The ecology of FAIR Digital Objects, with special attention to roundtripping and benchmarking across the research ecosystem

Daniel Mietchen ‡, §

‡ Ronin Institute of Independent Scholarship, Montclair, United States of America
§ Institute for Globally Distributed Open Research and Education (IGDORE), Jena, Germany

Corresponding author: Daniel Mietchen (daniel.mietchen@ronininstitute.org)

Abstract

The more findable, accessible, interoperable and reusable (i.e. aligned with the FAIR Principles outlined by Wilkinson et al. 2016) a Digital Object is, the more likely it is to interact with other entities in the research ecosystem and beyond. As long as the interoperability of these entities is not perfect (and it rarely is), a variety of interactions with a given Digital Object (e.g. split, merge, aggregation, transformation, backup, upload, download, or updates of content, metadata, storage or permissions) will mean a variety of representations of it, with some closer to the original than others. This has consequences for how the information about Digital Objects or contained in them can move around the research ecosystem. In some contexts, multiple representations of a given original (or aspects of it) might exist, creating the need to assess similarities, differences and relationships and to include them in curation, management, education, dissemination and preservation workflows. In other contetxts, the sole copy of a Digital Object might exist on a legacy system with limited alignment to the FAIR Principles, which creates the need for generating more readily accessible backup copies and to adapt some of them for inclusion in contemporary workflows.

In this presentation, we will look at the suitability of sets of FAIR Digital Objects to serve as indicators for several aspects of FAIRness across different elements of the research ecosystem. These sets could involve existing FAIR Digital Objects (e.g. data management plans, as per Mietchen 2021) as well as new or hypothetical ones, and inclusion or exclusion with respect to a given set could be defined using a wide range of criteria pertaining to the ecosystem elements of interest. Taking inspiration from tracing, monitoring, benchmarking and roundtripping activities in various research fields, we will then explore how far, how well and how quickly such sets - or their content - can travel through multiple elements of the research ecosystem (e.g. different databases or software pipelines or different stages of the research cycle) and what this means in terms of potential improvements to the FAIR Digital Objects themselves, to the sets and their

©
.

contents, to the way the FAIR assessments are implemented or to relevant elements of the research ecosystem.

Keywords

data use, data reuse, data reusability, metadata standards, ontology alignment, multilinguality

Presenting author

N/A

Presented at

First International Conference on FAIR Digital Objects, presentation

Conflicts of interest

References

- Mietchen D (2021) Now that machine actionable DMPs are becoming available, what could we do with them? Zenodo https://doi.org/10.5281/zenodo.5654899
- Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific data 3: 160018. https://doi.org/10.1038/sdata.2016.18