International FAIR Digital Objects Implementation Summit 2024
Extended Abstracts
https://doi.org/10.52825/ocp.v5i.1173

Published: 18 Mar. 2025

Combining FAIR Digital Object Implementation Concepts in a Real-Life Application

Benedikt Heinrichs[1,*], Sirieam Hunke[1,*], and Marius Politze[1,*]

[1]IT Center, RWTH Aachen University, Germany
*Correspondence: Benedikt Heinrichs, heinrichs@itc.rwth-aachen
Sirieam Hunke, hunke@itc.rwth-aachen.de
Marius Politze, politze@itc.rwth-aachen.de

Abstract. This work presents how to link different implementations of FAIR Digital Objects (FDOs) in a real-life application. First the real-life application and a use case is presented, afterward the different concepts are detailed, and finally the linking step is outlined.

Keywords: FAIR Digital Object, Research Data Management, Metadata, Persistent Identifier

1. The Real-Life Application (Coscine)

With the FAIR [1] research data management (RDM) platform Coscine [2], researchers from participating institutes, or with an ORCID, can easily make use of research data and metadata management, collaboration and archiving. The stored research data is annotated with semantic metadata according to existing ontologies. In Coscine, projects and resources exist to structure the research process. With the help of persistent identifiers (PIDs), each project and resource in Coscine can be uniquely referenced. In this case, PIDs are identifiers consisting of a URL that references the actual data and prevents inconsistencies when data is moved.

2. Use Case

In material science, reference datasets are the base of every experiment. Converting a universal material science reference dataset to an FDO, eases the exchanging process for the community. A reference dataset must satisfy two things, be visible and be reusable for the material scientists' experiments. From data storing, creating metadata schemas and developing FDOs, the group Task Area Materials Data Infrastructure from NFDI-MatWerk supports the material scientists with services like Coscine.

3. First Concept: Projects and Resources

A new way of approaching complex datasets in Coscine can be solved through implementing the FDO concept based on the Research Data Alliance (RDA) recommendations [3]. To implement an RDA recommended FDO in Coscine, each project PID and

resource PID needs an attached record with fixed types that are describing the content of the FDO and assists in machine-readability and utilization. The information that is stored in the PID record is represented by a Kernel Information Profile (KIP) that is registered in the Data Type Registry (DTR) [4]. This KIP is globally available and is responsible for the machine-readability [3]. Such a KIP has been specifically defined for Coscine. All in all, a PID, a PID Record, the Coscine KIP, and the data type will form a Coscine FDO. Resolving the Coscine PID will convert the Coscine FDO to user and machine-interpretable information.

4. Second Concept: Research Data and Metadata

The second concept for FDOs is more data-oriented than the first one. This means that it follows in the FDO definition that a binary string of data is at the center, semantic metadata is attached to it, operations surrounding it are defined, and a PID exists to uniquely define the FDO. In practice, this should be achieved with technologies like the FAIR Data Point (FDP) [5], which, given a suitable architecture, can provide standardized interfaces to interact with such FDOs. For the individual parts of the FDO, the binary string of data is provided as the research data stored in the real-life application. The metadata provided to this research data and all relevant provenance metadata acts as the semantic metadata. The operations are provided by the API and the FDP, and the PID is represented by using a PURL [6].

5. Linking Concepts

An idea for linking these concepts is that the first concept for projects and resources act as catalogs that contain each other and contain information about their respective entity. In the linking step between the first and second concept, the resources need to point to the research data by using the PURL PIDs of the FDOs. With this linking, it can be ensured that an intelligent agent has a chance to navigate the FDOs to get the requested information. Applying different technologies implementing the FDOs, there is an issue to understanding the FDOs with a single client. Our conclusion is that there is a need for clients that can understand different FDO implementations to fully understand the linked concept.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study. The code implementing our claims can be found here: https://git.rwth-aachen.de/coscine

Author contributions

- Benedikt Heinrichs (Conceptualization, Investigation, Software, Writing – original draft)
- Sirieam Hunke (Conceptualization, Investigation, Software, Writing – original draft)
- Marius Politze (Project administration, Supervision, Writing – review & editing)

Competing interests

The authors declare that they have no competing interests.

Funding

The work was supported with resources granted by NFDI-MatWerk, funded by Deutsche Forschungsgemeinschaft (DFG) with project number 460247524.

References

[1] M. D. Wilkinson et al., "The fair guiding principles for scientific data management and stewardship," *Scientific data*, vol. 3, p. 160 018, 2016. DOI: 10.1038/sdata.2016.18.

[2] M. Politze, F. Claus, B. D. Brenger, M. A. Yazdi, B. Heinrichs, and A. Schwarz, "How to manage it resources in research projects? towards a collaborative scientific integration environment," *European journal of higher education IT*, vol. 1, no. 2020/1, p. 5, 2020, ISSN: 2519-1764. DOI: 10.18154/RWTH-2020-11948. [Online]. Available: https://publications.rwth-aachen.de/record/808269.

[3] *Rda recommendation on pid kernel information profile final*, 2019. [Online]. Available: https://www.rd-alliance.org/system/files/RDA%20Recommendation%20on%20PID%20Kernel%20Information_final.pdf.

[4] Sirieam Hunke and Thomas Jejkal, *Coscine kip in dtr*, 2023. [Online]. Available: http://dtr-test.pidconsortium.net/#objects/21.T11148/8882327b7c25331e3cdd.

[5] da Silva Santos, Luiz Olavo Bonino, K. Burger, R. Kaliyaperumal, and M. D. Wilkinson, "Fair data point: A fair-oriented approach for metadata publication," *Data Intelligence*, vol. 5, no. 1, pp. 163–183, 2023. DOI: 10.1162/dint_a_00160.

[6] K. E. Shafer, S. L. Weibel, and E. Jul, "The purl project," *Journal of Library Administration*, vol. 34, no. 1-2, pp. 123–125, 2001, ISSN: 0193-0826.