

FAIR Implementation Profiles for the Social Sciences and Their Use Cases

Shuai Wang^{1,*}, Angelica Maineri², Navroop K. Singh¹, and Tycho Hofstra¹

¹Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

²Erasmus University Rotterdam, Rotterdam, the Netherlands

*Correspondence: Shuai Wang, shuai.wang@vu.nl

Abstract. FAIR Implementation Profiles (FIPs) are created and published to capture decisions made by communities on data curation and management. Despite the fact that many FIPs were created in domains such as medicine and 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

Convergence in FAIR implementation is hindered by heterogeneity in choices across researchers, projects, communities, and infrastructures. A FAIR Implementation Profile (FIP) consists of a structured set of questions and answers about communities' decisions about the use of FAIR Enabling Resources (FERs) [1]. FERs include repositories, identifier services, registries, knowledge representation languages, licenses for data and metadata, etc. For each question, FERs of a specific type are expected as answers (e.g. a registry for F4, a metadata schema for F2, structured vocabularies for I2) [1]. FIPs can be published as nanopublications with persistent identifiers (the nanopublication URI, and the PID of the object), the object itself (the content of the assertions), metadata (provenance metadata information), and other related publication information. Thus, FIPs can be considered as FDO implementation if augmented with the FDO Framework aligned statements [2]. More detailed analysis and comparison between nanopublications and FDOs have been proposed by E. A. Schultes et al [3]. Members of the FIPs and Practice Group of the FAIR Digital Objects Forum (<https://fairdo.org/wg/fdo-fipp/>) contribute to the enhancement of the FIPs and FERs and their representation as FDOs as defined in the FDO Framework. Many FIPs have been created but only a few are about social sciences communities. The social sciences offer interesting cases since they have a long-standing tradition of data sharing but adopt heterogeneous standards and resources in doing so. In this paper, we

provide a summary of recent advances in FIPs 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 and make data more FAIR in multiple ways.

Recently, six FIPs in social sciences were compared [4]. Among them, three FIPs were reused or improved from existing attempts, and published using the FIP Wizard (<https://fip-wizard.ds-wizard.org/>). 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 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) (<https://odisseei-data.nl/en/media-content-analysis-lab/>) 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 LGBTQ+ thesauri and controlled vocabularies of different languages. More specifically, members of the community use Homosaurus (<https://homosaurus.org/>) and related resources for indexing digital records, literature, and heritage objects related to the topic of LGBTQ+.

2. Use Cases

2.1 Use Case 1: Comparison of FIPs Between Communities

One can study how communities differ from each other by examining differences in declared FERs across FIPs [4], [5]. A summary of the overlapping FERs between communities can be made in the form of a table, the so-called convergence matrix [5]. In Wang et al. (2024) [4] such a table shows, for instance, that the SEH and LGBTQVoc communities share some FERs and are closer together than ESS and AUSSI-ESS. The results of the analysis help facilitate discussion over FAIR implementation choices within and across communities, and provide practical recommendations to strengthen connections across data from different sources.

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 [6]). A recent effort is the development of a customized interface with the related information extracted from FIPs as suggestions [7]. The study aimed to understand how researchers can take such suggestions into account when writing DMPs. 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 provide suggestions for the reuse of legacy data (i.e. data upcycling) in the SSHOC-NL Socio-Economic History Community

[8]. 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. However, such suggestions can be outdated if the FIPs used are outdated or new implementation decisions have not yet been captured by the FIPs.

3. Conclusion and Future Work

This extended abstract provides a summary of some published FIPs in social sciences. Three recent use cases are included to demonstrate how these FIPs could be used in practice. 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 in the future. The methodology of integration of FIP information into the legacy data upcycling workflow can be further explored [9]. Publishing FIPs (and FERs) with augmented statements aligned with the FDO Framework as FDOs could be a future work. Comparing a selected FIP with the result of the FAIR assessment of the 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 with DOI [10.5281/zenodo.10418879](https://doi.org/10.5281/zenodo.10418879). The knowledge model corresponding to the DMP template of the Vrije Universiteit Amsterdam, the survey, and all related documents are also available on Zenodo with DOI [10.5281/zenodo.10285647](https://doi.org/10.5281/zenodo.10285647)

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 (<https://fairexpertisehub.org/>), which was funded by the PDI-SSH grant.

Statement of Competing Interest

There is no competing interest, as far as the authors are aware.

Acknowledgments

The authors appreciate the help from colleagues including Mark Bruyneel, Stephanie van de Sandt, Ronald Siebes, Jacco van Ossenbruggen, Rick Mourits, Tobias Kuhn, and Barbara Magagna. The suggestions from the reviewers are highly appreciated.

References

- [1] B. Magagna, E. Schultes, and T. Kuhn, *FAIR Implementation Profile (FIP) ontology*, 2020. [Online]. Available: <https://w3id.org/fair/fip>.
- [2] L. O. Bonino da Silva Santos, T. P. Sales, C. M. Fonseca, and G. Guizzardi, "Towards a conceptual model for the fair digital object framework," in *Formal Ontology in Information Systems*. IOS Press, Dec. 2023, ISBN: 9781643684697. DOI: [10.3233/faia231131](https://doi.org/10.3233/faia231131). [Online]. Available: <http://dx.doi.org/10.3233/FAIA231131>.
- [3] E. A. Schultes et al., "The comparative anatomy of nanopublications and fair digital objects," *Research Ideas and Outcomes*, vol. 8, e94150, 2022. DOI: [10.3897/rio.8.e94150](https://doi.org/10.3897/rio.8.e94150). eprint: <https://doi.org/10.3897/rio.8.e94150>. [Online]. Available: <https://doi.org/10.3897/rio.8.e94150>.
- [4] S. Wang, A. Maineri, N. K. Singh, and T. Kuhn, "FAIR Implementation Profiles for Social Science," in *Metadata and Semantic Research*, E. Garoufallou and F. Sartori, Eds., DOI: 10.1007/978-3-031-65990-4_26, Cham: Springer Nature Switzerland, 2024, pp. 284–290, ISBN: 978-3-031-65990-4.
- [5] 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](https://doi.org/10.1007/978-3-030-65847-2_13).
- [6] 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](https://doi.org/10.3233/FC-221515).
- [7] S. Wang et al., *Aligning Data Management Plans with Community Standards using FAIR Implementation Profiles*, the preprint of the paper to be presented at the 18th International Conference on Metadata and Semantics Research (MTRS) conference is available at <https://shuai.ai/static/files/paper/mtsr-fip2dmp.pdf>, Zenodo, Apr. 2024. DOI: [10.5281/zenodo.10927435](https://doi.org/10.5281/zenodo.10927435). [Online]. Available: <https://doi.org/10.5281/zenodo.10927435>.
- [8] 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>.
- [9] 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](https://doi.org/10.1108/JD-12-2022-0271).