FIPs and Practice

Barbara Magagna[‡], Erik Anthony Schultes[§,|], Marek Suchánek[¶], Tobias Kuhn[#]

‡ GO FAIR Foundation, Leiden, Netherlands
§ Leiden University, Leiden, Netherlands
| Leiden Center for Data Science, Leiden, Netherlands
¶ Czech Technical University in Prague, Faculty of Information Technology, Prague, Czech Republic
# Vrije Universiteit Amsterdam, Department of Computer Science, Amsterdam, Netherlands

Corresponding author: Barbara Magagna (barbara@gofair.foundation)

Abstract

There is no doubt that FAIR Data and Services (GO FAIR Foundation 2019b) are needed to enable data intensive research and innovation. While the FAIR Guiding Principles (Wilkinson et al. 2016) specify the expected behaviors of digital artifacts they do not specify the technology choices actualizing these behaviors, leaving maximum freedom to operate for communities of practices. In recent years, EU research funding including projects associated with the European Open Science Cloud (European Commission 2020) have been coupled to the FAIRness of the actual data and service landscape and promote the development of FAIR implementation roadmaps. Begining in 2019, ENVRI-FAIR (ENVRI Community 2019) used a text-based questionnaire approach with more than 50 questions aiming to capture details on how the FAIR Principles were implemented in the Research Infrastructure data architectures. The responses however, were not directly usable for downstream analysis without substantial post-processing and harmonization (Magagna et al. 2020). At the same time, the GO FAIR Foundation (GO FAIR Foundation 2022) was leading the development of the FAIR Implementation Profile (FIP) concept. A FAIR Implementation Profile is a list of declared technology choices intended to implement each of the FAIR Guiding Principles, made as a collective decision by the members of a particular community of practice (Sustkova et al. 2020, Schultes et al. 2020). The FIP provided a more structured and efficient way of cataloging FAIR-realted implementations, but to scale the approach it was necessary to develop an ontology linked to machine-actionable questionnaires.

In a collaborative effort, the FIP ontology (Kuhn et al. 2020) was developed in order to specify questions prompting FAIR Implementation Communities (FICs) to explicitly declare for each of the FAIR Principles, the FAIR Enabling Resources (FER) that the community uses to implement them. FERs are defined as digital objects that provide functions needed to achieve some aspect of FAIRness. Twelve different types of FERs were identified (Magagna and Schultes 2022), such as identifier service type (FAIR Principle F1), lookup-service type (F4) or structured vocabulary type (I2). The FIP as a whole is interpreted as community specific metadata and as such suitable for addressing the directive to reuse

"domain-relevant community standards" given in Principle R1.3. To meet the requirement explicitly expressed by the ENVRI communities the FER concept has been extended to include its status as: already available for deployment or are under development; deployed currently or are planned to be deployed in the future; planned replacement of one FER by another in future.

The Data Stewardship Wizard (DSW, ds-wizard.org), initially developed as a data management planning (DMP) tool in ELIXIR (Hooft et al. 2016), was desinged with all the required features providing a machine-actionable questionnaire (Pergl et al. 2019). Using knowledge models (KMs, Codevence 2020b) and Jinja2-based document templates (Anonymous 2007), it provides a versatile way to define decision trees in the form of smart questionnaires as well as their transformations to practically any textual and machine-readable document, for example, JSON or RDF or even FAIR Digital Object-like nanopublications. These capabilties were leveraged to build the FAIR Implementation Profile Wizard (FIP Wizard, Codevence 2020) where different KMs were used to create nanopublications within the Wizard environment for 1) FAIR Enabling Resources, 2) FAIR Implementation Communities, 3) Metadata Longevity Plans (the FER type for Principle A2), and 4) the FAIR Implementation Profiles themselves. Users open a questionnaire, fill it in, and then they can preview and publish a nanopublication. The relationships between these FIP-related nanopublications are made by linking the Uniform Resource Identifiers (URI, W3C 2001) of the nanopublications via drop-down lists and autocomplete features in the FIP Wizard questionnaire interface. This functionality is achieved using real-time API calls to a nanopublication querying service as the user types. As curatorial feedback, the user can see various metadata related to the matching nanopublications (e.g. description, timestamp, or approval badge). A custom submission service allows the submission of the nanopubs to a FIP triple store (Kuhn 2020). SPARQL queries can then help to produce the FIP matrix, a cross table for FERs and Communities illuminating convergence opportunitesd by FER reuse over the technological and community landscapes.

In this way FIPs are published by the FIP Wizard as FAIR (machine-readable) and Open data, which can then serve as a reference for practical FAIR data stewardship activities conducted by members of that, and other communities. A FIP Wizard module linking the FIP to the DMP based on a mapping of the underlying KMs is under development and will translate the FIP into clear community-specific directives that data stewards can subsequently implement. FIP publication also encourages FIP reuse and repurposing by other communities, which saves time ‘reinventing the wheel’ and simultaneously drives convergence on FAIR implementation choices. Over time, FIPs need to be updated to fit the purposes of the community and to accomodate the ongoing development of FAIR technologies, such as FAIR Digital Objects. The FIP Wizard supports systematic versioningand it is anticipated that this revision legacy can later be mined for insights into FAIR-related technology trends.

The FAIR Implementation Profiles and Practice (FIPP) working group (GO FAIR Foundation 2019) focuses on the role of FAIR Implementation Profiles (FIP) in the FAIR Digital Object (FDO) space. FIPs impact the FDO development in 2 principal ways:

1. FIPs provide a **socio-technical approach** for driving the explicit and systematic community agreements on the use of FAIR implementations including domain-relevant community standards and emerging technologies like FDOs. Hence, FIPs can catalyze FAIR convergence both within and between disciplinary domains. The FIPP WG activities have leveraged FIPs in gaining commitment to FAIR practices, and in some cases, early interest in the deployment of FDOs among circumscribed communities.
2. Once a community of practice declares a FIP (in machine-readable format), the **FIP itself becomes a type of FAIR metadata** that describes the technical and semantic composition of FDOs that will be subsequently created within that community following FAIR Principle R1.3. The FIPP WG explores how FIPs may be used to instruct computation agents to interpret, interoperate with, and perform operations on FDOs as they are constructed under given FIPs.

Keywords

FAIR Principles, FAIR Implementation Profile, FAIR Digital Object, FAIR Enabling Resources, FIP Wizard, Community

Presenting author

Barbara Magagna

Presented at

First International Conference on FAIR Digital Objects, presentation

Conflicts of interest

References

- Anonymous (2007) Jinja. https://jinja.palletsprojects.com/en/3.1.x/. Accessed on: 2022-9-02.
- Codevence (2020a) FIP Wizard. https://fip-wizard.ds-wizard.org/. Accessed on: 2022-9-02.
- Codevence (2020b) Data Stewardship Wizard terminology. https://guide.ds-wizard.org/general/terminology. Accessed on: 2022-9-02.

- ENVRI Community (2019) Environmental Research Infrastructure building FAIR services for research, innovation and society. https://envri.eu/home-envri-fair/. Accessed on: 2022-9-02.
- European Commission (2020) European Open Science Cloud. https://eosc.eu/about-eosc. Accessed on: 2022-9-02.
- GO FAIR Foundation (2019a) FAIR Convergence Matrix & FAIR Implementation Profiles. https://www.go-fair.org/today/fair-matrix/. Accessed on: 2022-9-02.
- GO FAIR Foundation (2019b) The Internet of FAIR Data and Services. https://www.go-fair.org/resources/internet-fair-data-services/. Accessed on: 2022-9-02.
- GO FAIR Foundation (2022) GO FAIR Foundation team. https://www.gofair.foundation/about. Accessed on: 2022-9-02.
- Hooft R, Pergl R, Suchánek M (2016) ELIXIR Data Stewardship Knowledge Model. https://github.com/ds-wizard/ds-km. Accessed on: 2022-9-03.
- Kuhn T (2020) SPARQL endpoint. https://virtuoso.nps.petapico.org/sparql. Accessed on: 2022-9-02.
- Kuhn T, Magagna B, Schultes E (2020) FAIR Implementation Ontology (FIP). https://peta-pico.github.io/FAIR-nanopubs/fip/index-en.html. Accessed on: 2022-7-10.
- Magagna B, Adamaki A, Liao X, Rabissoni R, Zhao Z (2020) ENVRI-FAIR D5.1 Requirement analysis, technology review and gap analysis of environmental RIs. Zenodo https://doi.org/10.5281/zenodo.3884998
- Magagna B, Schultes E (2022) FAIR Enabling Resource (FER) Types. https://osf.io/vst8a. Accessed on: 2022-7-10.
- Pergl R, Hooft R, Suchánek M, Knaisl V, Slifka J (2019) “Data Stewardship Wizard”: A Tool Bringing Together Researchers, Data Stewards, and Data Experts around Data Management Planning. Data Science Journal 18 https://doi.org/10.5334/dsj-2019-059
- Schultes E, Magagna B, Hettne KM, Pergl R, Suchánek M, Kuhn T (2020) Reusable FAIR Implementation Profiles as Accelerators of FAIR Convergence. Lecture Notes in Computer Science138-147. https://doi.org/10.1007/978-3-030-65847-2_13
- Sustkova HP, Hettne KM, Wittenburg P, Jacobsen A, Kuhn T, Pergl R, Slifka J, McQuilton P, Magagna B, Sansone S, Stocker M, Imming M, Lannom L, Musen M, Schultes E (2020) FAIR Convergence Matrix: Optimizing the Reuse of Existing FAIR-Related Resources. Data Intelligence 2: 158-170. https://doi.org/10.1162/dint_a_00038
- W3C (2001) URIs, URLs, and URNs: Clarifications and Recommendations 1.0. https://www.w3.org/TR/uri-clarification/. Accessed on: 2022-9-02.
- Wilkinson M, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J, da Silva Santos LB, Bourne P, Bouwman J, Brookes A, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo C, Finkers R, Gonzalez-Beltran A, Gray AG, Groth P, Goble C, Grethe J, Heringa J, ’t Hoen PC, Hooft R, Kuhn T, Kok R, Kok J, Lusher S, Martone M, Mons A, Packer A, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S, Schultes E, Sengstag T, Slater T, Strawn G, Swertz M, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3 (1). https://doi.org/10.1038/sdata.2016.18