

Considerations for the Conduction and Interpretation of FAIRness Evaluations postprint

Authors: Azevedo, Ricardo de Miranda, Dumontier, Michel, Michel Dumontier, Azevedo, Ricardo de Miranda

Date: 2022-11-18T00:00:00+00:00

Abstract

The FAIR principles were received with broad acceptance in several scientific communities. However, there is still some degree of uncertainty on how they should be implemented. Several self-report questionnaires have been proposed to assess the implementation of the FAIR principles. Moreover, the FAIRmetrics group released 14, general-purpose maturity for representing FAIRness. Initially, these metrics were conducted as open-answer questionnaires. Recently, these metrics have been implemented into a software that can automatically harvest metadata from metadata providers and generate a principle-specific FAIRness evaluation. With so many different approaches for FAIRness evaluations, we believe that further clarification on their limitations and advantages, as well as on their interpretation and interplay should be considered.

Full Text

Preamble

Related to other papers in this special issue: 2 (p10); 9 (p87)

Addressing FAIR principles: F, A, I, R

Considerations for the Conduction and Interpretation of FAIRness Evaluations

Ricardo de Miranda Azevedo^{1†}, Mark D. Wilkinson² & Michel Dumontier¹

¹Institute of Data Science, Maastricht University, Universiteitssingel 60, Maastricht 6229 ER, The Netherlands

²Centro de Biotecnología y Genómica de Plantas UPM-INIA (CBGP), Autopista M-40 (Km 38) 28223-Pozuelo de Alarcón (Madrid), Comunidad de Madrid 28040, Spain

Keywords: FAIR metrics; FAIR principles; FAIR data; FAIRness; Implementation; Maturity indicators; Scientific community

Citation: R. de Miranda Azevedo, M. Wilkinson & M. Dumontier. Considerations for the conduction and interpretation of FAIRness evaluations. *Data Intelligence* 2(2020), 285–292. doi: 10.1162/dint_a_{00051}

ABSTRACT

The FAIR principles were received with broad acceptance in several scientific communities. However, there is still some degree of uncertainty on how they should be implemented. Several self-report questionnaires have been proposed to assess the implementation of the FAIR principles. Moreover, the FAIRmetrics group released 14 general-purpose maturity indicators for representing FAIRness. Initially, these metrics were conducted as open-answer questionnaires. Recently, these metrics have been implemented into software that can automatically harvest metadata from metadata providers and generate a principle-specific FAIRness evaluation. With so many different approaches for FAIRness evaluations, we believe that further clarification on their limitations and advantages, as well as on their interpretation and interplay, should be considered.

† Corresponding author: Ricardo de Miranda Azevedo (E-mail: ricardo.mrma@gmail.com, ORCID: 0000-0002-7641-6446).

1. INTRODUCTION

The FAIR (Findable, Accessible, Interoperable and Reusable) guiding principles for scientific data management were published in 2016 and quickly gained popularity among the scientific community, being cited more than 600 times in scholarly publications [1]. Moreover, the implementation of the FAIR principles has become a prerequisite for many funders [2,3], demonstrating how widely the principles have been adopted. Nonetheless, their implementation is still not fully understood by the research community [4].

The original FAIR publication does not exemplify the implementation of the principles. Instead, the authors suggested that scientific communities interested in enhancing data reuse should engage in defining which implementations are considered to be FAIR [5,6]. For this reason, the need to objectively evaluate the “FAIRness” of research resources is highly warranted [4].

Recently, several tools were developed for conducting FAIRness evaluations. These tools come in different forms, such as questionnaires [6-10], checklists [11,12], and semi-automated evaluators [13]. A detailed analysis of each of the existing FAIR assessment tools was recently published by the FAIR data maturity model workgroup of the Research Data Alliance (RDA) [14]. Moreover, the “FAIRsharing” group has released the “FAIRassist.org” website, which provides a summary of these different tools according to execution type, key features, organization, target objects, and whether they include reading materials [15].

Despite these recent efforts, we believe that further clarification can aid researchers in conducting and interpreting FAIRness evaluations. Therefore, the objectives of this manuscript are to (I) briefly describe the different types of FAIRness evaluations; (II) summarize the advantages and limitations of each type; and (III) provide considerations on how to use and interpret FAIRness evaluations.

2. TYPES OF FAIRNESS EVALUATION

In FAIRassist.org [15], four categories are used to describe how a certain tool is executed: manual - questionnaire; manual - checklist; semi-manual; and automated. For the sake of simplicity, we will collapse both “manual-questionnaire” and “manual-checklist” into a single category (“discrete-answer questionnaire-based evaluations”).

2.1 Discrete-answer questionnaire-based evaluations

This approach consists of online self-report questionnaires and checklists where respondents need to indicate their implementation choice (for a given principle) from a predefined set of answers. Questions aim to reflect each of the principles or some additional related concept that is not FAIR (e.g., data quality or openness). Questions are grouped according to each principle (i.e., F, A, I, R). The output of the evaluation is a sum score [8,16], a weighted sum score [12], or a visual score [7] that is automatically generated after completion.

Discrete-answer questionnaire-based tools may present several limitations: (I) the set of answers is not extensible, meaning that it is not possible to include additional standards; (II) items do not provide evidence of a FAIR implementation, but only intention; and (III) score computation is arbitrary. Nonetheless, these approaches have proven particularly useful as educational resources for respondents with little knowledge of FAIR. Moreover, these questionnaires could potentially be adapted for testing knowledge on data stewardship. Studies assessing the psychometric properties of these questionnaires are warranted.

2.2 Open-answer questionnaire-based evaluation

Open-answer questionnaires for evaluating FAIRness were first introduced by the FAIRmetrics focus group (a subset of the authors of the original FAIR paper) and assessed using a spreadsheet [6,17]. The authors published a design framework and exemplar metrics for representing the FAIRness of digital resources [6]. It consists of 14 “general-purpose” metrics reflecting the FAIR principles and their subparts (Wilkinson et al., 2018). Later on, the authors decided to rename the metrics to FAIR “maturity indicators,” as the evaluations are supposed to indicate specific points that can be improved rather than “measure” FAIRness.

Open-answer questionnaires for FAIRness evaluations have similar limitations to

discrete-answer category questionnaires: they are time-consuming and subject to respondent bias [13]. Another important limitation is the lack of instructional material to aid respondents in conducting the evaluations. Respondents need to dig into many different resources to attempt to provide answers. Moreover, the psychometric properties of these tools have also not been evaluated, which compromises their validity and reliability, and therefore any attempts to compute total scores should be regarded as experimental.

The advantages of this approach are: (I) answers need to include a statement that evidences the FAIR implementation (e.g., a URI of a metadata record of the adopted standard), which helps to ensure the FAIRness of the standards used by the resources; (II) it allows scientific communities to create additional maturity indicators; and (III) it can be filled in on multiple occasions (spreadsheet version).

2.3 Semi-automated evaluation

The “FAIR Maturity Evaluation Service” approach consists of a web-based metadata harvester [13,18] that parses metadata of a digital resource assigned to a global unique identifier (GUID). The respondent needs to enter: (I) a brief description of the resource to be evaluated; (II) the GUID of the resource; and (III) their ORCID [19]. Moreover, a “collection” of maturity indicators must be selected prior to conducting the evaluation [13]. A collection is basically a group of maturity indicator tests and should be compiled to address the needs of the resource or scientific community engaged in performing the evaluations. For instance, one could be only interested in “Findability” and “Accessibility” at the initial stage of a funded project.

Additionally, a short description of the resource being evaluated must be provided. Descriptions must include a minimum of 40 words, preferably including the title of the resource, the metadata provider, the collection used, a short description, and the ORCID code of the respondent.

The semi-automated approach has several advantages: (I) it is automated, which is supposed to reduce respondent bias; (II) it is open-source; and (III) it ensures transparency as the evaluations are open and the evaluators identifiable. However, the semi-automated approach also has some important limitations: (I) it requires that the resource has some kind of metadata provider available (and therefore is not useful for projects in the developmental phase); (II) it depends on the compatibility between software and metadata provider; and (III) it performs differently when comparing two different identifiers of the same resource. Nonetheless, potential users of the FAIR Maturity Evaluation Service are welcome to test and provide feedback to the development team.

3. CONSIDERATIONS FOR CONDUCTING AND INTERPRETING FAIRNESS EVALUATIONS

3.1 What does the outcome of a FAIRness evaluation tell us?

FAIRness can be interpreted as “a continuum of behaviors” exhibited by a resource that can lead to machine discoverability and (re)use [5]. Therefore, “absolute FAIRness” is conceptually not achievable (and will probably never be). Nonetheless, some tools for FAIRness evaluations attempt to provide overall scores.

However, besides not being a validated calculation, the FAIRmetrics group opposed the idea of computing total scores. Regardless of the methodology for computing overall sum scores, different digital objects would need different score computations. Therefore, overall scores should not be used for evaluative purposes.

In opposition to overall scores, the assessment of specific maturity indicators can provide a starting point to define implementation choices that are missing for that resource. Moreover, the output of each maturity indicator test also reflects whether the standard has been properly stated in the resource’s metadata or if it still needs to be implemented in the code of the FAIR maturity service. These outputs can be useful for data stewards internally appraising the resource they represent.

It should be taken into account that there are many possible reasons explaining a failed FAIR maturity test. The following questions should be raised: Does the metadata about the resource contain a statement that addresses the principle failing a maturity test? Does the metadata scheme cover that principle? Is that principle relevant for that resource type or scientific community to which it belongs? Most importantly, however, does the semi-automated evaluation software support that standard? The answers to these fundamental questions should aid evaluators in understanding why their resource failed a test and in further tailoring a strategy to address it.

3.2 What type of FAIRness evaluation should I use?

If the interest is solely in evaluating the FAIRness of the resource for internal appraisal, then the recommendation is to start with a semi-automated evaluation. However, it should be taken into account that the output of the evaluation might not reflect the needs of your scientific community. Although the existing FAIR maturity indicators are supposedly going to meet the needs of many communities, some reasoning is needed prior to using the semi-automated approach for official purposes.

Different approaches should be complementary rather than concurrent. The outcome of open-answer questionnaires can be useful to capture information on standards that still need to be implemented in the automated maturity tests. Moreover, as demonstrated in the FAIR principles explained and the GO-FAIR

matrix, questionnaires can be used for gathering valuable data on what the implementation choices are for that given scientific community. The existing maturity indicators shall work for many researchers; however, if scientific communities want to go further and develop their own maturity tests, then it should be their role to do so.

3.3 Resource scored positive using questionnaire, but negative on the automated evaluation; what should I do?

Researchers are genuinely more interested in assessing their own digital objects rather than those from others. However, it is important that the community ensures that the standards used are also FAIR. Therefore, members of scientific communities should consider identifying whether the standards adopted by their community (i.e., identifiers, formats, protocols, policies, etc.) comply with the FAIR principles. Metadata about the standards is crucial for ensuring their FAIRness.

FAIRsharing (www.fairsharing.org) [20], Identifiers.org (www.identifiers.org) [21], and Biportal (www.biportal.org) [22] are examples of metadata registries for metadata standards, including file formats, ontologies, identifier schemas, as well as for maturity indicators. Therefore, communities should ensure that the standards they select as best practices for FAIRness also contain proper metadata.

[Figure 1: see original paper]. Workflow of the interplay between manual and automated assessments.

4. CONCLUSION

The present manuscript provided a short description of different types of FAIRness evaluations, pointing to their advantages and limitations. Moreover, reflections on three fundamental points were provided: (I) evaluations should be assessed at the maturity indicator level rather than at the overall level of FAIRness; (II) the open-answer questionnaire approach should be used complementary to the automated approach rather than concurrently, and data stewards should figure out what leads to inconsistencies between both approaches; and (III) communities should take the role of implementing standards in the semi-automated evaluator. Standards must also be FAIR (and therefore have an identifier, metadata descriptors, etc.).

ACKNOWLEDGEMENTS

M. Dumontier was supported by grants from NWO (400.17.605; 628.011.011), NIH (3OT3TR002027-01S1; 1OT3OD025467-01; 1OT3OD025464-01), H2020-EU EOSLife (824087), and ELIXIR, the research infrastructure for life-science data. R. de Miranda Azevedo was supported by grants from H2020-EU

EOSClife (824087), and ELIXIR, the research infrastructure for life-science data.

REFERENCES

- [1] M.D. Wilkinson, M. Dumontier, I. Jan Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg ...& B. Mons. The FAIR Guiding Principles for scientific data management and stewardship. *Nature* 3(2016), 160018. doi:10.1038/sdata.2016.18.
- [2] D. Spichtinger & J. Siren. The Development of Research Data Management Policies in Horizon 2020. *Research Data Management -A European Perspective*. 2017. doi: 10.1515/9783110365634-002.
- [3] FAIR principles & FAIRmetrics -ZonMw. Available: <https://www.zonmw.nl/en/research-and-results/fair-data-data-management/data-management-at-zonmw/fair-principles-fairmetrics/>.
- [4] T. Hill. Turning FAIR into reality: Review. *Learned Publishing* 3(32) (2019), 283-286. doi: 10.1002/leap.1234.
- [5] B. Mons, C. Neylon, J. Velterop, M. Dumontier, L.O. Bonino da Silva Santos & M.D. Wilkinson. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. *Information Services & Use* 37(1) (2017), 49-56. doi: 10.3233/ISU-170824.
- [6] M.D. Wilkinson, S.A. Sansone, E. Schultes, P. Doorn, L.O. Bonino da Silva Santos & M. Dumontier. A design framework and exemplar metrics for FAIRness. *Scientific Data* 5(2018), 180118. doi: 10.1038/sdata.2018.118.
- [7] 5 Star Data Rating tool -OzNome. Available at: <https://research.csiro.au/oznome/tools/oznome-5-star-data/>.
- [8] FAIR data self-assessment tool. Available at: <https://www.andis.org.au/working-with-data/fairdata/fair-data-self-assessment-tool>.
- [9] FAIR data assessment tool-Data Impact Blog. Available at: <http://blog.ukdataservice.ac.uk/fair-data-assessment-tool/>.
- [10] J. Daniel, B. Clarke, L.L. Wang, A. Jones, P.L. Wojciechowicz, D. Torre, K.M. Jagodnik ...& A. Ma' ayan. FAIRshake: Toolkit to evaluate the findability, accessibility, interoperability, and reusability of research digital resources. *bioRxiv* 2019. Available at: <https://www.biorxiv.org/content/10.1101/657676v1>.
- [11] R. David. How to operationalize and to evaluate the FAIRness in the crediting and rewarding processes in data sharing: A first step towards a simplified assessment grid. Available at: <https://hal.archives-ouvertes.fr/hal-01943521/>.
- [12] GARDIAN FAIR Metrics. Available at: <https://gardian.bigdata.cgiar.org/metrics.php#!/>.
- [13] M.D Wilkinson, M. Dumontier, S.A. Sansone, L. O. Bonino da Silva Santos, M. Prieto, P. McQuilton ...& E. Schultes. Evaluating FAIR-compliance through

an objective, automated, community-governed framework. bioRxiv. 2018. Available at: <https://www.biorxiv.org/content/10.1101/418376v2.abstract>.

[14] Results of an analysis of existing FAIR assessment tools. Available at: <https://www.rd-alliance.org/group/fair-data-maturity-model-wg/outcomes/results-analysis-existing-fair-assessment-tools>.

[15] FAIRassist.org. Available at: <http://fairassist.org>.

[16] Checklist to evaluate FAIRness of data(sets). Available at: <https://docs.google.com/forms/d/e/1FAIpQLSF>

[17] Manual FAIRness assessment. Available at: https://docs.google.com/spreadsheets/d/1mJgYmiyLQdXrR_

[18] The FAIR maturity evaluation service. Available at: <https://w3id.org/AmIFAIR>.

[19] D. Butler. Scientists: Your number is up. Nature 485(7400) (2012), 564. doi:10.1038/485564a.

[20] S.A. Sansone, P. McQuilton & P. Rocca-Serra. FAIRsharing: Working with and for the community to describe and link data standards, repositories and policies. bioRxiv, 2018. Available at: <https://www.biorxiv.org/content/10.1101/245183v1.abstract>.

[21] N. Juty, C. Laibe & N. Le Novere. Identifiers.org/MIRIAM Registry annotation and cross-referencing framework. Nature Proceedings. 2011. doi: 10.1038/npre.2011.6396.1.

[22] P.L. Whetzel, N.F. Noy, N.H. Shah, P.R. Alexander, C. Nyulas, T. Tudorache & M.A. Musen. BioPortal: Enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. Nucleic Acids Research 39(2011), W541-5. doi: 10.1093/nar/gkr469.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv –Machine translation. Verify with original.