1st Conference on Research Data Infrastructure

Harmonising RDM

https://doi.org/10.52825/CoRDI.v1i.401

Published: 07 Sept. 2023

Which FAIR Are You?

A Detailed Comparison of Existing FAIR Metrics in the Context of Research Data Management

Mario Moser[1] [https://orcid.org/0000-0001-9325-4074],
Jonas Werheid[1] [https://orcid.org/0009-0003-6022-2633],
Tobias Hamann[1] [https://orcid.org/0000-0002-8021-5524],
Anas Abdelrazeq[1] [https://orcid.org/0000-0002-8450-2889], and
Robert H. Schmitt[1 2] [https://orcid.org/0000-0002-0011-5962]

[1]Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Germany
[2]Fraunhofer Institute for Production Technology (IPT), Aachen, Germany

Abstract: In data management the high-level FAIR principles are interpreted and implemented in various FAIR metrics. While this specific interpretation is intended, it leads to the situation of several metrics with different evaluation results for the same digital object. This work conducts an organizational-formal comparison, showing up elements like categories of importance in the considered metrics, as well as a content-wise comparison of selected metrics how they differ in their interpretation. The results give orientation especially to everyone in science aiming to find the right metric to make their data FAIR.

Keywords: Research Data Management - RDM - FDM; FAIR principles; FAIR metrics - comparison

1 Introduction

The FAIR principles [1] have been widely adopted as a guideline for making scientific data and scholarly digital objects more findable, accessible, interoperable, and reusable. Scholarly digital objects include data in a narrow sense as well as software, repositories, and workflows among others [1], [2]. The four foundational principles are described in more detail in fifteen guiding principles. These principles are interpreted and implemented in several FAIR metrics in order to meet domain-specific requirements or to focus on special types of digital objects [3]. This poses challenges for several addressees in research: scientists might want to select a metric to design their digital objects FAIR; initiatives working on developing new (specific) metrics need to understand existing ones and identify gaps; and research funding agencies might want to select an appropriate metric to measure the FAIRness of published results. Therefore, a detailed comparison of FAIR metrics is necessary to investigate differences and how they affect the data management and evaluation results for digital objects.

Various FAIR metrics are used in FAIR assessment tools [4]. While comparisons of FAIR assessment tools and their outcomes already exist (e. g. [5], [6], [7], [8], [9], [10]), they usually do not focus on the underlying metrics themselves as the reasons for differences observed. At least the tool evaluation in [11] includes comparison of the metrics Maturity Indicators (MI) and FAIRsFAIR (FsF). The tool comparison in [12] mentions the used metrics as one aspect. The development of the European Open Science Cloud (EOSC) FAIR metrics [13] made comparisons to metrics developed previously, but no complete overview comparing several metrics is currently available.

In our comparison, metrics that satisfies the following conditions are selected: Metrics that

- are generally applicable, i. e. that are not limited to one specific type of digital objects, since this reduces comparability;
- follow the structure of the fifteen FAIR guiding principles; and
- have a separate documentation following the characteristics in [14]'s table 1, such as a metric identifier and name as well as what and how is it measured.

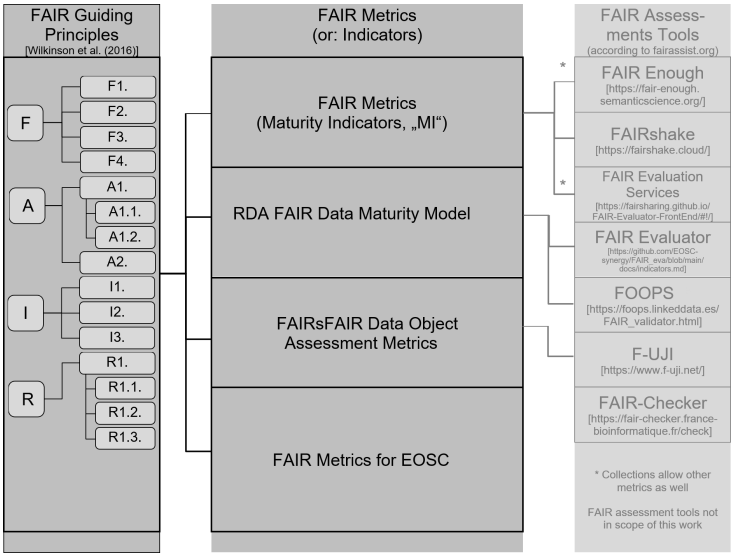
Metrics are identified based on scientific literature research as well as on considering metrics used in publicly available FAIR assessment tools. (Although the assessment tools themselves are mentioned in this paper, they are not compared.) For the identified metrics the comparison divides into two types: 1) An organizational-formal comparison focuses on characteristics like the evaluated digital objects, evaluation scale and metric item weighting factors. 2) Starting from each guiding principle a one-by-one content-wise comparison of each FAIR metric is conducted among the other metrics as well as with the (general) guiding principles itself.

2 Results

Four general FAIR metrics are identified for this work, presented in the middle column in figure 1, including their optional usage in assessment tools. As stated before, specific metrics are excluded (e. g. for FAIRness of software: [15], FAIR4RS ([16], [17]), [18]). The FAIRplus (F+) indicators [19] as well as several self-assessment checklists have been not been incorporated due to a missing structural similarity to the FAIR guiding principles.

The organizational-formal comparison reveals that metrics focus on digital objects in general (MI, RDA, FsF) or more specific on the web (EOSC). This outcome is not surprising since the concept of FAIR metrics is to specify the general FAIR principles for a certain usage. While e. g. the FAIR MIs are generally applicable, there is another metric rephrased to focus explicitly on software. The assessment tools "FOOPS" uses the (general) RDA metric for the evaluation of ontologies. RDA prioritizes the relevance of their metric items in three categories ("useful", "important" and "essential"). Following the design framework template by [14] reveals room for improvement in documentation.

The content-wise comparison of FAIR metric elements shows different numbers of metric items per metric: While the MIs are with 14 items quite close to the FAIR guiding principles, on the opposite site the RDA contains 41 items in their metric. The different number of metric items is mainly caused by splitting into multiple and more fine-granular metric items for one guiding principle: Data vs. metadata are distinguished; principles are divided into multiple parts (e. g. RDA-I1-01: "uses knowledge representation expressed in standardised format" and in addition "uses machine-understandable knowledge representation"); principles (e. g. A1.) with sub-principles (e. g. A1.1., A1.2.)



MI [20] [21], RDA [22] [23], FAIRsFAIR [24], EOSC [13]

Figure 1. FAIR metrics identified by usage in FAIR assessment tools

are either described on all levels or on subprinciple level only; human as well as machine readability is in some metrics explicitly distinguished. The EOSC that has marked some items "temporarily", which might evolve over time. The development of the EOSC FAIR metric has been influenced by the RDA's and FAIRsFAIR's metrics, what is expressed in similarity of the content. E. g. RDA-F1-01D demands that "Data is identified by a persistent identifier" [22] while FsF-F1-02D in a similar way states that "Data is assigned a persistent identifier" [24]. Therefore the RDA FAIR Data Maturity Model has an extended (direct/indirect) influence on the development of the presented subsequent metrics by FAIRsFAIR and EOSC. A high-level overview of the above-mentioned results is shown in figure 2.

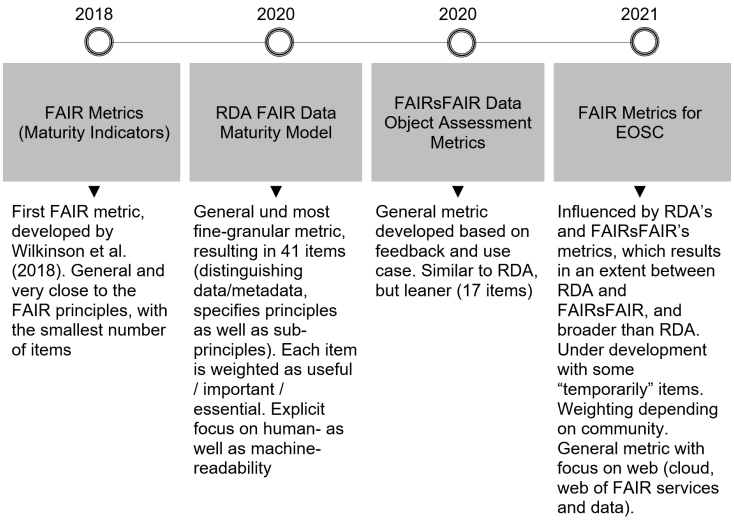


Figure 2. Simplified FAIR metrics overview with the results from the comparison

3 Outlook

Our detailed comparison of FAIR metrics reveals the differences on an organizational-formal level as well as content-wise. Differences are thematised and thus show the (intended) different interpretation and implementations of the FAIR principles.

The outcome can be one part of the explanation why different FAIR assessment tools lead to different results for the same digital object. For scientists, initiatives working on FAIR, and research funders, the results might be useful to decide which FAIR metric to use in their work. When developing a new metric, the comparison can help to identify gaps or to check typical design elements of metrics (like weighting categories, aspects to (not) focus on etc.).

It has to be mentioned that this analysis represents the current situation at the moment of execution. Due to the continuous evolution of FAIR, an ongoing evaluation of differences will be required. Existing metrics might change, and new metrics might be developed for specific purposes. This is necessary to adapt to new future developments and requirements. The selection of metrics in this work has been rather strict, so more metrics exist that are not covered here. But more important than *which* FAIR you use is *that* your data is somehow FAIR.

Author contributions

Mario Moser: Conceptualization, Investigation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing;
Jonas Werheid: Investigation, Formal analysis, Writing – original draft;
Tobias Hamann: Supervision, Writing – review & editing;
Anas Abdelrazeq: Supervision, Writing – review & editing;
Robert Schmitt: Funding acquisition, Supervision

Competing interests

The authors declare that they have no competing interests.

Funding

The authors would like to thank the Federal Government and the Heads of Government of the Länder, as well as the Joint Science Conference (GWK), for their funding and support within the framework of the NFDI4Ing consortium. Funded by the German Research Foundation (DFG) - project number 442146713.

Acknowledgements

The authors would like to thank the reviewers for their valuable and constructive feedback.

References

[1] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, *et al.*, "The FAIR Guiding Principles for scientific data management and stewardship," *Sci Data*, vol. 3, no. 160018, Mar. 2016. DOI: 10.1038/sdata.2016.18.

[2] S. Bechhofer, D. De Roure, M. Gamble, C. Goble, and I. Buchan, "Research Objects: Towards Exchange and Reuse of Digital Knowledge," *Nature Precedings*, Jul. 2010. DOI: 10.1038/npre.2010.4626.1.

[3] A. Jacobsen, R. de Miranda Azevedo, N. Juty, *et al.*, "FAIR Principles: Interpretations and Implementation Considerations," *Data Intelligence*, vol. 2, no. 1-2, pp. 10–29, Jan. 2020. DOI: 10.1162/dint_r_00024.

[4] A. Devaraju and R. Huber, "An automated solution for measuring the progress toward FAIR research data," *Patterns*, vol. 2, no. 11, Nov. 2021. DOI: 10.1016/j.patter.2021.100370.

[5] M. D. Wilkinson, S.-A. Sansone, G. Marjan, J. Nordling, R. Dennis, and D. Hecker, "FAIR Assessment Tools: Towards an "Apples to Apples" Comparisons," *Zenodo*, Dec. 2022. DOI: 10.5281/zenodo.7463421.

[6] K. Peters-von Gehlen, H. Höck, A. Fast, D. Heydebreck, A. Lammert, and H. Thiemann, "Recommendations for Discipline-Specific FAIRness Evaluation Derived from Applying an Ensemble of Evaluation Tools," *Data Science Journal*, vol. 21, no. 7, Mar. 2022. DOI: 10.5334/dsj-2022-007.

[7] N. Krans, A. Ammar, P. Nymark, E. Willighagen, M. Bakker, and J. Quik, "FAIR assessment tools: evaluating use and performance," *NanoImpact*, vol. 37, p. 100 402, 2033. DOI: 10.1016/j.impact.2022.100402.

[8] C. Bahim, M. Dekkers, and B. Wyns, "Results of an Analysis of Existing FAIR Assessment Tools," *Zenodo*, May 2019. DOI: 10.15497/rda00035.

[9] E. González, A. Benítez, and D. Garijo, "FAIROs: Towards FAIR Assessment in Research Objects," in *Linking Theory and Practice of Digital Libraries*, G. Silvello, O. Corcho, P. Manghi, *et al.*, Eds., Cham: Springer International Publishing, 2022, pp. 68–80. DOI: 10.1007/978-3-031-16802-4_6.

[10] D. Slamkov, V. Stojanov, B. Koteska, and A. Mishev, "A Comparison of Data FAIRness Evaluation Tools," in *Ninth Workshop on Software Quality Analysis, Monitoring, Improvement, and Applications (SQAMIA 2022)*, Oct. 2022. [Online]. Available: https://www.researchgate.net/publication/364308377.

[11] C. Sun, V. Emonet, and M. Dumontier, "A comprehensive comparison of automated FAIRness Evaluation Tools," *Semantic Web Applications and Tools for Healthcare and Life Sciences (SWAT4HCLS)*, Dec. 2022, Additional material at https://doi.org/10.5281/zenodo.5539823. [Online]. Available: https://ceur-ws.org/Vol-3127/paper-6.pdf.

[12] E. Kontsioti. "The Road to FAIRness: An Evaluation of FAIR Data Assessment Tools." (2023), [Online]. Available: https://www.thehyve.nl/articles/evaluation-fair-data-assessment-tools (visited on 04/26/2023).

[13] European Commission and Directorate-General for Research and Innovation, J. M. Aronsen, O. Beyan, *et al.*, *Recommendations on FAIR metrics for EOSC*, S. Jones and F. Genova, Eds. Publications Office, 2021. DOI: 10.2777/70791. [Online]. Available: https://op.europa.eu/o/opportal-service/download-handler?identifier=ced147c9-53c0-11eb-b59f-01aa75ed71a1&format=pdf&language=en&productionSystem=cellar&part=.

[14] M. D. Wilkinson, S.-A. Sansone, E. Schultes, P. Doorn, L. O. Bonino da Silva Santos, and M. Dumontier, "A design framework and exemplar metrics for FAIRness," *Scientific Data*, vol. 5, no. 1, Jun. 2018. DOI: 10.1038/sdata.2018.118.

[15] A.-L. Lamprecht, L. Garcia, M. Kuzak, *et al.*, "Towards FAIR principles for research software," *Data Science*, vol. 3, no. 1, pp. 37–59, Jun. 2020. DOI: 10.3233/DS-190026.

[16] N. P. Chue Hong, D. S. Katz, M. Barker, *et al.*, "FAIR Principles for Research Software (FAIR4RS Principles)," *Zenodo*, May 2022. DOI: 10.15497/RDA00068.

[17] M. Barker, N. P. Chue Hong, D. S. Katz, *et al.*, "Introducing the FAIR Principles for research software," *Scientific Data*, vol. 9, no. 622, pp–pp, Oct. 2022. DOI: 10.1038/s41597-022-01710-x.

[18] D. S. Katz, M. Gruenpeter, and T. Honeyman, "Taking a fresh look at FAIR for research software," *Patterns*, vol. 2, no. 1, Mar. 2021. DOI: 10.1016/j.patter.2021.100222.

[19] FAIRplus. "FAIRplus Indicators V0.1." (Oct. 2020), [Online]. Available: https://fairplus.github.io/fairification-results/2020-10-11-FAIRplus-indicators-v0.1/ (visited on 04/26/2023).

[20] M. D. Wilkinson, M. Dumontier, S.-A. Sansone, *et al.*, "Evaluating FAIR maturity through a scalable, automated, community-governed framework," *Scientific Data*, vol. 6, no. 174, Sep. 2019. DOI: https://doi.org/10.1038/s41597-019-0184-5.

[21] FAIR Metrics Group. "GitHub FAIRMetrics." (2022), [Online]. Available: https://github.com/FAIRMetrics/Metrics (visited on 04/26/2023).

[22] RDA FAIR Data Maturity Model Working Group, "FAIR Data Maturity Model: specification and guidelines," 2020. DOI: 10.15497/rda00050.

[23] C. Bahim, C. Casorrán-Amilburu, M. Dekkers, *et al.*, "The FAIR Data Maturity Model: An Approach to Harmonise FAIR Assessments," *Data Science Journal*, vol. 19, no. 1, pp. 1–7, Oct. 2020. DOI: 10.5334/dsj-2020-041.

[24] A. Devaraju, R. Huber, M. Mokrane, *et al.*, "FAIRsFAIR Data Object Assessment Metrics (0.5)," *Zenodo*, Apr. 2022. DOI: 10.5281/zenodo.6461229.