Future of Scholarly Communication

Forging an inclusive and innovative research infrastructure for scholarly communication in the social sciences and humanities

CHAPTER

03

The road to FAIR Social Sciences and Humanities

Authors:
Elena Giglia, Karla Avanço, Arnaud Gingold
Contributors:
Internal review: Marina Angelaki, Milica Ševkušić
Partners involved:
University of Turin, OpenEdition, IBL PAN
Partners involved (workshops):
University of Coimbra, DARIAH, University of Göttingen

Introduction

The FAIR principles (Findability, Accessibility, Interoperability, and Reusability) are a set of foundational guidelines aimed at improving the management of digital scholarly resources for both humans and machines. In considering digital objects as a whole, focusing on data management and reuse, and allowing for cross-disciplinary research, the FAIR principles provide an innovative approach for Social Sciences and Humanities' (SSH) practices. More fragmented, less data-centric, and in some fields, dealing with physical objects, the SSH environment still needs some guidance in exploring the full potential of FAIR principles. As a major component for integrating the European Open Science Cloud (EOSC), FAIR principles are also, more broadly, an important tool for open science.

Therefore, OPERAS-P Task 6.3 was aimed at providing the knowledge base together with the views of the community to find the most suitable way to make FAIRification of SSH data possible, using a threefold approach:

- **Speak the same language**

The first step was to focus on the different kinds of data in SSH and the issues arising from implementing FAIR principles via a thorough review of the rich and growing literature about data in SSH.

- **Work with the community**

Through focus groups and workshops, the task engaged stakeholders in unveiling perceptions about FAIR data, the needs of the different communities, and the challenges facing FAIRification regarding various disciplines.

- **Showing the road ahead**

Based on the review and the workshops, the task suggested some directions for OPERAS concerning the FAIRification of data, including FAIRification tool prototypes and measures to further engage the community both in discussions and implementation.

Task 6.3 worked in synergy with parallel activities carried out by the CO-OPERAS Implementation Network (IN) within the GOFAIR initiative, which was aimed at the FAIRification of SSH data and publications. CO-OPERAS IN is coordinated by OpenEdition and UniTo, with the regular contributions from Huma-Num.

Main Findings

Capturing the SSH in transition

FAIR principles, which are essentially digital, mostly data-centric, and oriented towards automated processes, are to a certain point an adequate device for capturing current SSH research practices. The levels of awareness, skills, and engagement concerning FAIR principles characterise, beyond disciplines, distinct communities and reveal the transitional period of the SSH global landscape. To avoid falling into mere opacity, the agnosticism of generic principles like FAIR requires the various communities' specificities to be taken into account.

Everything is data, or could be

Within the FAIR framework, the concept of data is intended to be as universal as possible, including datasets, publications, software, etc. Although accepted in various SSH fields – such as social sciences, history, linguistics, and digital humanities – the notion of data continues to be regularly discussed in the SSH context. The literature review and the research community workshops have also shown ways of integrating SSH's various data types and scientific methods into a coherent digital ecosystem. For instance, it seems possible to make a broad distinction between source and result data, or to consider the well-established processes of resources' curation and the management of the social sciences and humanities as FAIR-enabling practices.

Advocating for FAIR adoption means explaining its benefits

From previous findings, it appears that advocacy for FAIR principles remains the first step when considering the SSH landscape. In order to expand the adoption of FAIR principles within the SSH, it is the principles' final purpose that should be outlined, showing how they increase the quality of research and can integrate even convergent, although pre-existing, practices.

Coordination is key for FAIRification

The various workshops and communications made it obvious that the FAIRification of SSH implied a wide diversity of actors. All the actors involved in the data generation process should also be actors of the "FAIRification-chain": researchers, data stewards, repository managers, librarians, and publishers. Regarding the direct actors of data generation, there is, more specifically, a need to converge on metadata standards, potentially by sharing a minimal metadata set. Coordination at a broader level is also required to ensure a consistent FAIR ecosystem for the SSH. With that prospect, OPERAS intends to collaborate with SSH ERICs Cessda, CLARIN, and DARIAH as well as with projects like SSHOC and EOSC-Pillar.

FAIR stories, examples, and tools need more visibility

SSH researchers feel that there is a lack of recognition when their research relies concretely on implementations of FAIR principles. This poor reward system for researchers making their data actually findable, accessible, interoperable, and reusable obviously hinders the research community's engagement with the adoption of FAIR principles. Moreover, this deprives the community

of examples and stories that they could take inspiration from. Through CO-OPERAS IN, OPERAS carried out initiatives to address this issue. First, a blog is currently being established that will offer a common space for discussion and experience sharing. Second, in order to provide guidance for the FAIRification of publications, a major object of SSH research, on-going work will provide a FAIRification toolkit dedicated to publishing platforms.

Challenges

FAIR awareness is still low

General knowledge regarding specific but major aspects of FAIR, such as persistent identifiers, metadata interoperability standards, and open licenses, is very uneven in the SSH environment, if not simply lacking. Given that the acronym is most certainly becoming more widely known, it is obvious that FAIRification requires more than such a superficial understanding. Advocacy efforts and dedicated training are therefore required to increase the level and the quality of FAIR awareness. However, this may also rely on an effort of "translation," adapting the FAIR analytical grid to well-established and functional practices that are convergent with, if not identical to, FAIR principles.

Diversity and complexity of the SSH

As mentioned above, the SSH landscape does not offer a coherent or uniform landscape. However, rather than "fragmentation," we should simply speak, in this case, of the diversity and complexity of the SSH research environment. When it comes to data, the diversity increases immediately because of the various typologies and methods involved. It implies a slight adjustment of the objective: instead of bringing all the SSH communities into a single (and impossible) model, we should look for similarities that could work as hooks – able to connect together all the different parts of a rich and lively environment.

FAIR is not open, but it supports open science

This is a twofold challenge. FAIR is regularly described as distinct from openness; just as FAIR has regularly become the companion of open science policies. The statement "open as possible, closed as necessary," although handy for general presentations, offers unfortunately poor guidance for concrete FAIR implementations. In fact, open licensing allows for reuse in degrees that are not entirely described by the pair "open/closed." This represents the first challenge in terms of communication. Moreover, dealing with humans' creations and phenomena, SSH sources and outputs often include personal information, property rights, or even sensitive data. Reusability, as outlined by SSH researchers, is thus characterized by legal challenges that require specific guidance and expertise.

Incentives and rewards

Another challenge already mentioned concerns the global reward system of contemporary research. Incentives to adopt FAIR principles and rewards for FAIR implementations can only partially rely on research networks and infrastructures. The FAIRness level and quality of research should be part of funders and policymakers' assessment processes.

Recommendations

Target audience	Recommendation
Research performing organisations 	**Consider the research communities' needs** in order to align implementation policies with research practices and purposes. **Preserve the domain specificities** regarding data, digital objects, methods, etc., and adapt FAIR implementations accordingly. **Preserve the multilingualism and bibliodiversity** of the SSH environment. **Address the sustainability of FAIR** services to ensure data reuse in the longer term.
Scholars 	**Collaborate on common minimal metadata** sets, allowing for cross-disciplinary research and the building of FAIR digital objects.
Scholars and research infrastructures 	**Produce an inventory of existing FAIRification tools**, enriched with new tools dedicated to SSH specific objects, such as publications or cultural heritage materials.
Funders and Policy makers 	**Acknowledge and reward** FAIRification of data and publications, making them part of the research assessment.
Publishers 	**Use the FAIR principles as a grid** to assess the FAIRness of publishing systems, to enhance both content visibility and the quality of publishers' information systems.
OPERAS 	**Expand and improve advocacy** by offering both explanations of FAIR principles and examples of FAIR tools and services. **Provide training on FAIR data and metadata** that takes into account the disciplines, data types, and the existing standards' specificities. **Coordinate with other infrastructures and projects** involved in FAIRification to offer consistent guidance to their respective communities.

Data and further reading

- The full report from this task is available at OPERAS Innovation Lab community on Zenodo.
- Reports from the national workshops on Definition of Data for FAIR SSH (in chronological order): Porto (international), Turin, Coimbra, Göttingen, Paris, Brussels (international).
- T. Biro, E. Giglia, "Humanities and Data: for a community-driven path towards FAIRness", March 2020, recording from the Berlin OpenScience conference.

FAIR data in the SSH

Main findings

- A snapshot of the SSH transitional period
- Everything is data, or could be
- Advocating for FAIR adoption means explaining its benefits
- For FAIRification, coordination is key
- FAIR stories, examples and tools need more visibility

Challenges

- FAIR awareness is still low
- Diversity and complexity of the SSH
- FAIR is not open but supports open science
- Incentives and rewards

Recommendations

RESEARCH PERFORMING ORGANISATIONS

- Consider the research communities' needs to align implementation policies with research practices
- Preserve the domain specificities and adopt FAIR
- Preserve multilingualism and bibliodiversity
- Address the sustainability of their FAIR services
- Produce an inventory of existing SSH FAIRification tools

SCHOLARS

- Collaborate on common minimal metadata sets allowing for cross-disciplinary research and the building of FAIR digital objects

FUNDERS

- Acknowledge and reward FAIRification practices

PUBLISHERS

- Use the FAIR principles as a grid to assess publication FAIRness

OPERAS

- Improve advocacy and enhance training for FAIR adapted to SSH
- Coordinate with other infrastructures and projects involved in FAIRification to offer consistent guidance