International FAIR Digital Objects Implementation Summit

TIBconferencesession

TIB-OP will set DOI with \TIBdoi

Published: 1970-01-01

FAIR Implementation Profiles for the Social Sciences and Their Use Cases

Shuai Wang [1][https://orcid.org/0000-0002-1261-9930], Angelica Maineri[2][https://orcid.org/0000-0002-6978-5278], Navroop K. Singh[1][https://orcid.org/0000-0001-9131-3528] , and Tycho Hofstra[1][https://orcid.org/0009-0009-7320-864X]

[1]Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[2]Erasmus University Rotterdam, the Netherlands

Abstract: FAIR Implementation Profiles (FIPs) are created and published to capture decisions made by communities on data curation and management. Despite that many FIPs were created in domains such as medicine, environmental science, few FIPs were available in social sciences. This extended abstract reports recent advances in creating and using FIPs in social sciences. It consists of a summary of available FIPs in social sciences. Three use cases were included to demonstrate how FIPs can be used to guide data management for researchers, organisations, and communities. Finally, we envision some future work on FIP development.

Keywords: FAIR Implementation Profile, data management plan, data management

1 FAIR Implementation Profiles for the Social Sciences

A FAIR Implementation Profile (FIP) [1] is a special kind of FAIR Digital Object (FDO) that consists of questions and answers about communities' decisions about the use of FAIR enabling resources (FERs) regarding the FAIR principles. FERs include repositories, identifier services, registries, knowledge representation languages, licenses for data and metadata, etc. Many FIPs have been created but only a few stem from social sciences communities. The social sciences offer interesting cases since they have a long-standing tradition of data sharing but adopt heterogeneous standards in doing so. In this paper, we provide a summary of recent advances in FAIR implementation profiles for social sciences and outline how FIPs can capture the decisions of communities in social sciences. More specifically, use cases are included to demonstrate how FIPs can be used to improve data management in multiple ways.

Recently, six FIPs in social sciences were compared [2]. Among them, three FIPs were reused or upgraded from existing attempts, and published using the FIP Wizard[1]. The original Social Science Survey Research (SSSR) was based on an outdated knowledge model. The other two FIPs corresponding to the European Social Survey (ESS) and the Australian Social Survey International – ESS (AUSSI-ESS), respectively, were sourced from a report produced by the WorldFAIR project. In addition, three new

[1]https://fip-wizard.ds-wizard.org/

FIPs were published. First, the SSHOC-NL Socio-Economic History (SEH) community studies social and economic history using a variety of historical sources. Second, the Media Content Analysis Lab (MCAL)[2] is a community of communication science scholars that aims to facilitate the sharing and analysis of large digital media content collections. Lastly, the LGBTQ+ Linked Open Vocabulary (LGBTQVoC) community consists of contributors from various countries who create multilingual LGBTQ+ controlled vocabularies[3] for indexing digital records, literature, and heritage of LGBTQ+ objects in various languages.

2 Use Cases

2.1 Use Case 1: Comparison of FAIR Implementation Between Communities

By listing the FERs used in FIPs on different aspects of the FAIR principles, one can study how communities differ from each other [1], [2]. For example, while Dataverse[4] and the DANS SSH Data Station[5] overlap for SEH and MCAL, the corresponding entry for LGBTQVoc is missing. This indicates that the LGBTQVoc community has not published their data in any public registry. A summary of the overlapping FERs between communities can be made in the form of a table [2]. It was reported that the SEH and LGBTQVoc communities are closer by sharing some FAIR Implementation decisions but not between ESS and AUSSI-ESS. The analysis can be used for further discussion on the convergence of communities.

2.2 Use Case 2: Aligning community standards in DMPs using FIPs

It can be different for researchers to consider community standards while writing Data Management Plans (DMPs) due to lack of awareness of data standards and practices of one or multiple communities. It has been proposed that FIPs can be used to serve as suggestions for aligning research data management with community standards (see also [3]). A recent effort is the development of a customized interface with the related information extracted from FIPs as suggestions. The study aimed to understand how researchers can take such suggestions into account when writing DMPs [4]. The findings affirm the potential of FIPs as a valuable resource to harmonize research data with community standards.

2.3 Use Case 3: Using FIP as guidance for the upcycling of legacy data

Most recently, it was demonstrated how a FIP can be used for the reuse of legacy data (i.e. data upcycling) in the SSHOC-NL Socio-Economic History Community [5]. The FIP can guide data upcycling by offering suggestions on which standards and technologies should be adopted. For instance, FIP can give explicit advice on repositories to store new versions of upcycled data and suggest what PID (persistent ID) should be used.

[2] https://odissei-data.nl/en/media-content-analysis-lab/

[3] Members of the community translate terms in Homosaurus (https://homosaurus.org/) and use related resources for the curation of linked data.

[4] https://dataverse.nl/

[5] https://ssh.datastations.nl/

3 Conclusion and Future Work

This extended abstract provides a summary of published FIPs in social sciences. Three recent use cases are included to demonstrate how these FIPs could be used. More FIPs are expected to be published in social sciences. Thus, it would be possible to do a detailed comparison between multiple communities that are closely related and discuss their convergence. The integration of FIP information into the legacy data upcycling workflow can be further explored [6]. Comparing an FIP with the result of the FAIR assessment of datasets of the corresponding community is a logical continuation of the work we describe in this paper.

Data availability statement

All FIPs and their reports mentioned in the paper are available on Zenodo.[6] The knowledge model corresponding to the DMP template of the Vrije Universiteit Amsterdam, the survey, and all related documents are also available on Zenodo[7].

Author contributions

A.M.: Conceptualization, Data curation, Formal analysis, Methodology, Project administration, and Writing - review & editing. N.K.S.: Data curation and Resources. S.W.: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Supervision, Writing - original draft, and Writing - review & editing. T.H.: Resources and Validation.

Funding

The project was carried out at the FAIR Expertise Hub,[8] funded by the PDI-SSH grant.

Acknowledgements

The authors appreciate help from colleagues, especially Mark Bruyneel, Stephanie van de Sandt, Ronald Siebes, Jacco van Ossenbruggen, Rick Mourits, and Tobias Kuhn.

References

[1] E. Schultes *et al.*, "Reusable FAIR Implementation Profiles as accelerators of FAIR Convergence," in *Advances in Conceptual Modeling*, G. Grossmann and S. Ram, Eds., Cham: Springer International Publishing, 2020, pp. 138–147, ISBN: 978-3-030-65847-2. DOI: 10.1007/978-3-030-65847-2_13.

[2] S. Wang *et al.*, "FAIR Implementation Profiles for Social Science," in *Proceeding of 17th Int Conf on Metadata and Semantics Research*, Springer, 2023.

[3] K. Hettne *et al.*, "FIP2DMP: Linking data management plans with fair implementation profiles," *FAIR Connect*, vol. 1, pp. 23–27, Jan. 2023. DOI: 10.3233/FC-221515.

[4] S. Wang and N. K. Singh, *Aligning Data Management Plans with Community Standards using FAIR Implementation Profiles*, Zenodo, Dec. 2023. DOI: 10.5281/zenodo.10286752.

[6] The dataset and related documents are with DOI 10.5281/zenodo.10418879.
[7] DOI 10.5281/zenodo.10285647.
[8] https://fairexpertisehub.org/

[5] S. Wang *et al.*, "Towards rigorous data upcycling using the fair implementation profile for the sshoc-nl socio-economic history community," en, 2023, Presented at the Criteria and Methods for Upcycling Data Collections in Social and Economic History Online Workshop (legacy4reuse). [Online]. Available: https://shuai.ai/static/files/paper/upcycling.pdf.

[6] W. Scheltjens, "Upcycling historical data collections : A paradigm for digital history?" *Journal of documentation*, pp. 1–31, 2023, ISSN: 0022-0418. DOI: 10.1108/JD-12-2022-0271.