1st Conference on Research Data Infrastructure

Connecting RDM

https://doi.org/10.52825/CoRDI.v1i.263

Published: 07 Sept. 2023

FDOs to enable Cross-Silo Work

Peter Wittenburg[1][https://orcid.org/ 0000-0003-3538-0106], Ulrich Schwardmann[2], Christoph Blanchi[3], Claus Weiland[4]

[1] FDO Forum, International

[2] FDO Forum, International

[3] DONA Foundation, Switzerland

[4] Senckenberg Institute, Germany

Abstract. In this paper we describe with the help of two examples how FAIR Digital Objects can be used to bridge between different dataspaces, repositories using different technologies, and model worlds. The first example is the digital specimen as defined by DiSSCO which offers access to a variety of information from different repositories in a persistent way. The second example results from a collaboration with the I4.0 Asset Administration Shell experts where we integrate the two domains using the Digital Product Pass as an example. In this example we also show how FDOs can be used to implement secure access on shared data as it occurs in almost all supply chain processes.

Keywords: FAIR Digital Objects, Secure Data, Data Management, Data Interoperability

1. Introduction

Extensive funding programs especially in Europe (EOSC, NFDI, Gaia-X, etc.) are intensifying the work on data infrastructures in research and industry with the intention to create deep knowledge on all levels from policy to development to finally manage the digital transformation. In parallel the major information companies are also investing huge sums to provide proprietary infrastructure services in a highly competitive landscape. The motivation for these investments becomes increasingly more obvious and urgent: How to democratise the ability to apply statistical methods such as Deep Learning based on AI-ready data.

Each of the many projects being funded create their specific dataspaces aligned with a specific technology stack and a set of rules. If we extend this increasing European fragmentation to the global level we can imagine that we will have an interoperability challenge at several levels – from data modelling to semantics. This situation can be compared with the lack of network technology interoperability a few decades ago when a wide variety of networking solutions were developed until a consensus was built around the internet technology: (1) The agreement on “datagrams” as autonomous entities travelling within the global network. (2) The agreement on TCP/IP as a unifying interface protocol.

2. FAIR Digital Objects

The FAIR Digital Object model (FDO, figure 1) is analogous to the datagrams as autonomous, self-contained and machine actionable entities that persistently bind all information necessary to facilitate its processing across a wide range of different dataspaces independently of their

respective technological choices and sets of rules [1]. Therefore, FDOs are a candidate for achieving basic global data interoperability. The FDO's common data model emerged from

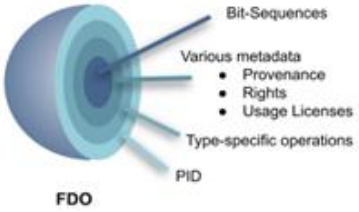

Figure 1. Schematic indication of the FDO Concept.

efforts started within the RDA covering multiple scientific disciplines. Over the last three years the major focus of the FDO work was on combining the basic object model with the FAIR principles and to make it machine actionable. As a result the FDO specifications require that each digital entity be assigned a globally unique, persistent identifier that is resolvable into a predictable, profile specified, set of attribute value pairs that are easily readable and interpretable by any client. Each attribute will be defined and registered in open registries [2]. These attributes can be used to specify a wide range of properties such as checksums to verify the integrity of data, or signatures to verify the authenticity of a claim, or references to resources (data, metadata, configurations, software services, etc.) hosted on repositories of measurable trustworthiness.

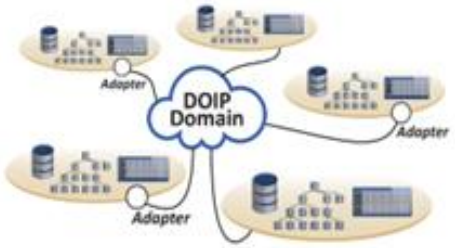

Figure 2. Possibility of creating an integrated FDO domain by using a unique protocol.

In the same way that datagrams are transported on the Internet using TCP/IP, a unifying protocol we call DOIP [3] provides a standard approach for interacting with any FDOs on the Internet independently of the specific technologies used by the various service providers. DOIP acts as the interoperability glue between data service providers of different sorts (figure 2). As an example, we can refer to the development of a DOIP adapter that allowed us to connect the B2Share repository [4] with the CORDRA repository [5] both using different technologies and data models. Other groups have also been using DOIP in a wide range of projects and at large scale. We are currently in the process of integrating some of these different projects in a comprehensive testbed and building on work from KIT and GWDG. This testbed will include repositories from different disciplines and countries/regions to demonstrate the usefulness and scalability of the approach. The goal is to provide validators to enable anyone to connect their repository.

3. Interconnecting with FDOs

The Digital Specimen (DS, figure 3) developed in the context of the DiSSCo RI demonstrates the huge potential of FDOs in managing digital artefacts which are created and curated by numerous institutions [6]. The DS combines a wide range of information of different types (images like scans, sequence data, geolocations, taxonomic classifications, etc.) acting as a Digital Twin of a physical specimen in a scientific collection The DS concept makes also use of another feature of FDOs: direct links between type and registered operations. In DiSSCo, many

FDOs are accessed by clients using operations like feature extraction from digitised content. Such operations can be registered in a registry managed by the resource provider or in open registries managed by other communities of practice.

Intensive interactions with experts who designed and developed the Industry 4.0 Asset Administration Shell (AAS) solution resulted in the design of a smart solution for the common

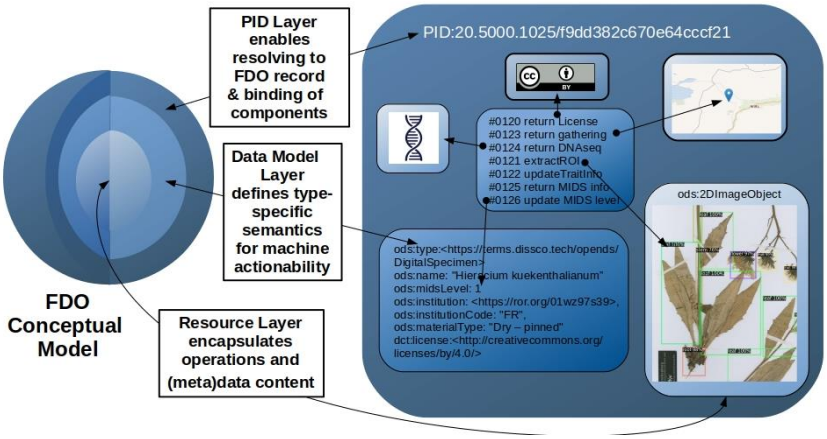

Figure 3. Illustration of the DiSSCO Digital Specimen concept.

challenge of enabling many distributed actors to operate on a shared data structure in a highly secure manner [7]. An example application motivated by industry is the development of an FDO based tracking of the transportation related green-gas emissions integrated with the Dig-

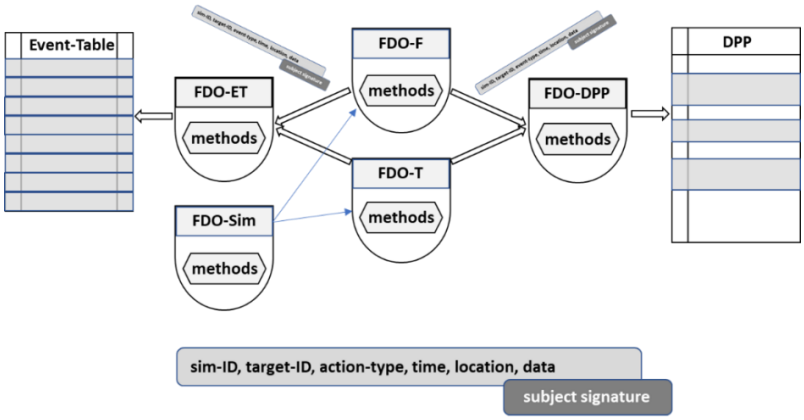

Figure 4. Concept to implement secure sharing of the Digital Product Pass which modelled using the I4.0 AAS technology.

ital Product Pass. All data structures and actors including the DPP are modelled as FDOs and the use of registered typed operation relations is used to enforce secure interactions (figure 4). Operations on FDOs can be compared to methods as they are known from object-oriented programming. Standard security technologies such as digital signatures certificates and PKI infrastructures are core to this approach. This solution demonstrates that industry standards such as I4.0 AAS and FDOs can be combined to achieve smart solutions working across country borders. Stakeholders from industry see much potential here to fill the gap between Operational Technology and Information Technology, which is a key issue for industry.

3. Conclusion

In this contribution based on concrete examples we show that FDOs

- due to their persistent and machine actionable bundling of information are a candidate for establishing the emerging Global Integrated Dataspace
- provide a simple yet extensible solution to structure the huge amount of distributed data sources around concepts such as the digital specimen and to provide this information in a persistent way
- can be used to implement secure interactions on shared resources in a smart way.

FDOs is an open, licence and property rights free concept involving a variety of implementations that will be turned into an International Standard with the help of DIN to enable broad application.

Author contributions

All authors contributed to the whole paper.

Competing interests

The authors declare that they have no competing interests.

Acknowledgement

We would like to acknowledge the work of the FDO community and the excellent collaboration with our colleagues from RWTH Aachen who were involved in the specification and implementation of the I4.0 AAS concept.

References

1. I. Anders, et al., “FAIR Digital Object Technical Overview”, https://zenodo.org/record/7824714#.ZEDULs7P3b0
2. I. Anders, et al., “FDO Forum FDO Requirement Specifications”, https://zenodo.org/record/7782262#.ZEDURs7P3b0
3. DONA, “Digital Object Interface Protocol Specification”, https://www.dona.net/sites/default/files/2018-11/DOIPv2Spec_1.pdf
4. EUDAT. “B2SHARE”, https://b2share.eudat.eu/
5. DONA, “CORDRA – DOIP Client Library”, https://www.cordra.org/documentation/client/doip-java.html
6. DiSSCO, “What is a Digital Specimen?”, https://dissco.tech/2020/03/31/what-is-a-digital-specimen/
7. P. Wittenburg, C. Blanchi, “Interacting FDOs for Secure Processes”, https://www.researchgate.net/publication/364508074_Interacting_FDOs_for_Secure_Processes