation methodologies, each of which can define their own questions or metrics, using the aspects underlying the indicators, so that the result of an evaluation is comparable to an evaluation using another methodology.

There are two different perspectives that evaluation methodologies can have:

1. **Measuring progress**: in this perspective, the emphasis lies on delivering a measure of the extent to which a resource under evaluation meets the requirements expressed in an indicator, giving an indication of which steps may be taken to achieve full satisfaction of an indicator.
2. **Measuring pass-or-fail**: in this perspective, the emphasis lies on determining whether a resource under evaluation meets the requirement of an indicator on a binary, pass-or-fail scale, providing a measure of how a resource under evaluation performs in reaching a particular target level of FAIRness.

The model may be used during the development of Research Data Management Plans before any data and metadata have been produced to specify the level of FAIRness that the resources are expected to achieve. It can also be used after the production of data resources to test what the achieved level of the resources is. Data producers, i.e. researchers, and data publishers can use the model to determine where their practices could be improved to achieve a higher level of FAIRness, while project managers and funding

CC-BY 4.0 Creative Commons

agencies can use the model to determine whether the resources achieve a pre-defined, expected level of FAIRness.

Details of the two evaluation methods are given in section 6.

CC-BY 4.0 Creative Commons

3 Implementation

3.1 Flexibility of the framework

Certain indicators may be less important or even irrelevant to some, less data-intensive disciplinary communities. Still, it is essential that different scholarly fields have equal chances to comply with the FAIR indicators. As such, one cannot highlight enough the importance of the responsible and careful implementation of these metrics to minimize unintended consequences.

Yet, flexibility is at the core of the FAIR principles. Several communities have often tailored their assessment approaches to fit their practices. For examples of such assessment approaches, see the Results of an Analysis of Existing FAIR Assessment Tools[8].

In this context, the framework was developed with a descriptive purpose rather than a prescriptive purpose. In other words, the usage of this document and more particularly the usage of the indicators and their priorities may potentially differ from one person to another, from one organisation to another or even from one community to another.

As such, certain aspects of the framework, which can be seen as conflicting due to different perceptions and interpretations, may be adapted or simply not considered.

3.2 Indicators

Disparities across communities led to discussion as to the relevance of some indicators;

- In alignment with the interpretations of the Principle F1[9], the Working Group decided that both metadata and data should be identified with a PID in order to be found. However, requiring that a PID should be assigned separately to both data and metadata may not align with some existing practice where a PID resolves to a landing page that may contain the metadata of the object and the identifier (e.g., URL) to access the actual data contents. Assumptions on the way that data and metadata objects are identified in practice will influence the implementation and results of automated assessments, and therefore it is up to the evaluator to determine which solution suits best.
- With regard to the divergent interpretations, some FAIR principles explicitly state '*where necessary*'. Indicators tied to such sub-principles give the possibility to communities to choose whether they are applicable or not. This observation is also valid for indicators deemed not relevant for a specific community.

[8] RDA FAIR Data Maturity Model Working Group. Results of an Analysis of Existing FAIR Assessment Tools. DOI: 10.15497/RDA00035

[9] "Principle F1 states that digital resources, i.e., data and metadata, must be assigned a globally unique and persistent identifier in order to be found and resolved by computers" in: Jacobsen, A., de Miranda Azevedo, R., Juty, N. et. al. FAIR Principles: Interpretations and Implementation Considerations. Data Intelligence. Vol. 2. Issue 1-2 (2020). DOI: 10.1162/dint_r_00024.

CC-BY 4.0 Creative Commons

3.3 Priorities

Priority levels may also be open to interpretation, and priorities may be changed depending on the context.

As an example, it is understood from principle I1[10] that knowledge representation should be machine-understandable. Yet, data coming from humanities fields, especially from outside of Digital Humanities, will often not be expressed in a machine understandable knowledge representation (RDF[11], SKOS[12] or LOD[13]) by nature but instead, it is often expressed in natural language, even if encoded using machine readable methods (e.g. TEI[14]). Therefore, it becomes quite clear that the indicator treating machine-understandable knowledge representation will be less relevant according to the humanities.

3.4 Evaluation methods

Throughout the development of the indicators and priorities, the evaluation method was developed and refined. The evaluation method was designed solely to propose a means to evaluate the FAIRness of a digital object and provide results in a meaningful way. However, it is bound to evolve together with the rest of the framework.

While refining the evaluation method, some shortfalls and considerations have been identified;

- The weight of each FAIR area is not the same. For starters, all FAIR areas have a different number of indicators, which introduces a bias. Furthermore, these FAIR areas do not have the same proportions of indicators ranked essential, important or useful. Essential indicators are even completely absent from the Interoperability FAIR area – which makes it possible to reach a high level of FAIRness in the absence of any essential Interoperability capacity when using the proposed pass-and-fail method.
- Upgrading or downgrading priorities according to the context – as explained in the section above – may have repercussions on the evaluation method and more specifically on the results it yields.
- It is likely that FAIR practices can vary a lot from one community to another, eventually because their requirements can be different.

Finally, striking out an indicator or even downgrading the indicator's priority are not devoid of consequences. It will inevitably impact the comparison of manual and automated test outcomes. Therefore, is it recommended to the communities to carefully consider how each indicator applies to their data and their practices. Yet, it is essential to ensure a high level of compliance to the framework to produce results that can be combined and compared amongst peers.

[10] Jacobsen, A., de Miranda Azevedo, R., Juty, N. et. al. FAIR Principles: Interpretations and Implementation Considerations. Data Intelligence. Vol. 2. Issue 1-2 (2020). DOI: 10.1162/dint_r_00024.
[11] W3C. Resource Description Framework (RDF). https://www.w3.org/RDF/
[12] W3C. SKOS Simple Knowledge Organization System. https://www.w3.org/2004/02/skos/
[13] W3C. Linked Open Data. https://www.w3.org/egov/wiki/Linked_Open_Data
[14] TEI: Text Encoding Initiative. https://tei-c.org/

CC-BY 4.0 Creative Commons

4 Indicators

4.1 List of indicators

Table 1 FAIR data maturity model indicators

FAIR	ID	Indicator		Priority
F1	RDA-F1-01M	Metadata is identified by a persistent identifier	●●●	Essential
F1	RDA-F1-01D	Data is identified by a persistent identifier	●●●	Essential
F1	RDA-F1-02M	Metadata is identified by a globally unique identifier	●●●	Essential
F1	RDA-F1-02D	Data is identified by a globally unique identifier	●●●	Essential
F2	RDA-F2-01M	Rich metadata is provided to allow discovery	●●●	Essential
F3	RDA-F3-01M	Metadata includes the identifier for the data	●●●	Essential
F4	RDA-F4-01M	Metadata is offered in such a way that it can be harvested and indexed	●●●	Essential
A1	RDA-A1-01M	Metadata contains information to enable the user to get access to the data	●●	Important
A1	RDA-A1-02M	Metadata can be accessed manually (i.e. with human intervention)	●●●	Essential
A1	RDA-A1-02D	Data can be accessed manually (i.e. with human intervention)	●●●	Essential
A1	RDA-A1-03M	Metadata identifier resolves to a metadata record	●●●	Essential
A1	RDA-A1-03D	Data identifier resolves to a digital object	●●●	Essential
A1	RDA-A1-04M	Metadata is accessed through standardised protocol	●●●	Essential
A1	RDA-A1-04D	Data is accessible through standardised protocol	●●●	Essential
A1	RDA-A1-05D	Data can be accessed automatically (i.e. by a computer program)	●●	Important
A1.1	RDA-A1.1-01M	Metadata is accessible through a free access protocol	●●●	Essential
A1.1	RDA-A1.1-01D	Data is accessible through a free access protocol	●●	Important
A1.2	RDA-A1.2-01D	Data is accessible through an access protocol that supports authentication and authorisation	●	Useful
A2	RDA-A2-01M	Metadata is guaranteed to remain available after data is no longer available	●●●	Essential
I1	RDA-I1-01M	Metadata uses knowledge representation expressed in standardised format	●●	Important
I1	RDA-I1-01D	Data uses knowledge representation expressed in standardised format	●●	Important
I1	RDA-I1-02M	Metadata uses machine-understandable knowledge representation	●●	Important
I1	RDA-I1-02D	Data uses machine-understandable knowledge representation	●●	Important
I2	RDA-I2-01M	Metadata uses FAIR-compliant vocabularies	●●	Important
I2	RDA-I2-01D	Data uses FAIR-compliant vocabularies	●	Useful
I3	RDA-I3-01M	Metadata includes references to other metadata	●●	Important
I3	RDA-I3-01D	Data includes references to other data	●	Useful
I3	RDA-I3-02M	Metadata includes references to other data	●	Useful

CC-BY 4.0 Creative Commons

FAIR	ID	Indicator		Priority
I3	RDA-I3-02D	Data includes qualified references to other data	●	Useful
I3	RDA-I3-03M	Metadata includes qualified references to other metadata	●●	Important
I3	RDA-I3-04M	Metadata include qualified references to other data	●	Useful
R1	RDA-R1-01M	Plurality of accurate and relevant attributes are provided to allow reuse	●●●	Essential
R1.1	RDA-R1.1-01M	Metadata includes information about the licence under which the data can be reused	●●●	Essential
R1.1	RDA-R1.1-02M	Metadata refers to a standard reuse licence	●●	Important
R1.1	RDA-R1.1-03M	Metadata refers to a machine-understandable reuse licence	●●	Important
R1.2	RDA-R1.2-01M	Metadata includes provenance information according to community-specific standards	●●	Important
R1.2	RDA-R1.2-02M	Metadata includes provenance information according to a cross-community language	●	Useful
R1.3	RDA-R1.3-01M	Metadata complies with a community standard	●●●	Essential
R1.3	RDA-R1.3-01D	Data complies with a community standard	●●●	Essential
R1.3	RDA-R1.3-02M	Metadata is expressed in compliance with a machine-understandable community standard	●●●	Essential
R1.3	RDA-R1.3-02D	Data is expressed in compliance with a machine-understandable community standard	●●	Important

CC-BY 4.0 Creative Commons

4.2 Indicators for Findable

RDA-F1-01M Metadata identified by a persistent identifier	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *F1 (meta)data are assigned a globally unique and eternally persistent identifier*.
Description of the indicator RDA1-F1-01M	This indicator evaluates whether or not the metadata is identified by a persistent identifier. A persistent identifier ensures that the metadata will remain findable over time and reduces the risk of broken links.
Assessment details	The persistence of an identifier is determined by the commitment of the organisation that assigns and manages the identifier, so the evaluation of this indicator needs to take into account the persistence policy of that organisation. Such a commitment could be expressed by a university or research institute, by a research infrastructure or by an organisation that issues formal identifiers, such as the International DOI Foundation. A possible way to evaluate this indicator is to verify that the identifier used for the metadata is listed in a registry service like the RDA-endorsed FAIRsharing.[15]

RDA-F1-01D Data identified by a persistent identifier	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *F1 (meta)data are assigned a globally unique and eternally persistent identifier*.
Description of the indicator RDA-F1-01D	This indicator evaluates whether or not the data is identified by a persistent identifier. A persistent identifier ensures that the data will remain findable over time and reduces the risk of broken links.
Assessment details	The persistence of an identifier is determined by the commitment of the organisation that assigns and manages the identifier, so the evaluation of this indicator needs to take into account the persistence policy of that organisation. Such a commitment could be expressed by a university or research institute, by a research infrastructure or by an organisation that issues formal identifiers, such as the International DOI Foundation. A possible way to evaluate this indicator is to verify that the identifier used for the data is listed in a registry service like the RDA-endorsed FAIRsharing.

[15] https://fairsharing.org/standards/?q=&selected_facets=type_exact:identifier%20schema

CC-BY 4.0 Creative Commons

RDA-F1-02M Metadata is identified by a globally unique identifier	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *F1 (meta)data are assigned a globally unique and eternally persistent identifier*.
Description of the indicator RDA-F1-02M	The indicator serves to evaluate whether the identifier of the metadata is globally unique, i.e. that there are no two identical identifiers that identify different metadata records.
Assessment details	Global uniqueness of identifiers should be evaluated based on a description of how identifiers are assigned. Such a description should make it clear that the mechanism for assigning identifiers cannot possibly assign the same identifier to different resources, or assign an identifier that has already been assigned via some other mechanism/organisation. A possible way to evaluate this indicator is to verify that the identifier used for the data is listed in a registry service like the RDA-endorsed FAIRsharing.

RDA-F1-02D Data is identified by a globally unique identifier	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *F1 (meta)data are assigned a globally unique and eternally persistent identifier*.
Description of the indicator RDA-F1-02D	The indicator serves to evaluate whether the identifier of the data is globally unique, i.e. that there are no two people that would use that same identifier for two different digital objects.
Assessment details	Global uniqueness of identifiers should be evaluated based on a description of how identifiers are assigned. Such a description should make it clear that the mechanism for assigning identifiers cannot possibly assign the same identifier to different resources or assign an identifier that has already been assigned via some other mechanism/organisation. A possible way to evaluate this indicator is to verify that the identifier used for the data is listed in a registry service like the RDA-endorsed FAIRsharing.

RDA-F2-01M Rich metadata is provided to allow discovery	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *F2: Data are described with rich metadata*.
Description of the indicator RDA-F2-01M	The indicator is about the presence of metadata, but also about how much metadata is provided and how well the provided metadata supports discovery.
Assessment details	This indicator can be evaluated by verifying that metadata is provided. The amount of metadata to be provided may also be part of the metadata policy of the repository where the data is published.

CC-BY 4.0 Creative Commons

RDA-F3-01M Metadata includes the identifier for the data	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *F3: Metadata clearly and explicitly include the identifier of the data they describe.*
Description of the indicator RDA-F3-01M	The indicator deals with the inclusion of the reference (i.e. the identifier) of the digital object in the metadata so that the digital object can be accessed.
Assessment details	This indicator can be evaluated by verifying that the identifier of the data is included in the metadata element that is specified for that purpose in the metadata standard used, for example in an "about" or "describes" predicate, or a Link Relation[16] such as "describes"/"describedBy".

RDA-F4-01M Metadata is offered in such a way that it can be harvested and indexed	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *F4: (Meta)data are registered or indexed in a searchable resource*.
Description of the indicator RDA-F4-01M	The indicator tests whether the metadata is offered in such a way that it can be indexed. In some cases, metadata could be provided together with the data to a local institutional repository or to a domain-specific or regional portal, or metadata could be included in a landing page where it can be harvested by a search engine. The indicator remains broad enough on purpose not to limit the way how and by whom the harvesting and indexing of the data might be done.
Assessment details	This indicator can be evaluated by verifying that the metadata is made available for indexing. This is the case when the metadata is in fact harvested and indexed, for example in a general search engine or in a more restricted index, such as an institutional repository or a discipline-specific portal.

[16] https://www.iana.org/assignments/link-relations/link-relations.xhtml

CC-BY 4.0 Creative Commons

4.3 Indicators for Accessible

RDA-A1-01M Metadata contains information to enable the user to get access to the data	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *A1: (Meta)data are retrievable by their identifier using a standardised communication protocol*.
Description of the indicator RDA-A1-01M	The indicator refers to the information that is necessary to allow the requester to gain access to the digital object. It is about (i) restrictions to access the data (i.e. access to the data may be open, restricted or closed), (ii) the actions to be taken by a person who is interested to access the data, in particular when the data has not been published on the Web or (iii) specifications that the resources are available through specified authentication/authorisation system including single sign-on providers such as eduGAIN[17] or through specialised solutions such as proposed for EPOS[18].
Assessment details	The indicator can be evaluated by looking at the information that is provided in the metadata about (i) access conditions according to the metadata standard used, (ii) information that describes the actions to be taken. This can be included in the metadata or in some other place, for example on a landing page of the digital object or (iii) the requirements to be satisfied in order to gain access to the data. This may be machine-understandable information in which case the evaluation could be automated by processing the information. If it is human-readable information, the evaluator can follow the instructions and verify that this allows access to the data.

[17] https://edugain.org/
[18] https://www.epos-ip.org/progress-epos-authentication-and-authorisation-solutions

CC-BY 4.0 Creative Commons

RDA-A1-02M Metadata can be accessed manually (i.e. with human intervention)	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *A1: (Meta)data are retrievable by their identifier using a standardised communication protocol*.
Description of the indicator RDA-A1-02M	The indicator refers to any human interactions that are needed if the requester wants to access metadata. The FAIR principle refers mostly to automated interactions where a machine is able to access the metadata, but there may also be metadata that require human interactions. This may be important in cases where the metadata itself contains sensitive information. Human interaction might involve sending an e-mail to the metadata owner, or calling by telephone to receive instructions.
Assessment details	The indicator can be evaluated by looking for information about the way that metadata can be accessed with human intervention, either in documentation, for example in a landing page, or in metadata about the metadata in cases where there is multi-layered metadata, for example using CatalogRecord in DCAT.

RDA-A1-02D Data can be accessed manually (i.e. with human intervention)	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *A1: (Meta)data are retrievable by their identifier using a standardised communication protocol*.
Description of the indicator RDA-A1-02D	The indicator refers to any human interactions that are needed if the requester wants to access the digital object. The FAIR principle refers mostly to automated interactions where a machine is able to access the digital object, but there may also be digital objects that require human interactions, such as clicking on a link on a landing page, sending an e-mail to the data owner, or even calling by telephone.
Assessment details	The indicator can be evaluated by looking for information in the metadata that describes how access to the digital object can be obtained through human intervention.

CC-BY 4.0 Creative Commons

RDA-A1-03M Metadata identifier resolves to a metadata record	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *A1: (Meta)data are retrievable by their identifier using a standardised communication protocol*.
Description of the indicator RDA-A1-03M	This indicator is about the resolution of the metadata identifier. The identifier assigned to the metadata should be associated with a resolution service that enables access to the metadata record.
Assessment details	The indicator can be evaluated by checking that the metadata can be accessed using its identifier. The evaluator or evaluation tool may also want to verify that the resolution delivers the correct metadata record.

RDA-A1-03D Data identifier resolves to a digital object	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *A1: (Meta)data are retrievable by their identifier using a standardised communication protocol*.
Description of the indicator RDA-A1-03D	This indicator is about the resolution of the identifier that identifies the digital object. The identifier assigned to the data should be associated with a formally defined retrieval/resolution mechanism that enables access to the digital object or provides access instructions for access in the case of human-mediated access. The FAIR principle and this indicator do not say anything about the mutability or immutability of the digital object that is identified by the data identifier -- this is an aspect that should be governed by a persistence policy of the data provider.
Assessment details	The indicator can be evaluated by invoking the mechanism specific to the protocol (e.g. GET for HTTP) and verifying that this delivers the digital object.

RDA-A1-04M Metadata is accessed through standardised protocol	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *A1: (Meta)data are retrievable by their identifier using a standardised communication protocol.*
Description of the indicator RDA-A1-04M	The indicator concerns the protocol through which the metadata is accessed and requires the protocol to be defined in a standard.
Assessment details	This indicator can be evaluated by looking at the way the metadata can be accessed. Common metadata access protocols are HTTP and FTP, Atom[19], OAI-PMH[20] and Web Services Metadata Exchange[21].

[19] Internet Engineering Task Force. The Atom Publishing Protocol. https://tools.ietf.org/html/rfc5023
[20] Open Archives Initiative. Protocol for Metadata Harvesting. https://www.openarchives.org/pmh/
[21] W3C. Web Services Metadata Exchange (WS-MetadataExchange). https://www.w3.org/TR/ws-metadata-exchange/

CC-BY 4.0 Creative Commons

RDA-A1-04D Data is accessible through standardised protocol	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *A1: (Meta)data are retrievable by their identifier using a standardised communication protocol.*
Description of the indicator RDA-A1-04D	The indicator concerns the protocol through which the digital object is accessed and requires the protocol to be defined in a standard.
Assessment details	This indicator can be evaluated by looking at the way the data can be accessed. Common data access protocols are HTTP and FTP, DAP[22] and JSON-RPC[23].

RDA-A1-05D Data can be accessed automatically (i.e. by a computer program)	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *A1: (Meta)data are retrievable by their identifier using a standardised communication protocol.*
Description of the indicator RDA-A1-05D	The indicator refers to automated interactions between machines to access digital objects. The way machines interact and grant access to the digital object will be evaluated by the indicator.
Assessment details	This indicator can be evaluated by resolving the link to the data, e.g. by resolving the persistent identifier and verifying that the data is reached. In the common case that the identifier is an HTTP URI, this can be done using the HTTP GET method. The evaluator or evaluation tool may also want to verify that the resolution delivers the correct data.

RDA-A1.1-01M Metadata is accessible through a free access protocol	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *A1.1: The protocol is open, free and universally implementable*.
Description of the indicator RDA-A1.1-01M	The indicator tests that the protocol that enables the requester to access metadata can be freely used. Such free use of the protocol enhances data reusability.
Assessment details	The indicator can be evaluated on the basis of information provided about whether the use of the protocol is free of charge. Common examples are HTTP and FTP.

[22] https://cdn.earthdata.nasa.gov/conduit/upload/512/ESE-RFC-004v1.1.pdf
[23] https://www.jsonrpc.org/

CC-BY 4.0 Creative Commons

RDA-A1.1-01D Data is accessible through a free access protocol	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *A1.1: The protocol is open, free and universally implementable*.
Description of the indicator RDA-A1.1-01D	The indicator requires that the protocol can be used free of charge which facilitates unfettered access.
Assessment details	This indicator can be evaluated by verifying that the protocol is free of charge. This is the case for most protocols in use, for example HTTP and FTP.

RDA-A1.2-01D Data is accessible through an access protocol that supports authentication and authorisation	● Useful
Principle to which the indicator relates	This indicator is linked to the following principle: *A1.2: The protocol allows for an authentication and authorisation where necessary.*
Description of the indicator RDA-A1.2-01D	The indicator requires that if the data or local environment indicates a degree of additional protection then the access protocol must support authentication and authorisation of people and/or machines?
Assessment details	The indicator can be evaluated by assessing whether an authentication and authorisation process is present in the protocol (e.g. HMAC).

RDA-A2-01M Metadata is guaranteed to remain available after data is no longer available	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *A2: Metadata should be accessible even when the data is no longer available*.
Description of the indicator RDA-A2-01M	The indicator intends to verify that information about a digital object is still available after the object has been deleted or otherwise has become unavailable. If possible, the metadata that remains available should also indicate why the object is no longer available.
Assessment details	The indicator can be evaluated on the basis of information provided about the life cycle of metadata and data, which should indicate that the metadata will remain available if the data is no longer available. This information is likely to be available from the repository where the metadata and data are stored.

CC-BY 4.0 Creative Commons

4.4 Indicators for Interoperable

RDA-I1-01M Metadata uses knowledge representation expressed in standardised format	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation*.
Description of the indicator RDA-I1-01M	The indicator serves to determine that an appropriate standard is used to express knowledge, for example, controlled vocabularies for subject classifications.
Assessment details	The indicator can be evaluated by looking at information describing the way metadata values are expressed using controlled vocabularies, verifying that the standard used is appropriate for the domain and the type of digital object. Deciding on the appropriateness of the knowledge representation may be based on its inclusion in a registry like the one developed by FAIRsharing.

RDA-I1-01D Data uses knowledge representation expressed in standardised format	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation*.
Description of the indicator RDA-I1-01D	The indicator serves to determine that an appropriate standard is used to express knowledge, in particular the data model and format.
Assessment details	The indicator can be evaluated by looking at information about the data model and format, verifying that the standard used is appropriate for the domain and the type of digital object. Deciding on the appropriateness of the knowledge representation may be based on its inclusion in a registry like the one developed by FAIRsharing.

CC-BY 4.0 Creative Commons

RDA-I1-02M Metadata uses machine-understandable knowledge representation	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation*.
Description of the indicator RDA-I1-02M	This indicator focuses on the machine-understandability aspect of the metadata. This means that metadata should be readable and thus interoperable for machines without any requirements such as specific translators or mappings.
Assessment details	This indicator can be evaluated by looking at the knowledge representation model used for the expression of the metadata. Examples are RDF, OWL, JSON-LD and SKOS. Information about models and formats can be looked up in a registry like the RDA-endorsed FAIRsharing (see for example : https://fairsharing.org/standards/?q=&selected_facets=type_exact:model/format).

RDA-I1-02D Data uses machine-understandable knowledge representation	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation*.
Description of the indicator RDA-I1-02D	This indicator focuses on the machine-understandability aspect of the data. This means that data should be readable and thus interoperable for machines without any requirements such as specific translators or mappings.
Assessment details	This indicator can be evaluated by looking at the knowledge representation model used for the expression of the data. Examples are RDF, OWL, JSON-LD, Data Cube[24], the Generalized Data Model for clinical research[25] and SKOS. Information about models and formats can be looked up in a registry like the RDA endorsed FAIRsharing (see for example: https://fairsharing.org/standards/?q=&selected_facets=type_exact:model/format).

[24] W3C. The RDF Data Cube Vocabulary. W3C Recommendation 16 January 2014. https://www.w3.org/TR/vocab-data-cube/

[25] Danese, M.D., Halperin, M., Duryea, J. et al. The Generalized Data Model for clinical research. BMC Med Inform Decis Mak 19, 117 (2019). DOI: 10.1186/s12911-019-0837-5

CC-BY 4.0 Creative Commons

RDA-I2-01M Metadata uses FAIR-compliant vocabularies	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *I2: (Meta)data use vocabularies that follow the FAIR principles*.
Description of the indicator RDA-I2-01M	The indicator requires the vocabulary used for the metadata to conform to the FAIR principles, and at least be documented and resolvable using globally unique and persistent identifiers. The documentation needs to be easily findable and accessible.
Assessment details	The indicator can be evaluated by verifying that each of the vocabularies used in the metadata is documented and resolvable using globally unique and persistent identifiers, with the documentation being easily findable and accessible. Typically, the reference to the specification of the vocabularies used will be included in the documentation of the digital object or the repository where it is kept.

RDA-I2-01D Data uses FAIR-compliant vocabularies	● Useful
Principle to which the indicator relates	This indicator is linked to the following principle: *I2: (Meta)data use vocabularies that follow the FAIR principles*.
Description of the indicator RDA-I2-01D	The indicator requires the controlled vocabulary used for the data to conform to the FAIR principles, and at least be documented and resolvable using globally unique and persistent identifiers. The documentation needs to be easily findable and accessible.
Assessment details	The indicator can be evaluated by verifying that each of the vocabularies used in the data is documented and resolvable using globally unique and persistent identifiers, with the documentation being easily findable and accessible. Typically, the reference to the specification of the vocabularies used will be included in the documentation of the digital object or the repository where it is kept.

CC-BY 4.0 Creative Commons

RDA-I3-01M Metadata includes references to other metadata	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *I3: (Meta)data include qualified references to other (meta)data*.
Description of the indicator RDA-I3-01M	The indicator is about the way that metadata is connected to other metadata, for example through links to information about organisations, people, places, projects or time periods that are related to the digital object that the metadata describes.
Assessment details	The indicator can be evaluated by looking at the occurrence of references to other metadata, for example ORCID[26] for people or Geonames[27] for places.

RDA-I3-01D Data includes references to other data	● Useful
Principle to which the indicator relates	This indicator is linked to the following principle: *I3: (Meta)data include qualified references to other (meta)data*.
Description of the indicator RDA-I3-01D	This indicator is about the way data is connected to other data, for example linking to previous or related research data that provides additional context to the data.
Assessment details	The indicator can be evaluated by looking at the presence of references to other data in the data. For example, there may be links to other resources in cells in a spreadsheet, or in RDF-based data.

RDA-I3-02M Metadata includes references to other data	● Useful
Principle to which the indicator relates	This indicator is linked to the following principle: *I3: (Meta)data include qualified references to other (meta)data*.
Description of the indicator RDA-I3-02M	This indicator is about the way metadata is connected to other data, for example linking to previous or related research data that provides additional context to the data. Please note that this is not about the link from the metadata to the data it describes; that link is considered in principle F3 and in indicator RDA-F3-01M.
Assessment details	The indicator can be evaluated by looking at the presence of references to other data in the metadata.

[26] https://orcid.org/
[27] https://www.geonames.org/

CC-BY 4.0 Creative Commons

RDA-I3-02D Data includes qualified references to other data	● Useful
Principle to which the indicator relates	This indicator is linked to the following principle: I3: (Meta)data include qualified references to other (meta)data.
Description of the indicator RDA-I3-02D	This indicator is about the way data is connected to other data. The references need to be qualified which means that the relationship role of the related resource is specified, for example that a particular link is a specification of a unit of measurement, or the identification of the sensor with which the measurement was done.
Assessment details	The indicator can be evaluated by validating the presence of references with specification of the relationship role that the related resource has with the data object.

RDA-I3-03M Metadata includes qualified references to other metadata	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *I3: (Meta)data include qualified references to other (meta)data*.
Description of the indicator RDA-I3-03M	This indicator is about the way metadata is connected to other metadata, for example to descriptions of related resources that provide additional context to the data. The references need to be qualified which means that the relationship of the related resource is specified, for example person Y *is the author of* dataset X.
Assessment details	The indicator can be evaluated by looking at the presence of references with specification of the relationship that the related resource has to the described resource.

RDA-I3-04M Metadata includes qualified references to other data	● Useful
Principle to which the indicator relates	This indicator is linked to the following principle: *I3: (Meta)data include qualified references to other (meta)data*.
Description of the indicator RDA-I3-04M	This indicator is about the way metadata is connected to other data. The references need to be qualified which means that the relationship role of the related resource is specified, for example dataset X is derived from dataset Y.
Assessment details	The indicator can be evaluated by looking at the presence of references with specification of the relationship role that the related resource has with the described resource.

CC-BY 4.0 Creative Commons

4.5 Indicators for Reusable

RDA-R1-01M Plurality of accurate and relevant attributes are provided to allow reuse	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *R1: (Meta)data are richly described with a plurality of accurate and relevant attributes*.
Description of the indicator RDA-R1-01M	The indicator concerns the quantity but also the quality of metadata provided in order to enhance data reusability.
Assessment details	This indicator can be evaluated with the help of standards registries such as the RDA-endorsed FAIRsharing (see for example: https://fairsharing.org/standards/?q=/format&selected_facets=type_exact:reporting%20guideline).

RDA-R1.1-01M Metadata includes information about the licence under which the data can be reused	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *R1.1: (Meta)data are released with a clear and accessible data usage license*. More information about that principle can be found here.
Description of the indicator RDA-R1.1-01M	This indicator is about the information that is provided in the metadata related to the conditions (e.g. obligations, restrictions) under which data can be reused. In the absence of licence information, data cannot be reused.
Assessment details	This indicator can be evaluated by looking in the metadata for licence information. This information may be in human-readable text; machine-understandability of the information is covered in indicator RDA-R1.1-03M.

RDA-R1.1-02M Metadata refers to a standard reuse licence	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *R1.1: (Meta)data are released with a clear and accessible data usage license*.
Description of the indicator RDA-R1.1-02M	This indicator requires the reference to the conditions of reuse to be a standard licence, rather than a locally defined licence.
Assessment details	The indicator can be evaluated by verifying that the licence is indeed a standard licence. Examples of standard licences are: Creative Commons licences, Open Data Commons.

CC-BY 4.0 Creative Commons

RDA-R1.1-03M Metadata refers to a machine-understandable reuse licence	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *R1.1: (Meta)data are released with a clear and accessible data usage license.*
Description of the indicator RDA-R1.1-03M	This indicator is about the way that the reuse licence is expressed. Rather than being a human-readable text, the licence should be expressed in such a way that it can be processed by machines, without human intervention, for example in automated searches.
Assessment details	The indicator can be evaluated by verifying that the link to the licence resolves to a machine-understandable expression of the conditions. An example of such a machine-understandable expression is the RDF expression of Creative Commons licences, or the various serialisations of the Open Data Rights Language (ODRL).

RDA-R1.2-01M Metadata includes provenance information according to community-specific standards	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *R1.2: (Meta)data are associated with detailed provenance.*
Description of the indicator RDA-	This indicator requires the metadata to include information about the provenance of the data, i.e. information about the origin, history or workflow that generated the data, in a way that is compliant with the standards that are used in the community for which the data is curated.
Assessment details	The indicator can be evaluated by verifying that the provenance information follows the community standard. A RDA-endorsed service like FAIRsharing could be helpful to identify the relevant standards.

RDA-R1.2-02M Metadata includes provenance information according to a cross-community language	● Useful
Principle to which the indicator relates	This indicator is linked to the following principle: *R1.2: (Meta)data are associated with detailed provenance.*
Description of the indicator RDA-R1.2-02M	This indicator requires that the metadata provides provenance information according to a cross-domain language.
Assessment details	The indicator can be evaluated by assessing whether a cross-domain language is used for provenance information (such as PROV-O).

CC-BY 4.0 Creative Commons

RDA-R1.3-01M Metadata complies with a community standard	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *R1.3: (Meta)data meet domain-relevant community standards.*
Description of the indicator RDA-R1.3-01M	This indicator requires that metadata complies with community standards.
Assessment details	The indicator can be evaluated by verifying that the metadata follows a community standard. A service like the RDA-endorsed FAIRsharing or the Metadata Standards Catalog could be helpful to identify the relevant standards.

RDA-R1.3-01D Data complies with a community standard	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *R1.3: (Meta)data meet domain-relevant community standards.*
Description of the indicator RDA-R1.3-01D	This indicator requires that data complies with community standards.
Assessment details	The indicator can be evaluated by verifying that the data follows a community standard. A service like the RDA-endorsed FAIRsharing could be helpful to identify the relevant standards.

RDA-R1.3-02M Metadata is expressed in compliance with a machine-understandable community standard	●●● Essential
Principle to which the indicator relates	This indicator is linked to the following principle: *R1.3: (Meta)data meet domain-relevant community standards.*
Description of the indicator RDA-R1.3-02M	This indicator requires that the metadata follows a community standard that has a machine-understandable expression.
Assessment details	This indicator can be evaluated by verifying that the community standard used for the metadata has a machine-understandable expression.

CC-BY 4.0 Creative Commons

RDA-R1.3-02D Data is expressed in compliance with a machine-understandable community standard	●● Important
Principle to which the indicator relates	This indicator is linked to the following principle: *R1.3: (Meta)data meet domain-relevant community standards.*
Description of the indicator RDA-R1.3-02D	This indicator requires that the data follows a community standard that has a machine-understandable expression.
Assessment details	This indicator can be evaluated by verifying that the community standard used for the data has a machine-understandable expression.

CC-BY 4.0 Creative Commons

5 Priorities

Among the set of indicators for FAIRness, 20 of the indicators are classified as *Essential*, 14 *Important* and 7 *Useful (Figure 1)*.

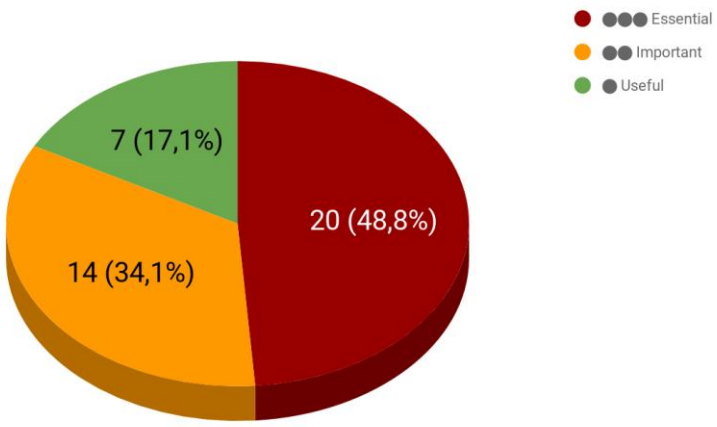

Figure 1 Distribution of priorities

Table 2 Distribution of priorities per FAIR area

	Principle				
Priority	Findable	Accessible	Interoperable	Reusable	Grand Total
Essential	7	8	0	5	20
Important	0	3	7	4	14
Useful	0	1	5	1	7
Grand Total	7	**12**	**12**	**10**	**41**

CC-BY 4.0 Creative Commons

5.1 Essential indicators

Table 3 Essential indicators

ID	Essential indicators
RDA-F1-01M	Metadata is identified by a persistent identifier
RDA-F1-01D	Data is identified by a persistent identifier
RDA-F1-02M	Metadata is identified by a globally unique identifier
RDA-F1-02D	Data is identified by a globally unique identifier
RDA-F2-01M	Rich metadata is provided to allow discovery
RDA-F3-01M	Metadata includes the identifier for the data
RDA-F4-01M	Metadata is offered in such a way that it can be harvested and indexed
RDA-A1-02M	Metadata can be accessed manually (i.e. with human intervention)
RDA-A1-02D	Data can be accessed manually (i.e. with human intervention)
RDA-A1-03M	Metadata identifier resolves to a metadata record
RDA-A1-03D	Data identifier resolves to a digital object
RDA-A1-04M	Metadata is accessed through standardised protocol
RDA-A1-04D	Data is accessible through standardised protocol
RDA-A1.1-01M	Metadata is accessible through a free access protocol
RDA-A2-01M	Metadata is guaranteed to remain available after data is no longer available
RDA-R1-01M	Plurality of accurate and relevant attributes are provided to allow reuse
RDA-R1.1-01M	Metadata includes information about the licence under which the data can be reused
RDA-R1.3-01M	Metadata complies with a community standard
RDA-R1.3-01D	Data complies with a community standard
RDA-R1.3-02M	Metadata is expressed in compliance with a machine-understandable community standard

CC-BY 4.0 Creative Commons

5.2 Important indicators

Table 4 Important indicators

ID	Important indicators
RDA-A1-01M	Metadata contains information to enable the user to get access to the data
RDA-A1-05D	Data can be accessed automatically (i.e. by a computer program)
RDA-A1.1-01D	Data is accessible through a free access protocol
RDA-I1-01M	Metadata uses knowledge representation expressed in standardised format
RDA-I1-01D	Data uses knowledge representation expressed in standardised format
RDA-I1-02M	Metadata uses machine-understandable knowledge representation
RDA-I1-02D	Data uses machine-understandable knowledge representation
RDA-I2-01M	Metadata uses FAIR-compliant vocabularies
RDA-I3-01M	Metadata includes references to other metadata
RDA-I3-03M	Metadata includes qualified references to other metadata
RDA-R1.1-02M	Metadata refers to a standard reuse licence
RDA-R1.1-03M	Metadata refers to a machine-understandable reuse licence
RDA-R1.2-01M	Metadata includes provenance information according to community-specific standards
RDA-R1.3-02D	Data is expressed in compliance with a machine-understandable community standard

CC-BY 4.0 Creative Commons

5.3 Useful indicators

Table 5 Useful indicators

ID	Useful indicators
RDA-A1.2-01D	Data is accessible through an access protocol that supports authentication and authorisation
RDA-I2-01D	Data uses FAIR-compliant vocabularies
RDA-I3-01D	Data includes references to other data
RDA-I3-02M	Metadata includes references to other data
RDA-I3-02D	Data includes qualified references to other data
RDA-I3-04M	Metadata include qualified references to other data
RDA-R1.2-02M	Metadata includes provenance information according to a cross-community language

CC-BY 4.0 Creative Commons

6 Evaluation methods

6.1 Measuring progress

This approach is focused on the extent to which a resource under evaluation meets the requirement of the indicator, in order to answer the question:

> "How can the FAIRness of this data be improved?"

The indicator maturity levels 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 FAIRness progress per indicator is an evaluation of each indicator against these five levels of compliance. It gives the possibility to 'discard an indicator' as it might not be relevant for a particular community. The rationale of this approach is to give credit for evolution and help people to improve.

This approach may be most useful for data providers and publishers who want to do a self-assessment test to get a better idea on where to concentrate efforts to make their resources more FAIR.

A visualisation of the results of the evaluation of the indicators for all the FAIR areas could be presented as follows:

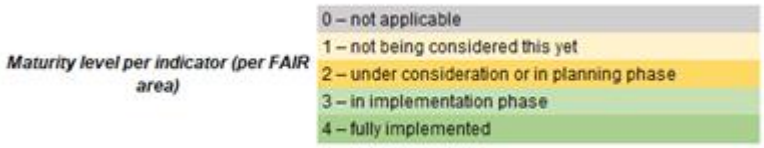

Figure 2 Five maturity levels

CC-BY 4.0 Creative Commons

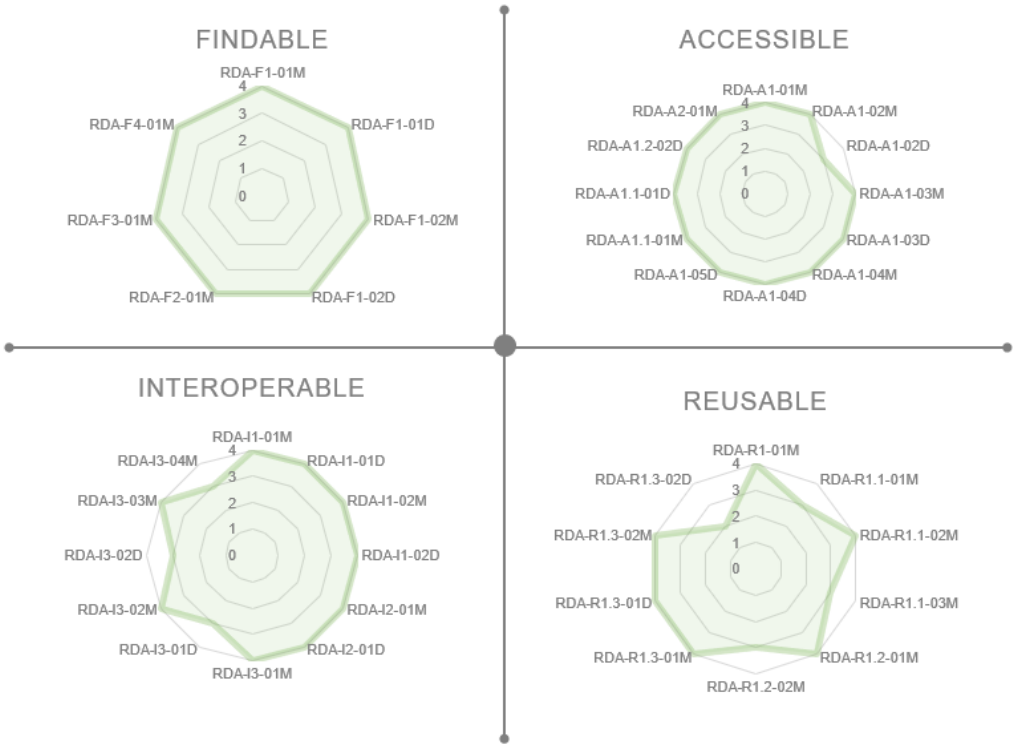

Figure 3 'Measuring progress' visualisation

The data provider or publisher could see from this visualisation that the resources are findable, and that the other areas are well developed with a small number of indicators requiring some more work.

6.2 Measuring pass-or-fail

This approach is focused on determining how a resource under evaluation performs on meeting the indicators across the FAIR areas. In that sense, it is a stricter evaluation as it gives a binary answer on each of the indicators, in effect only counting the indicators that reach top level 4 in the approach that measures progress in the previous section.

Here is an example to illustrate:

RDA-F4-01M Metadata is offered in such a way that it can be harvested and indexed.

- Metadata cannot be harvested and indexed **> FAIL**
- Metadata can be harvested and indexed **> PASS**

In addition to measuring the passing or failing on individual indicators, this approach measures the FAIRness per area by taking into account the priorities. It is measured per indicator – and aggregated per

CC-BY 4.0 Creative Commons

FAIR area. The level per area is determined based on the compliance to priorities. This is used to provide a 'measure of FAIRness'.

	Essential	Important	Useful
Level 0	○		
Level 1	●		
Level 2	●	◐	
Level 3	●	●	
Level 4	●	●	◐
Level 5	●	●	●

○ None of the indicators are satisfied
◐ Half of the indicators are satisfied
● All indicators are satisfied

Figure 4 Five compliance levels

This method may be most useful for external evaluators, for example community managers or funding agencies that want to verify that the resources that they manage or fund comply with a pre-defined level of FAIRness.

A visualisation of the results of this evaluation method could be as follows:

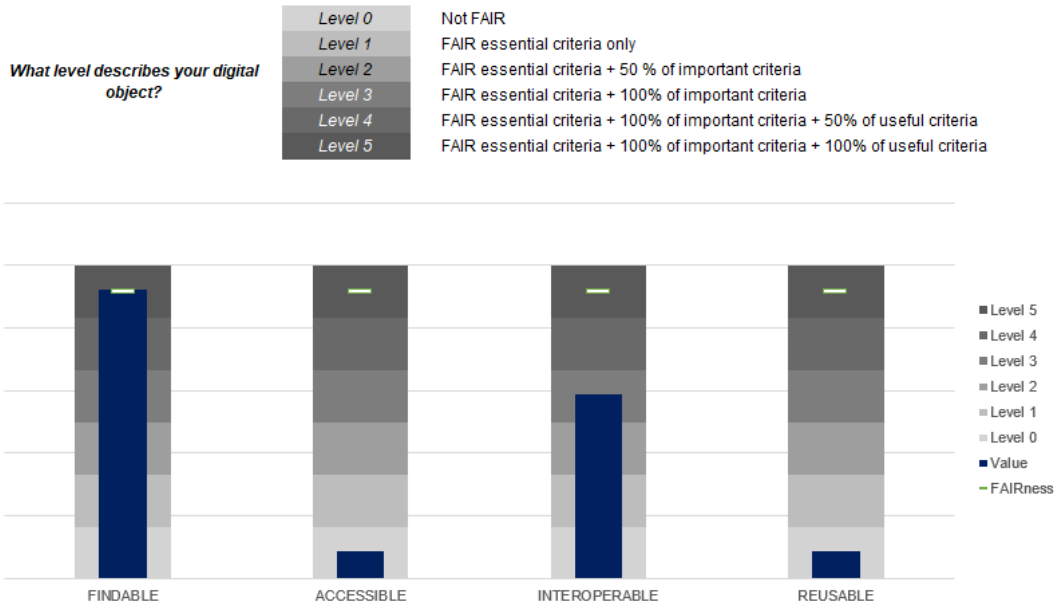

Figure 5 'Measuring pass or fail' visualisation

CC-BY 4.0 Creative Commons

In this visualisation, the evaluator can compare the level of FAIRness of a resource to an expected level for the community or the funding programme. In this example, it shows that the evaluated resource does not reach level 1, the minimum level of FAIRness for Accessible and Reusable.

The results shown on this visualisation relates to the same input provided for the “Measuring progress” method, as illustrated on Figure 3.

The data provider or publisher could see that from this graph, the digital object evaluated has a (1) **level 5** for Findability and (2) a **level 3** for Interoperability, which translate to (1) having all Findable indicators satisfied whereas (2) only the essential Interoperable indicators are satisfied. In other words, despite having almost all Interoperable criteria satisfied, as seen on Figure 3, a minority of the useful criteria are satisfied, which justifies a **level 3.** As for Accessibility and Interoperability, not all the essential criteria are satisfied which justifies a **level 0**.

6.3 Combined approach

The two approaches outlined in the sections above can be combined to bring together the benefits of both approaches.

The evaluation could first start with measuring the progress per indicator, leading to a set of radar charts.

The second step would then summarise all the indicators for which the highest level, i.e. when the requirement for an indicator has been fully implemented, is reached to give the data for the pass-or-fail results.

A FAIR evaluation tool is available which allows anyone to evaluate a resource and get results for both approaches. The tool can be downloaded here.

CC-BY 4.0 Creative Commons

7 Future maintenance

The FAIR Maturity Model described in this document is intended as a first version of the model. It is conceivable that, in the future, evolving experience with assessment methods for the FAIR principles and, possibly, further evolution of the FAIR principles may require changes to the model.

This document aims at informing a variety of assessment approaches which will in turn impact the perception one has of a specific digital object and potentially impact the reputation of the organisation who owns or curates it. The model should be further developed taking into account comments and contributions of a wide range of stakeholders and potentially extend the reliance on the references used as baseline to build the model (section 1.1). All the valuable contributions to FAIR should be considered as dependencies for the future confidence in this document. Referring to the preceding point, maintaining the model should be operated in all transparency.

The maintenance of the model will be taken up by a RDA Maintenance Working Group, a group that is defined by RDA[28] as having the purpose of managing the maintenance activities and supporting the adopters of the original Recommendation. This group will be tasked to gather further evidence of usage of the model and interact with research communities to incorporate further requirements towards future versions of the model. One of the principal tasks of the RDA Maintenance Working Group will be to solicit for implementation use cases from across different communities, monitor the proper implementation of the FAIR data maturity model and draw conclusions – which are expected to differ from one community to another.

Beyond the adoption of the model, the Maintenance Working Group might consider looking in more details at the indicators for Interoperable. These indicators present a challenge, notably due to some terms of which the Working Group was not able to fully grasp the meaning (e.g. knowledge representation).

Another piece of work would be to liaise with FAIRsharing to better align the requirements of assessment with the content of their registry (e.g. maintenance of standards).

In addition, Reusability may also need to be further discussed. For instance, whether non-standard licences (e.g. local licences) should be supported over standard licences and in what situation.

Furthermore, the Maintenance Working Group could consider looking at potential technologies and services to support the FAIR data maturity model future development and outreach.

All information related to the establishment of an RDA Maintenance Working Group and the maintenance activities within that context will be available on the RDA page of the Working Group.

[28] RDA. Creating or Joining an RDA Working Group. Section “Closing out a Working Group”, option b). https://www.rd-alliance.org/groups/creating-and-managing-rda-groups/creating-or-joining-rda-working-group.html

CC-BY 4.0 Creative Commons

8 Glossary

Table 6 Glossary

Term	Definition	Source	Related term	Used in
Access conditions	The exact conditions under which a Resource is accessible, expressed as requirements that a machine can understand to either automatically execute the requirements or alert the user to the requirements, such as the requirement to create a user account.	GO-FAIR		RDA-A1-01M
Access control	A way of limiting access to a system or to physical or virtual resources.	Techopedia	(related) Authorization	RDA-A1-01M
Authentication	The act of proving an assertion, such as the identity of a computer system user.	Wikipedia, also Techterms, Merriam-Webster, Cambridge Dictionary		RDA-A1.2-02D
Authorisation	The function of specifying access rights/privileges to resources, which is related to information security and computer security in general and to access control in particular.	Wikipedia, also Merriam-Webster, Cambridge Dictionary	(related) Access control	RDA-A1.2-02D
Automated (or automatic) access to data	Access using programmatic means, e.g. an API or SPARQL endpoint			RDA-A1-05D
Community standard, community-specific standard, domain/discipline standard	A standard for metadata or data that is recognised and widely used within a community.			RDA-F2-01M, RDA-R1-01M, RDA-R1.2-01M, RDA-R1.3-01M, RDA-R1.3-01D, RDA-R1.3-02M, RDA-R1.3-02D
Cross-domain language (for provenance)	An expression of provenance that can be used to represent and interchange provenance information generated in different systems and under different contexts, e.g. W3C PROV Ontology.		(broader) Provenance information	RDA-R1.2-02M
Data, digital object	A resource consisting of units of information, such as observations or measurements, being the primary subject of the FAIR			Throughout

CC-BY 4.0 Creative Commons

Term	Definition	Source	Related term	Used in
	evaluation.			
Discovery	The act of encountering something, often involving searching or navigating in a database.			RDA-F2-01M
Essential	Adjective used for an indicator that must be met in all cases for an object to be considered FAIR. This may be domain- or context-dependent.		Mandatory	
FAIR-compliant vocabulary	A vocabulary that is documented, resolvable and machine-understandable using globally unique and persistent identifiers.	GO-FAIR		RDA-I2-01M
Free access protocol	A protocol that can be used free of cost.	GO-FAIR	(broader) Protocol	RDA-A1.1-01M, RDA-A1.1-01D
Important	Adjective used for an indicator that should be met in most cases, unless there are circumstances where an object could still be FAIR if the requirement is not met, for example if the information is not applicable. This may be domain- or context-dependent.		Recommended	
Knowledge representation	Either (a) a set of concepts in a controlled vocabulary, an ontology or thesaurus, or (b) a data model, i.e. a well-defined framework to describe and structure metadata or data.	GO-FAIR	(narrower) Machine-understandable knowledge representation, Self-describing knowledge representation (related) Metadata, Vocabulary	All indicators under I1
Licence (re-use)	A legal document that specifies what a user can do with a resource.		(narrower) Standard licence, Machine-understandable licence	RDA-R1.1-01M, RDA-R1.1-02M, RDA-R1.1-03M
Machine-understandable knowledge representation	A knowledge representation that is expressed in such a way that a machine can take a decision on further actions.		(broader) Knowledge representation	RDA-I1-02M, RDA-I1-02D
Machine-understandable licence	A licence that is expressed in such a way that a machine can take a decision on further actions.		(broader) Licence	
Mandatory	Adjective used for an indicator that	RFC2119: MUST	Essential	

CC-BY 4.0 Creative Commons

Term	Definition	Source	Related term	Used in
	must be met. Used in earlier versions of the indicators. Now replaced by 'Essential'.			
Metadata	Information describing the characteristics of a data object including, for example, structural information describing data structures (e.g., data format, syntax, and semantics) and descriptive information describing data contents (e.g., information security labels).	NIST	(narrower) Metadata record, Metadata element, Metadata statement	
Metadata record	A set of metadata statements that is identified as a group.		(broader) Metadata (related) Metadata statement	
Metadata element	A property that is used to describe a characteristic, for example Dublin Core `dc:title`, Schema.org `schema:name` or DataCite Title.		(broader) Metadata	
Metadata statement	A single characteristic of a resource, usually expressed as a property-value pair, e.g. Title="My data".		(broader) Metadata	
Ontology	A formalised set of concepts relevant to a particular area of interest, representing rich and complex knowledge about things, groups of things, and relations between things, as well as a set of constraints about the usage of its terms.	W3C	(related) Vocabulary	
Open source access protocol	A protocol that is licensed under an Open Source licence, which implies it can be freely used, modified, and shared.	GO-FAIR	(broader) Protocol	
Optional	Adjective used for an indicator that may be evaluated at the discretion of the evaluator. Used in earlier versions of the indicators. Now redefined as 'Useful'	RFC2119: MAY	Useful	
Persistent identifier	An identifier that is guaranteed to remain valid over time. This guarantee requires an institutional commitment on the part of the publisher or maintainer of the identifier, and may include a guarantee that the identifier will		(related) Universally unique identifier	

CC-BY 4.0 Creative Commons

Term	Definition	Source	Related term	Used in
	continue to resolve to the same resource for a specific period of time.			
Protocol	A method by which a user or machine can gain access to a resource, for example HTTP(S) or FTP for access to resources on the Internet, or a telephone number and instructions to call the person or organisation that owns or manages the resource.	GO-FAIR	(narrower) Free access protocol, Open source access protocol	
Provenance information	Information about the origin and history of a resource. May include a description of the workflow that led to the resource, who generated or collected it and how it was processed.	GO-FAIR	(narrower) Cross-domain language (for provenance)	
Qualified reference	A reference that specifies its relationship to the resource that contains the reference.		(broader) Reference	
Recommended	Adjective used for an indicator that must be evaluated but does not have to be met necessarily, i.e. there could be certain circumstances that not meeting the requirement does not constitute a failure. Used in earlier versions of the indicators. Now replaced by 'Important'.	RFC2119: SHOULD	Important	
Reference	A link from one resource to another, for example a citation, or a URL, that allows a user to find or access the other resource.		(narrower) Qualified reference	
Resolution, resolving	The process in which an identifier is the input — a request — to a network service to receive in return a specific output of one or more pieces of current information (state data) related to the identified entity: e.g., a location (URL).	DOI Handbook		
Resource	Anything that is accessed and/or reused and of which the FAIRness is being assessed, including metadata and datasets.		(narrower) Metadata Data, digital object	
Reuse	The act of using an existing resource for a different purpose or in a different context. This may involve republishing and creating derivatives as far as allowed under			

CC-BY 4.0 Creative Commons

Term	Definition	Source	Related term	Used in
	the licence specified for reuse of the resource.			
Self-describing knowledge representation	A knowledge representation that comes with information that describes itself, e.g. with metadata about the knowledge representation such as a SKOS Concept Scheme.		(broader) Knowledge representation	
Standard	An agreed way of doing something, a norm. A standard provides the requirements, specifications, guidelines or characteristics that can be used for the description, interoperability, citation, sharing, publication, or preservation of all kinds of digital objects such as data, code, algorithms, workflows, software, or papers.	FAIRsharing FAQ		
Standard licence	A licence that is defined in a published and recognised specification.		(broader) Licence	
Standard vocabulary	A vocabulary that is defined in a standard. See for example the FAIRsharing standards list.		(broader) Vocabulary	
Standardised protocol	A protocol that is defined in a published and recognised specification.		(broader) Protocol	
Taxonomy	Orderly classification of things according to their relationships.	Wikipedia	(broader) Vocabulary	
Thesaurus	A list of subject headings or descriptors usually with a cross-reference system for use in the organization of a collection of documents for reference and retrieval		(broader) Vocabulary	
Universally unique identifier	An identifier that is guaranteed to uniquely identify a particular resource, irrespective of the context, i.e. making it impossible for the same identifier to refer to different resources. Not to be confused with UUID/RFC4122.		(related) Persistent identifier	
Useful	Adjective used for an indicator that may be met in some cases, increasing the FAIRness of a resource. This may be domain- or context-dependent.		Optional	

CC-BY 4.0 Creative Commons

Term	Definition	Source	Related term	Used in
Vocabulary, controlled vocabulary	A set of terms or concepts that can be used in the description of a resource. This includes taxonomies, ontologies and thesauri.	Adapted from the ANDS Vocabularies and research data Guide, section "What is a vocabulary?"	(narrower) Standard vocabulary (related) Ontology Taxonomy Thesaurus (related) Knowledge representation	

CC-BY 4.0 Creative Commons

9 Working Group materials

Table 7 Working group presentations and meeting reports

Date	Title	File
21/02/2019	Workshop #1 presentation	[Link to the document]
01/03/2019	Workshop #1 meeting report	[Link to the document]
03/04/2019	Workshop #2 presentation	[Link to the document]
06/05/2019	Workshop #2 meeting report	[Link to the document]
18/06/2019	Workshop #3 presentation	[Link to the document]
24/06/2019	Workshop #3 meeting report	[Link to the document]
12/09/2019	Workshop #4 presentation	[Link to the document]
19/09/2019	Workshop #4 meeting report	[Link to the document]
23/10/2019	Workshop #5 presentation	[Link to the document]
14/11/2019	Workshop #5 meeting report	[Link to the document]
04/12/2019	Workshop #6 presentation	[Link to the document]
11/12/2019	Workshop #6 meeting report	[Link to the document]
13/02/2020	Workshop #7 presentation	[Link to the document]
20/02/2020	Workshop #7 meeting report	[Link to the document]
24/03/2020	Workshop #8 presentation	[Link to the document]
03/04/2020	Workshop #8 meeting report	[Link to the document]
21/05/2020	Workshop #9 presentation	[Link to the document]
03/06/2020	Workshop #9 meeting report	[Link to the document]

CC-BY 4.0 Creative Commons