3rd FAIR Digital Objects Conference

TIBconferencesession

TIB-OP_will_set_DOI

Published: 1970-01-01

Integrating FAIR Digital Objects and Dataspace Technologies: Towards a Global Research Ecosystem

Julia Pampus[1], Markus Spiekermann[2], Michael Steinert[1], and Peter Koen[3]

[1]Fraunhofer Institute for Software and Systems Engineering ISST, Dortmund, Germany

[2]Huawei Technologies Deutschland GmbH, Munich, Germany

[3]Microsoft Corporation, Redmond, Washington

*Correspondence: Julia Pampus, julia.pampus@isst.fraunhofer.de

Abstract. Dataspaces are a technical approach to trust-based data sharing. This paper explores the integration of dataspace technologies with FAIR Digital Objects (FDOs). We propose a concept for a Globally Integrated Dataspace that connects research organizations, the Global Research Dataspace. Furthermore, we present a reference implementation based on international dataspace standards (ISO/IEC 20151) and the Eclipse Dataspace Components. Our approach demonstrates that combining FDOs with dataspace protocols enables interoperable, verifiable, and decentralized data sharing, fostering a scalable and trustworthy global research data ecosystem.

Keywords: Dataspaces, FAIR Digital Objects, Trust, Data Sharing

1. Introduction

The FAIR principles (findable, accessible, interoperable, and reusable) are central to research data management, but do not specify technologies, standards, or implementation solutions. This variability creates high integration barriers and costs, as existing solutions may vary across regions and heterogeneous infrastructures. In the research domain, unrestricted access to data is crucial for ensuring the quality of scientific work and facilitating mutual learning.

FDOs can be used as a technical baseline for globally connected research data [1]. FDOs are standardized, machine-readable units that contain data, metadata, and other digital resources. Their future management is guided by the STAP principles (sovereignty, transparency, accountability, and persistence). As an implementation approach for addressing these STAP principles, the scientific community is driving the idea of a Globally Integrated Dataspace (GID) [2], [3]. The concept proposes using dataspace technologies to manage protected FDOs with enhanced access control capabilities [3], serving as an additional layer to the existing FDO infrastructures.

In this work, we foster the vision of a GID but propose connecting FDO Repositories directly via dataspace technologies, replacing the original FDO Layer (cf. [2]). In

connecting organizations, dataspace technologies implement functional requirements such as publishing and discovering data offerings, negotiating data usage conditions (i.e. policies), and orchestrating data transfers, as specified by the Dataspace Protocol (DSP) [4]. We present the results of a research project, adopting the methodology of Consortium Research (CR). Over the course of six months and a consortium of experienced architects and engineers from industry and research organizations, we were able to use CR phases-based iterations to come up with a reference implementation of FDOs in dataspaces following ISO/IEC 20151 [5]. In addition, we transfer our learnings to develop a concept of a Global Research Dataspace. Due to our grounding in Design Science Research, we contribute to both the academic knowledge base and practitioners' guidelines, promoting the natural growth of the research ecosystem.

2. Global Research Dataspace

Being located across different federal regions, FDO Repositories must comply with diverse regional regulations and technologies [3]. Figure 1 illustrates a network of research organizations. We outline three scenarios: a 1-1-relation (Scenario A) and a 1-n-relation (Scenario B), each within the exact jurisdictional boundaries, and one across different jurisdictional boundaries, harmonized by supranational legislation (Scenario C). In all scenarios, (a) trust is mainly regulation-based; organizations outside the same jurisdiction are not trusted by default. In addition, (b) migrating data across infrastructures requires significant integration effort, often leading to centralized approaches.

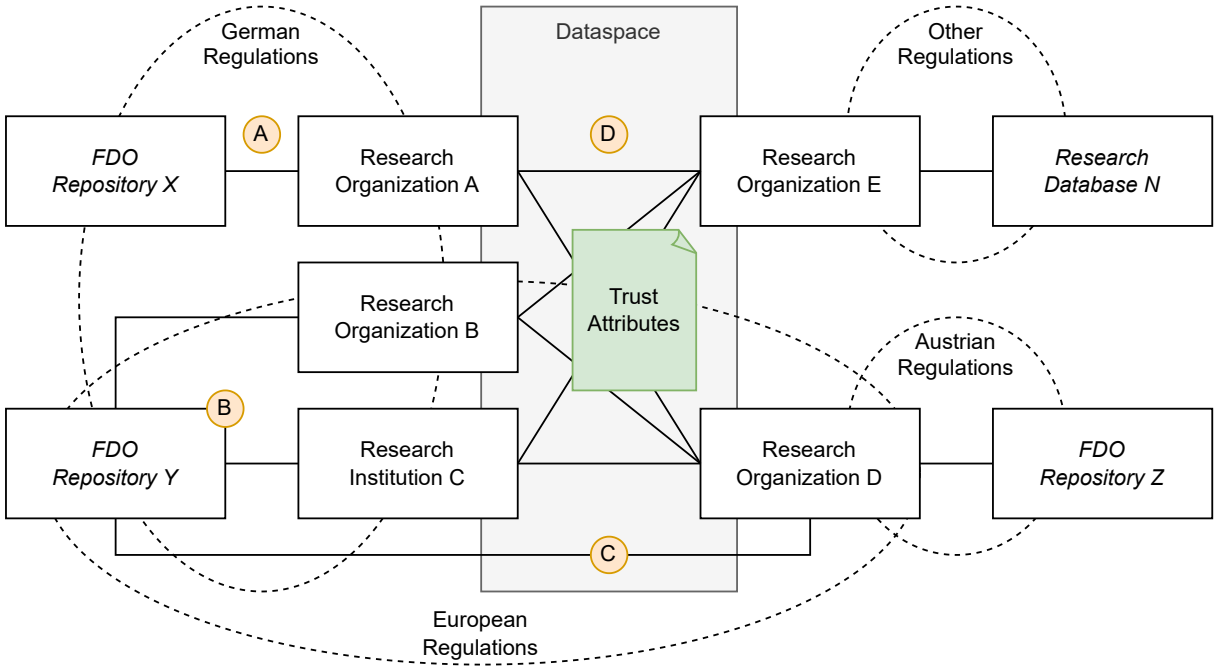

Figure 1. *Illustrative representation of connecting research organizations via a dataspace.*

Referring to the STAP principles (cf. Section 1), connecting research organizations via a dataspace (Scenario D) addresses both challenges as follows:

- **Sovereignty.** Organizations retain control over their data by deciding what data should be shared or consumed and under what conditions, respecting jurisdictional and organizational boundaries. Organizations autonomously define their own trust measures and decide whom to trust. The selected trust attributes are negotiated and validated throughout the data sharing processes. These attributes

are bound to trust anchors that governmental services may represent. However, they can be replaced or supplemented by other entities (e.g., public institutions or other organizations), thus opening, merging, and shifting trust boundaries.
- **Transparency.** Using common vocabularies (Data Catalog Vocabulary (DCAT) [6], Open Digital Rights Language (ODRL) [7]) and protocols (DSP [4]) for sharing information within the dataspace, FDOs are made accessible and resolvable across heterogeneous systems.
- **Accountability.** The implementation of observability measures allows tracking and attributing actions or changes to FDOs along their life cycle.
- **Persistence.** Following a decentralised architecture, FDOs are transmitted peer-to-peer and generally persist in their original repositories without central services acting as intermediaries.

3. Implementation

We describe the integration of FDOs into dataspaces as a five-part integration framework, addressing a) Model alignment, b) Policy representation, c) End-to-end flow, d) Governance and trust, e) Evaluation criteria:

- **Model alignment.** Each DCAT Dataset is re-cast as an FDO, whose persistent identifier (PID) replaces the conventional dataset ID. The dataspace connector serializes these objects as DCAT Catalog, inserting the PID as '@id' and linking to the corresponding FDO Type and Kernel Profile through 'dct:conformsTo'. This alignment establishes an explicit structural bridge between the object's FAIR identity and the Catalog vocabulary.
- **Policy representation.** Within each DCAT entry, the usage constraints of the FDO – such as purpose limitation, attribution duties, or retention period – are expressed as ODRL Policies. Policy engines can natively interpret these policies, aligning FAIR licensing semantics with the ODRL vocabulary.
- **End-to-end flow.** During discovery, the dataspace connector processes the Catalog by resolving each PID to its Kernel metadata and validating the declared type. The subsequent DSP Contract Negotiation process embeds the dataset ID into the Agreement payload, ensuring that the negotiated data usage conditions correspond to a uniquely identifiable digital object and version. During the DSP Transfer Process, the dataspace connector resolves the dataset ID to its current data address, verifies the Kernel metadata, and initiates a data transmission in accordance with the Agreement.
- **Governance and trust.** Lightweight type registries and PID resolution services maintain verifiable definitions for all referenced object types and profiles. Versioning and auditability are reinforced by embedding immutable links between object identity, policy, and provenance.
- **Evaluation criteria.** The dataspace is considered FDO-compliant when every published dataset has a resolvable PID, references its FDO Type and Kernel profile, exposes access rules as ODRL policies, and can be transferred and validated through PID resolution.

Our prototypical implementation of this concept is based on the Eclipse Dataspace Components (EDC)[1], an open-source framework for trust-based data sharing governed by the Eclipse Foundation.

[1] https://github.com/eclipse-edc (Last accessed: 2025-10-12)

4. Conclusions

By merging the robustness of the DSP-based dataspace infrastructure with FDOs, this work illustrates how globally resolvable, typed, and verifiable digital objects can circulate across heterogeneous infrastructures and different trust domains. With this, we emphasize that dataspaces, guided by trust, decentralization, and interoperability, address the STAP principles by design. By adhering to common standards, our approach is not limited to the research domain and can be transferred to any other context using FDOs.

Data availability statement

This work does not include any data.

Underlying and related material

This work is based on fundamental dataspace concepts and characteristics, following the work on ISO/IEC 20151 [5] and the Dataspace Manifesto [8].

Author contributions

Julia Pampus: Writing – original draft, Conceptualization

Markus Spiekermann: Writing – original draft, Conceptualization

Michael Steinert: Writing – review and editing, Conceptualization, Software

Peter Koen: Writing – review and editing, Conceptualization

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This research was partially funded within the Mission AI program, managed by acatech and funded by the former German Federal Ministry for Digital and Transport.

References

[1] Y. Zhang, C.-P. Klas, D. Schiffner, *et al.*, “An FDO-Based Implementation for a Standardized Data Exchange,” in *Open Conference Proceedings*, vol. 5, 2024.

[2] P. Wittenburg, G. Strawn, C. Blanchi, D. Broeder, and U. Schwardmann, “Layered Dataspaces in GIDS,” in *Open Conference Proceedings*, vol. 5, 2024.

[3] P. Wieder, P. Wittenburg, S. Bingert, *et al.*, “Data Economy and FDOs,” in *Open Conference Proceedings*, vol. 5, 2024.

[4] P. Koen, M. Kollenstart, J. Marino, *et al.*, *Dataspace Protocol 2025-1*, Accessed on 2024-10-12, 2025. [Online]. Available: https://eclipse-dataspace-protocol-base.github.io/DataspaceProtocol/.

[5] International Organization for Standardization, *ISO 20151 - Dataspace concepts and characteristics*, Draft standard by ISO, Status: Under development., 2025. [Online]. Available: https://www.iso.org/standard/86589.html.

[6] W3C DCAT Working Group, *Data Catalog Vocabulary (DCAT) - Version 3*, Accessed on 2024-10-12, 2024. [Online]. Available: https://www.w3.org/TR/vocab-dcat-3/.

[7] Open Digital Rights Language (ODRL) Community Group, *ODRL Vocabulary & Expression 2.2*, Accessed on 2025-10-12, 2018. [Online]. Available: https://www.w3.org/TR/odrl-vocab/.

[8] S. Steinbuss, P. Koen, and M. Spiekermann, *The Data Space Manifesto: A vision for trusted data sharing*, 2025. DOI: 10.5281/zenodo.15190876.