A Primer for FDO Profile Creation Based on Early Experience

Wouter Addink [iD]

Naturalis Biodiversity Center, Netherlands

Correspondence: Wouter Addink, wouter.addink@naturalis.nl

Abstract. This paper summarizes the contents of a lightning talk at the FDO Summit 2024 about the creation of a FDO Profile for the kernel attributes in a PID record. It describes the early experience that DiSSCo had with creating FDO Profiles for Digital Specimens and related digital objects. The Digital Specimen FDOs act as a digital surrogate for specimens held in natural science collections. The work done in DiSSCo resulted in a set of 13 core attributes used for all FDO Types in DiSSCo, complemented with additional domain specific attributes. These core attributes and reasons for deviation with the RDA PID Kernel recommendation are presented in the lightning talk.

Keywords: FDO Profile Attributes, Kernel Metadata, PID Profiles, FDO Forum, DiSSCo

1. Introduction

[FDO](#) Profiles, also known as PID Profiles, define metadata attributes about the PID (Persistent Identifier) and, optionally, about the object the PID represents. These profiles are themselves a FDO and are a requirement for FDO implementation. An FDO Profile is instantiated as a FDO record (PID record) containing the metadata attributes, also known as [Kernel attributes](#), with their values as key-value pairs to describe the PID and the thing it refers to (referent). An FDO Profile defines the set of attributes (also known as Kernel attributes) in a metadata record associated with a PID. A PID record which adheres to an FDO profile is known as an FDO record. The FDO profile defines attributes needed for PID management and machine actionability, including locating the instantiated object and/or object metadata if these exist. The instantiated object is a bit sequence that represents an [informational unit](#) such as a document, a dataset, a photo, a service, etc. A FDO record as defined by an FDO Profile should include metadata to support the [FAIR principles](#) which may include metadata describing the referent for disambiguation and reproducibility purposes. However, it should be limited to essential properties for quick inspections and immediate processing according to Anders et al. [1].

2. Which attributes need to be described in a FDO profile?

The FDO PID Profiles & Attributes paper describes what needs to be included in a FDO profile [2]:

- PID, name and address of the creating authority,
- Name and email address of the responsible person
- Date of PID Profile specification
- Number of kernel attributes used
- for each attribute: TypePID, Format, Cardinality

The main part of the FDO profile is the description of the kernel attributes. The FDO Forum aims to define a set of mandatory and optional attributes to be included in FDO Profiles, which may be extended with other optional attributes specified by communities relevant for a specific FDO Type. In the FDO - Kernel Attributes & Metadata Version 1.0, the FDO Forum concluded to only have two obligatory attributes knowing that different communities will use different semantic spaces. These are a reference to the PID Profile and the Type of the FDO and have been included as the mandatory attributes in the FAIR Digital Object Technical Overview [2]. The document also proposed a set of optional attributes but a final specification describing the set of optional attributes has not yet been established, so which ones to include? And where to include community defined metadata attributes: in the PID record, in an object's metadata record, or as part of the object itself? And how to point to that metadata, through a URL in the PID record, or through an operation described in the object's Type definition? The lighting talk will provide some answers to these questions, based on experiences with creating FDO profiles for digital objects in DiSSCo, the Distributed System of Scientific Collections.

3. Core attributes defined by DiSSCo

DiSSCo used the Recommendation on PID Kernel Information as defined by the Research Data Alliance (RDA) PID Kernel Information WG [3] as a starting point, but alterations were needed to support mutable digital objects and digital objects that are containers of other digital objects. Also the terminology was adjusted to FDO terminology. This resulted in a set of 13 mandatory kernel attributes as core attributes used for all FDO Types in DiSSCo complemented with additional domain specific attributes (see DiSSCo Kernel and Digital Specimen FDO Record attributes). The core attributes are listed in Table 1.These core attributes and reasons for deviation with the RDA PID Kernel recommendation are presented in the lightning talk.

Table 1. Mandatory attributes described in FDO profile for FDOs in DiSSCo

Index	Attribute	Example	Cardinality	Type
1	**fdoProfile** PID to a machine readable description of the attributes in the FDO record	https://hdl.handle.net/21.T11148/d8de0819e144e4096645	1/1	pid
2	**fdoRecordLicense** the licence for the FDO record, required to be always public domain	https://creativecommons.org/publicdomain/zero/1.0/	1/1	string
3	**digitalObjectType** PID to a description of the Type of digital object that defines the metadata, bit sequences (if any) and operations for the object	https://hdl.handle.net/21.T11148/894b1e6cad57e921764e	1/1	pid
4	**digitalObjectName** name of the object type for humans	digital specimen type 1	1/1	string
5	pid	https://doi.org/10.22/GEE-W3J-HL2	1/1	pid

	the PID of which the FDO record is part, in DiSSCo this is a Handle or DOI. It is recommended to store this pid also in the local collection management system for the specimen.			
6	pidIssuer In case of a DOI this is a PID for the DOI Registration Agency	https://hdl.handle.net/10.17183	1/1	pid
7	pidIssuerName	DataCite	1/1	string
8	issuedForAgent In the case of a digital specimen, this is a PID for DiSSCo as the agent responsible for serving the digital specimen object	https://hdl.handle.net/10.22	1/1	pid
9	issuedForAgentName	DiSSCo	1/1	string
10	pidRecordIssueDate date the pid record was created	2022-11-24	1/1	date
11	pidRecordIssueNumber starts with 1 and is incrementally increased by 1 every time the pid record is updated. Compatible with DOI schema requirements.	2	1/1	int
12	structuralType Nature of the digital object, compatible with DOI schema requirements. The nature of a digital specimen object is always "digital". Other digital objects (outside DiSSCo) could be of physical, performance or abstraction nature.	digital	1/1	literal
13	pidStatus A PID is considered to have a lifecycle, PID status indicates the status in the life cycle, e.g. draft, active, retired. PID statuses are described further in the PID infrastructure design. One of: one of: DRAFT, ACTIVE, RETIRED, OBSOLETE, FAILED, MERGED, SPLIT	DRAFT	1/1	vocab

Data availability statement

This submission describes the results of a request for comments (RFC) process, it is not based on data

Underlying and related material

The RFC document that was the basis for this paper can be found here: but on the RFC document available here: https://docs.google.com/document/d/1hj5M9Cko5LduDv1H8CUYdGocg_B3YUf6MafoZPHPaec.

Author contributions

The authors' contributions according to the CreDIT guidelines: Writing – original draft and Writing – review & editing.

Competing interests

The author declares that he has no competing interests.

Funding

Acknowledgement

References

[1] I. Anders., M. Hellström, S. Islam, T. Jejkal., L. Lannom., U. Schwardmann, P. Wittenberg, "FDO PID Profiles & Attributes", 2022, doi: https://doi.org/10.5281/zenodo.7825630

[2] I. Anders, C. Blanchi, D. Broder, M. Hellström, S. Islam, T., Jejkal, L. Lannom, K. Peters-von Gehlen, R. Quick, A. Schlemmer, U. Schwardmann, S. Soiland-Reyes, G. Strawn, D. van Uytvanck, C. Weiland, P. Wittenburg, C. Zwölf, "FAIR Digital Object Technical Overview", 2023, doi: https://doi.org/10.5281/zenodo.7824714

[3] T. Weigel, B. Plale, M. Parsons, G. Zhou, Y. Luo., U. Schwardmann, R. Quick, M. Hellström, K. Kurakawa, "RDA Recommendation on PID Kernel Information (Version 1)", 2018, doi: https://doi.org/10.15497/RDA00031