Towards an extensible FAIRness assessment of FAIR Digital Objects

Vincent Emonet‡, Remzi Çelebi‡, Jinzhou Yang‡, Michel Dumontier‡

‡ Institute of Data Science at Maastricht University, Maastricht, Netherlands

Corresponding author: Vincent Emonet (vincent.emonet@maastrichtuniversity.nl)

Abstract

The objective of the FAIR Digital Objects Framework (FDOF) is for objects published in a digital environment to comply with a set of requirements, such as identifiability, and the use of a rich metadata record (Santos 2021, Schultes and Wittenburg 2019, Schwardmann 2020). With the increasing prevalence of the FAIR (Findable, Accessible, Interoperable, Reusable) principles, and FAIR Digital Objects (FDO), used within different communities and domains (Wise et al. 2019), there will be a need to evaluate whether a FDO meets the requirements of the ecosystem in which it is used.

Without a dedicated framework, communities will develop isolated assessment systems from the ground up (Sun et al. 2022, Bahim et al. 2020), which will cost them time, and lead to FAIRness assessments with limited interoperability and comparability.

Previous work from the FAIR Metrics working group defined a framework for deploying individual FAIR metrics tests as separate services endpoints (Wilkinson et al. 2018, Wilkinson et al. 2019). To work in accordance with this framework, each test should take a subject URL as input, and return a score, either 0 or 1, a test version, and the test execution logs. A central service can then be used to assess the FAIRness of digital objects using collections of individual assessments. Such a framework could be easily extended, but there are currently no guidelines or tools to implement and publish new FAIRness assessments complying with this framework.

To amend this problem, we published the fair-test library in python and its documentation, which help with developing and deploying individual FAIRness assessments. With this library, developers define their metric tests using custom python objects, which will guide them to provide all required metadata for their test as attributes, and implement the test evaluation logic as a function. The library also provides additional helper functions for common tasks, such as retrieving metadata from a URL, or testing a metric test.

These tests can then be deployed as a web API, and registered in a central FAIR evaluation service supporting the FAIR metrics working group framework, such as FAIR

enough or the FAIR evaluator. Finally, users of the evaluation services will be able to group the registered metrics tests in collections used to assess the quality of publicly available digital objects.

There are currently as many as 47 tests that have been defined to assess compliance with various FAIR metrics, from which 25 have been defined using the fair-test library, including assessing if the identifier used is persistent, or if the metadata record attached to a digital object complies with a specific schema.

This presentation introduces a user-friendly and extensible tool, which can assess whether specific requirements are met for a digital resource. Our contributions are:

- Developing and publishing the fair-test library to make the development and deployment of independent FAIRness assessment tests easier.
- Developing and publishing tests in python for existing FAIR metrics: 23 generic tests covering most of the FAIR metrics, and 2 domain-specific tests for the Rare Disease research community.

We aim to engage with the FDO community to explore potential use-cases for an extensible tool to evaluate FDOs, and discuss their expectations related to the evaluation of digital objects.

Insights and guidelines from the FDO community would contribute to further improving the fair-test ecosystem. Among improvements that are currently being under consideration, we can cite improving the collaborative aspect of metadata extraction, or adding new metadata to be returned by the tests.

Keywords

FAIR evaluations, library, validation

Presenting author

Vincent Emonet

Presented at

First International Conference on FAIR Digital Objects, presentation

Conflicts of interest

References

- Bahim C, Casorrán-Amilburu C, Dekkers M, Herczog E, Loozen N, Repanas K, Russell K, Stall S (2020) The FAIR Data Maturity Model: An Approach to Harmonise FAIR Assessments. Data Science Journal 19 https://doi.org/10.5334/dsj-2020-041
- Santos LOBdS (2021) FAIR Digital Object Framework Documentation. https://fairdigitalobjectframework.org. Accessed on: 2022-7-08.
- Schultes E, Wittenburg P (2019) FAIR Principles and Digital Objects: Accelerating Convergence on a Data Infrastructure. Communications in Computer and Information Science3-16. https://doi.org/10.1007/978-3-030-23584-0_1
- Schwardmann U (2020) Digital Objects – FAIR Digital Objects: Which Services Are Required? Data Science Journal 19 https://doi.org/10.5334/dsj-2020-015
- Sun C, Emonet V, Dumontier M (2022) A Comprehensive Comparison of Automated FAIRness Evaluation Tools. SWAT4HCLS. 44-53 pp. URL: http://ceur-ws.org/Vol-3127/paper-6.pdf
- Wilkinson M, Sansone S, Schultes E, Doorn P, Bonino da Silva Santos LO, Dumontier M (2018) A design framework and exemplar metrics for FAIRness. Scientific Data 5 (1). https://doi.org/10.1038/sdata.2018.118
- Wilkinson M, Dumontier M, Sansone S, Bonino da Silva Santos LO, Prieto M, Batista D, McQuilton P, Kuhn T, Rocca-Serra P, Crosas M, Schultes E (2019) Evaluating FAIR maturity through a scalable, automated, community-governed framework. Scientific Data 6 (1). https://doi.org/10.1038/s41597-019-0184-5
- Wise J, de Barron AG, Splendiani A, Balali-Mood B, Vasant D, Little E, Mellino G, Harrow I, Smith I, Taubert J, van Bochove K, Romacker M, Walgemoed P, Jimenez R, Winnenburg R, Plasterer T, Gupta V, Hedley V (2019) Implementation and relevance of FAIR data principles in biopharmaceutical R&D. Drug Discovery Today 24 (4): 933-938. https://doi.org/10.1016/j.drudis.2019.01.008