

12 March 2020

The findings from the FAIRsFAIR investigation into persistent identifier usage and semantic interoperability across data infrastructures Europe-wide (D2.1 Report on FAIR requirements for persistence and interoperability 2019) reveal a multiplicity of technical solutions and wide variation both within and between scientific domains. The report, based on a review of projects and landmarks listed by the European Strategy Forum on Research Infrastructures (ESFRI) and various Research Data Alliance (RDA) groups, is now available for perusal and comment.

Research Methodology

Over a period of months and with the aim of describing the overall data management landscape and identifying commonalities and gaps regarding standards for and implementation of semantic interoperability, vocabularies and ontologies, metadata, and persistent identifiers, the FAIRsFAIR

© Copyright 2020 – FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558. The content of this document does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of such content.

team working on FAIR Practices: Semantics, Interoperability, and Services gathered information through a combination of desk research, survey data, and interviews with selected digital infrastructures.

Key sources of information for the desk research were:

- Infrastructures identified during the course of previous landscaping efforts and overviews
- FAIRsharing and the documentation of metadata and persistent identifiers
- Outputs from RDA groups and their associates, for example CODATA

To complement and validate information from the desk research, a survey aimed at data managers and data support experts was developed and administered to respondents in domains ranging from energy, environment, health, engineering, social sciences and computing to the humanities.

Then, during September 2019 five semi-structured expert interviews were conducted with domain-agnostic or generic realities namely **DataCite, Deutsches Klimarechenzentrum (DKRZ), DiSSCo, FAIRsharing and Figshare**. The aim of the interviews was to get a better understanding of:

- Practical implementations of semantic interoperability across infrastructures
- Perceived critical success factors for FAIR and semantic interoperability
- The most serious omissions in currently available tools and specifications

Key Findings

- **FAIRness is not currently clearly defined as a concept** and FAIR vocabularies, software, and services are largely undefined.
- **The landscape is diverse in all aspects.** Differences inside domains are often bigger than differences between domains. Refinement and application of the FAIR principles should be driven by research rather than technology to achieve the needed usability and the potentially huge benefits of FAIR data. While standardisation across domains will not solve all problems given the differing needs of each domain, community adoption and trust are critical success factors within domains.
- **Semantic artefacts are a key element in building interoperability and good quality (meta)data.** Given that needs and maturity levels differ across infrastructures and domains, and shared resources are a necessity, both local management and overall governance are critical. In particular, local data management services need to be involved in both reuse of reference metadata and enabling local modifications.
- Crosswalks, mappings and semantic application profiles should be published and registered in **machine readable formats.**
- **PID and data type registries should promote reuse rather than bulk creation of PIDs.** To support interoperability, they should be considered semantic artefacts and used mindfully.
- **Reuse of semantic artefacts should be promoted** by publishing application profiles. This should happen in machine readable formats in shared registries. Curated registries such as the EOSC Hub, FAIRsharing and re3data.org are important resources for promoting implementations of the FAIR data principles.
- **Data citation and machine actionable solutions should be developed in parallel.**
- The most popular, potentially most useful, and most complex approaches to improving FAIRness of data are based on technologies using Linked Data, the nature of which simultaneously encourages experimentation and hinders wide adoption. This situation should

© Copyright 2020 – FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558. The content of this document does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of such content.

improve as technology advances.

The FAIRsFAIR Report on FAIR requirements for persistence and interoperability 2019 is the first of three such reports to be published during the life of the project. Additional research undertaken by FAIRsFAIR during the same period – March to November 2019 assessed the current state of FAIR data policy and FAIR data practice in Europe.

Your Feedback Invited

In the light of the many landscaping, specification and "FAIRification" activities ongoing within the EOSC projects and elsewhere, and with a view to developing recommendations to be tabled in further reports, the authors invite you to peruse the document and enrich the content by offering your comments and suggestions **before 17 April 2020.**

You can also browse the meeting notes from the **associated webinar organised on 11 February 2020.**

© Copyright 2020 – FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558. The content of this document does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of such content.