FAIR DIGITAL OBJECTS FORUM

FDO Forum FDO Requirement Specifications Version 3.0

FDO Forum (FDOF) Proposed Recommendation 12. January 2023

Current and previous versions:
FDO Requirement Specification P3.0: **PR-RequirementSpec-3.0-??**
??
FDO Requirement Specification PR2.0: **PR-RequirementSpec-2.1-20221017**
https://docs.google.com/document/d/1aGA-TBr4XpORhMPtnf_--Nb4FYJccgeSvGmGh68jNws/edit

Editors:
G. Strawn, P. Wittenburg

Authors: I. Anders, Ch. Blanchi, Daan Broeder, M. Hellström, Sh. Islam, Th. Jejkal, L. Lannom, K. Peters, R. Quick, A. Schlemmer, U. Schwardmann, St. Soiland-Reyes, G. Strawn, D. van Uytvanck, C. Weiland, P. Wittenburg, C. Zwölf

Abstract

This FDO Requirement Specification document specifies the criteria which digital entities need to fulfill to be accepted as FAIR Digital Objects. These specifications need to be adhered to by all implementations that claim to be FDO compliant. The requirements are written at a level that allows for different implementations. The specifications are split into generic guidelines to characterize the landscape in which FDOs will exist and more specific requirements.

Status of this document

This FDO Requirement Specification Proposed Recommendation PR 3.0 is a new major version that results from the previous PR2.1 version, which was anchored in the agreements of the Paris meeting [1], and in addition it contains requirements which are extracted from all FDO Specification documents which have a Prepared Recommendation or Proposed Endorsed Recommendation status according to the FDO Document Process standard.

This FDO Requirements Specification might be amended with implementation guidelines for compliant technology stacks, such as Digital Object Architecture and Linked Data. While these

implementation examples are not part of these specifications, they can be used for illustration purposes.

Formal References

This document is based on the following official FDO Documents:

- FDO Configuration Types: https://docs.google.com/document/d/1LojKBfuzuI6tKSMSZM77qd45lHmMFAxIeZR7hl1DAUk/edit?usp=sharing
- FDO Machine Actionability: https://docs.google.com/document/d/1PPGW83siPDMLG5QRHsKM2cOUPT-O5SqS_jZvs_AO1Qw/edit?usp=sharing
- FDO PID Profiles and Attributes: https://docs.google.com/document/d/1iirLKoWiG9D2ZZ65RqVk89bCwSYwBVHBxk5gLmeNigI/edit?usp=sharing
- Mandatory and and Optional Kernel Attributes: https://docs.google.com/document/d/11CuSoNOpg3vYaoHqi4Yg6ncP3qOcUIC/edit?usp=sharing&ouid=105963437005859657844&rtpof=true&sd=true
- FDO Granularity, Versioning and Mutability: https://docs.google.com/document/d/1jc7FRtzosJBHzkl5oVh-1USnYr3tB_RKDGrDq4uynUg/edit?usp=sharing
- FDO Typing: https://docs.google.com/document/d/1X0hcOVIqP7iYIJf9u-7x3RwcXK8ecsauy0FZq_6-Bq0/edit?usp=sharing
- FDO Implementing Attributes, Profiles, Types (in progress): https://docs.google.com/document/d/1bS1wtBjleZUWFZQuXlH2mD3H-qft7cj_bfY1USy5jqQ/edit?usp=sharing
- FDO Upload Document: https://docs.google.com/document/d/1NkkdiIslxuCrvexSKtQXIPopMhE7nHOpd8X4m-3YJ5s/edit?usp=sharing
- FDO DOIP Endorsement: https://docs.google.com/document/d/1VJjpSztDm9HqR3zKrAeZXR55llorajkvfMUQ0HbbPTk/edit?usp=sharing

Acknowledgements

We would like to thank the Paris meeting participants for their initial contributions to the first external version which was edited by Luiz Bonino and Peter Wittenburg, the many contributions made by all FDO Working group members in finalizing the FDO Specification documents.#

1. Generic Guidelines

These generic guidelines are meant to describe the landscape in which FDOs can exist independent of any specific compliance requirements.

G1: FDOs should provide a path for long term infrastructure investments that is not tied to any particular technology stack.

G2: FDOs need to generate trust in accurate data survival over long periods of time, assuring researchers, funders, and developers that their significant effort in reusing them will be worthwhile.

G3: FDOs must offer compliance with the FAIR principles through measurable indicators of FAIRness.

Note: The FDO Forum will not become engaged in measuring FAIRness. Initiatives organised in the framework of RDA (https://www.rd-alliance.org/groups/fair-data-maturity-model-wg) or CODATA (https://codata.org/worldfair-global-cooperation-on-fair-data-policy-and-practice-a-major-two-year-project-starts-today/) are working on this topic.

G4: FDOs need to support machine actionability as being specified by FDO PR-MachineActionDef-2.0 [2] or later.

G5: FDOs need to support the abstraction principle, i.e., abstracting away details that are not needed at the basic object management level. At that level there is no need to distinguish among different types such as data, metadata, software, semantic assertions, etc., for data management operations.

G6: FDOs need to support stable bindings among all information entities required for machine navigation of the global data space through the use of global, unique, and resolvable persistent identifiers.

G7: FDOs need to support encapsulation**,** such that operations can be associated with FDOs of all types.

Note: A typing system for FDOs is currently being worked out by FDO Forum [3].

G8: FDOs need to support technology independence**,** allowing implementations using different technologies.

G9: FDOs need to comply with minimal agreed standards, e.g., for movement of FDOs between systems, for interaction with FDOs, etc., to guarantee FDO interoperability across heterogeneous systems.

Note: G9 needs to be amended when specific implementations have been made available as basis for FDOs and when additional standardization processes have been accomplished.

2. FDO Requirement Specifications

2.1 General Requirements

FDOR-GR1: A *PID*, standing for a globally unique, persistent and resolvable identifier, is assumed to be the basis for FAIR Digital Objects. Every FDO is assigned one or more PIDs.

FDOR-GR2: A PID resolves to a structured record (*PID Record)* compliant with a specified *PID Profile* which leads to resolution results that enable programmatic resolution from PID back to the FDO and its elements as specified by these requirements. PID Records represent the information characterising FDOs and together with their resolving PIDs they can themselves be FDOs.

Note: The term "PID Record" is synonymous to the term "FDO Record".

FDOR-GR3: If an FDO contains a structured bit-sequence, the structured PID record includes at least a reference to the location(s)[1] where the bit-sequence encoding the content of a FAIR-DO (FDO) can be accessed as well as the type definition of the FDO. The structured record may also contain PIDs pointing to *Metadata FDOs* describing properties of the target FDO.

FDO-GR4: The PID record needs to contain mandatory FDO attributes, may contain optional FDO attributes and attributes agreed upon by recognized communities. Values of attributes can be part of the PID record or they can be references. Expectations of which attribute values are contained within the PID record and which are references pointing from the record to external sources should be specified in the PID profile or definition of said attribute in a Data Type Registry.

FDO-GR5: Each FDO identified by a PID can be accessed or operated on using an interface protocol by specifying the PID of a registered supported operation.

FDO-GR6: Any basic FDO interface protocol offers standard Create, Read, Update, Delete (CRUD) operations on FDOs and a possibility to use extended/domain operations for specific FDO applications. The addition of an operation to list available extended/domain operation for specific FDO types is strongly recommended.

FDO-GR7: The relations between FDO Types and supported operations are maintained in separate registries.

FDO-GR8: Metadata can themselves be FDOs which describe the properties of the referenced FDO. They must be specified by a registered schema that refers to defined and registered metadata categories.

FDO-GR9: Metadata can be of different types such as descriptive, domain specific, provenance, system, access permissions, transactions, etc.

FDO-GR10: Metadata schemas are maintained by communities of practice and are FDOs. Such metadata schemas should therefore themselves follow FAIR principles.

FDO-GR11: A collection of FDOs is also an FDO. The content of collection FDOs describes its construction using an agreed formal language which specifies the relationships of the constituent members. An FDO may be a member of several collections.

FDO-GR12: Deletion of an FDO must lead to standardised and thus machine interpretable tombstone notes in metadata and PID records. The PID itself is not deleted.

[1] Note that 'location' here means the network access point providing FDO access, not necessarily the network storage location of the underlying bits.

FDO-GR13: The PID resolution and the FDO Layer information must be “machine actionable” i.e., are machine interpretable and belong to a type for which operations have been specified in symbolic grammar.

FDO-GR14: FDOs can be configured in different ways as long as the configurations are compliant with the FDO Specifications.

FDO-GR15: The granularity of FDOs is dependent on pragmatic utility decisions within the communities of practice Those communities define the level of useful entities to use in the corresponding application field.

2.2 PID Layer Specifications

FDO-PIDR1: In accordance with the FAIR principles the protocols to resolve a PID into useful information need to be standardised, freely accessible, and publicly usable.

FDO-PIDR2: The resolution of a PID needs to lead to predictable results which are specified by a machine actionable FDO profile.

FDO-PIDR3: The PID is resolved to a set of key attribute-value pairs defined by the FDO profile. The profile needs to be machine actionable, i.e., all attributes need to be defined and registered in open registries which are compliant with the overall FDO or community-specific standards.

FDO-PIDR4: The PID system which is used for the identification of FDOs must be global, robust, scalable, and demonstrate persistence.

FDO-PIDR5: The PID System which is used for the identification of FDOs must support high security capabilities. The owner, or owner-delegated agent, of a PID and its associated attribute-value pairs is the only actor allowed to make changes, to define accessibility to attribute-value pair information, and to request encryption of information.

FDO-PIDR6: Management access to the PID system needs to be secured by a public key infrastructure and if necessary the use of standardised certificates must be possible.

2.3 FDO Layer Specifications

This document does not include details about the schemas defining FDO profiles and attributes. For this we refer to the document on **Implementation of Attributes, Types, Profiles and Registries**.

FDO-FDOR1: The content of each FDO record must be structured according to an FDO profile in accordance with an FDO defined schema registered in a recognized registry.

FDO-FDOR2: The FDO record consists of a set of attribute-value pairs as defined by the FDO Profile and all used attributes need to be defined and registered according to the type specification schema.

FDO-FDOR3: Each FDO record needs to contain the mandatory kernel attributes, as defined by the FDO profile, including the type of the FDO.

FDO-FDOR4: FDO records can include a) optional kernel attributes which are defined by the FDO Forum or b) can include other community defined and registered attributes if machine actionability is guaranteed.

FDO-FDOR5: In the canonical case that an FDO includes bit-sequences encoding FDO content, the attribute-value set must contain the reference to these bit-sequences and a rich metadata description as requested by the FAIR principles.

FDO-FDOR6: Every FDO must have an FDO Type, facilitating automated machine validation and processing of the FDO.

FDO-FDOR7: An FDO may have included any sort of bit-sequences (data, metadata, software, semantic assertions, etc.) but each must be typed using a type-value pair in compliance with the FAIR principles.

FDO-FDOR8: FDOs can use any existing types or define new types, either completely new or derived from existing types.

FDO-FDO9: The type of an FDO is specified using an agreed upon syntactic description that may be structured and dependent on components that are themselves FDOs. This syntactic description is called FDO Type definition in the following and is an FDO itself and thus FAIR compliant.

FDO-FDO10: Every FDO Type and type is associated with a PID which resolves to the FDO that includes the type definition. So it is of type FDO type definition.

FDO-FDO11: The FDO Type Framework specifies methods for defining types for and attaching types to FDOs

2.4 Resource Layer Specifications

As has been mentioned, the resources referenced by the FDO are being managed by different service providers and we are confronted with a great amount of legacy data and a variety of practices. Therefore, we can only make general statements and should expect an extended period during which existing approaches will move towards interoperability and machine-navigability.

FDO-RESR1: All resources referenced by FDOs must be FAIR compliant.

FDO-RESR2: All resource and service providers should demonstrate trustworthiness by regularly assessing the quality of the procedures applied according to standards such as CoreTrustSeal.

3. References

[1] FDO Framework V1.02: https://github.com/GEDE-RDA-Europe/GEDE/blob/master/FAIR%20Digital%20Objects/FDOF/FAIR%20Digital%20Object%20Framework-v1-02.docx
[2] FDO Machine Actionability: PR-MachineActionDef-2.0-20220611
[3] Typing Document Ref. to come WD-TypingFDOs-1.0-20220310

4. Changes from previous versions

Version	Who	Date	Comment
FDO F V0.2	Paris Meeting recommendation	December 2019	This document was endorsed by the participants of the Paris Meeting after editing by L. Bonino and P. Wittenburg
WD0.3	FDO R Editors	February 2022	since the version FDO F V1.02 was broadly discussed in the FDO TSIG WG we assign it the version number 0.3, many changes have been applied to this version which we will not document in detail, we needed to remove the term “framework” since this term is now used by L. Bonino for his website and FDO Forum needed to urgently overcome huge confusions
WD1.0	FDOR Editors	March 2022	For FDOR3 a further note was added to indicate the necessity of three attribute categories: FDO mandatory, FDO optional, Community specified. - also the empty chapter 4 “Errata” was removed since we will use Google docs to document discussions.
PR2.0	FDOR Editors & Authors	August 2022	- The editors, G. Strawn and P. Wittenburg, suggested to include all major commenters as authors of this document which they accepted - Improvement of the G1 Formulation - Update of the reference in G4 - Reformulation of FDOR2 and adding a note to capture the comments made, introducing the synonym “FDO Profile” - change of the order in FDO3 and FDO4 including renaming them - Reformulation of the new FDOR3 to capture the comments made - Reformulation of the new FDO4 to capture the comments made - Improvement of the formulation of FDOR5 and FDOR7 to include "supported" - Improvement of the FDOR8 formulation - Extension of the FDOR11 formulation to capture the comments. - Improvement of the FDOR12 formulation
PR2.1	Editors	17.10.2022	- no further comments were made

PR3.0	Editors	12.1.2023	- major extension of PR2.1 - all requirements specifications defined in the various FDO Specification documents have been added to this version