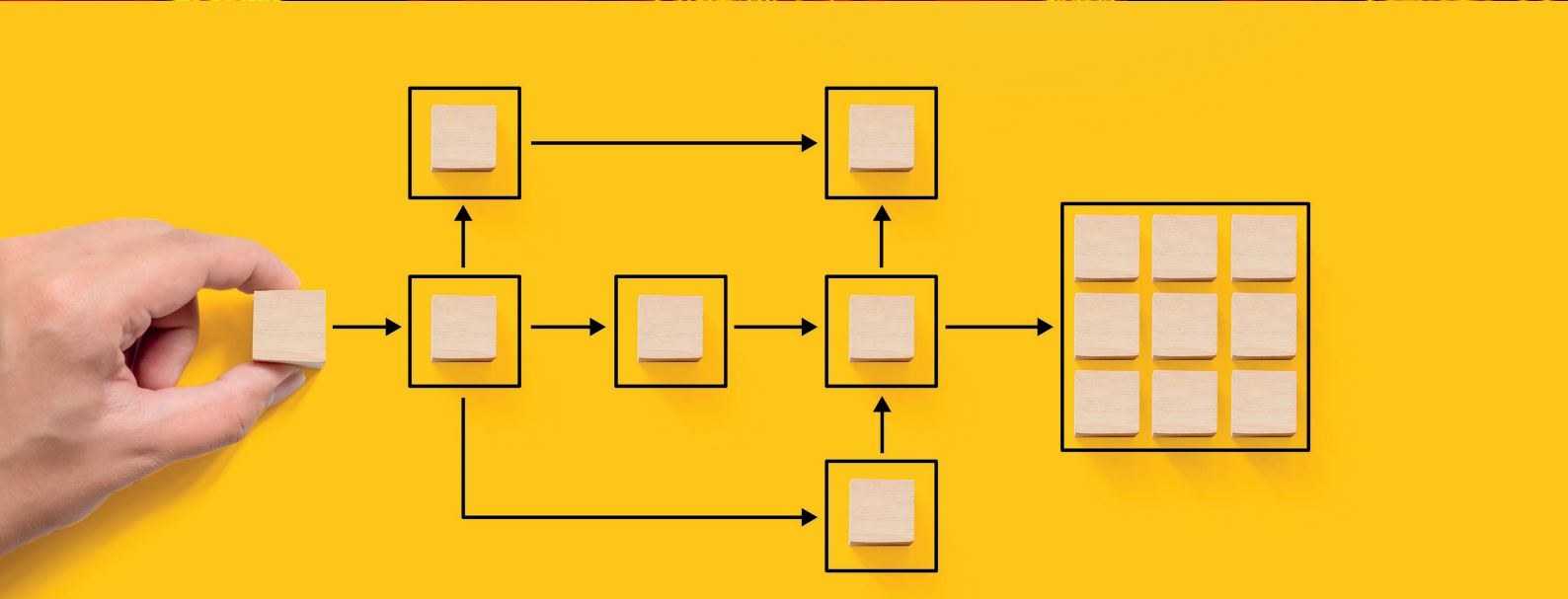
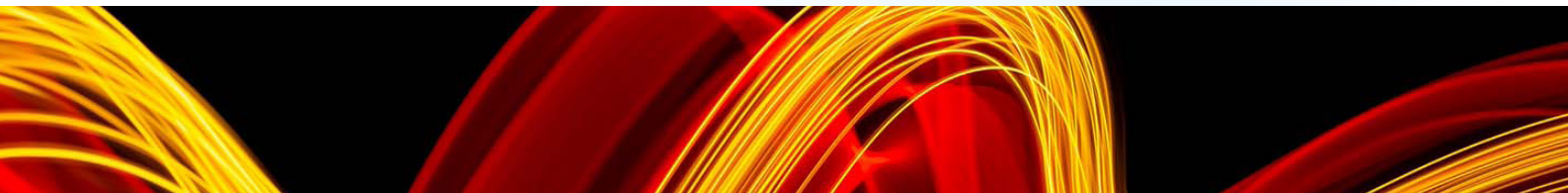


Applying FAIR Signposting and RO-Crate at NFDI4Health



Authors names, affiliations and ORCIDs:

- Matthias Löbe, Institute for Medical Informatics (IMISE), University of Leipzig, <https://orcid.org/0000-0002-2344-0426>
- Daniel Turner, Digital Curation Centre (DCC), <https://orcid.org/0000-0003-2216-8181>

Support action:

FAIR Signposting and RO-Crate. During the 3-month support action, participants took part in three virtual workshops to implement FAIR Signposting and RO-Crate to improve the discovery and consumption of their content and metadata. Participants benefited from interacting with mentors representing FAIR Signposting and RO-Crate.

Keywords:

FAIR Signposting, RO-Crate, NFDI4Health, FAIR Digital Objects

Summary:

Team members working on behalf of the national infrastructure for research data in Germany, NFDI4Health, explored ways in which FAIR Signposting and RO-Crate could be applied within the health domain. Work was concentrated on their own research data management system, FAIRDOME SEEK, which is especially useful for those with limited resources.



■ Introduction

The team of participants represented the NFDI4Health research infrastructure for personal health data within the consortium for the German National Research Data Infrastructure. Their objective is the “amalgamation of epidemiological, public health and clinical research.”¹ Within the consortium the research infrastructures collaborate toward ensuring FAIR-enabling practices, and the team worked together with other partners working for the medical informatics initiative and participating in the FAIR Assessment action.

FAIR Signposting was new to the team, while RO-Crate had arisen previously in FAIR Digital Objects discussions with Peter Wittenburg, who is also active in the national research data infrastructure. The support action was an opportunity to give this line of work higher priority, making time to look into the documentation and what could be done with it.

■ Approach taken

Early on, the generic approach of the support action would only fit basic needs, missing problems specific to the health domain but still offering valuable insights into the two technologies. The team also wanted to take a practical approach. Armed with a variety of tools and services, it was decided to concentrate on their own research data management system, FAIRDOME. Only one developer is working on the system and participated in the support action, pointing out that others in Manchester were also implementing the technologies in a similar fashion. This prompted further discussion between the two parties, bringing in those who were not part of the FAIR-IMPACT project. The external group confirmed that development was underway and this was part of their roadmap. Although not activated by default, the team already have signposting implemented and showed the NFDI4Health team how to activate it. Work is also underway on RO-Crate, but it was not part of the major fork and thus not available to the team participating in the support action. Given the differences between bioinformatics and medical informatics, the RO-Crate implementation might be of use following slight modifications.

■ Challenges encountered and addressed

There were great challenges, but these were not expected to impact the current project because an existing solution was being extended. The problems were nothing new, and no additional issues seem to arise from these implementations.

■ Impact

The team read the specifications, understood the aims of the methods, and saw how developers are implementing them. But to use this in medical informatics, the team would probably need some kind of graphical user interface and some harmonised ontological terminology that allows specifying certain things, since signposting and RO-Crate are infrastructure methods that enable machines or agents to read certain parts of a website or data repository records and extract basic metadata. The team explained RO-Crates to others as an intelligent archive format, like a zip

1 <https://www.nfdi4health.de/en/about-us.html>



file where they can have a collection of their research assets as a minimal collection of documents, datasets, and maybe scripts or models that helps understand what the dataset is about and could also be used as a supplement of a publication.

To be able to implement an automatic process to create RO-Crates, the team would need to advance first in the standardisation of the research data outputs. Currently these are different types of excel tables and not self-descriptive. The team could create a container with a version of the data in another format, provenance information describing how the data has been generated, or annotations about some concepts from an ontology that are not part of the paper. But to do it automatically, the team would first need to define a convention of folders and formats, and it would also require more interaction with the development team.

■ ■ Future plans

Progress is not so fast in large infrastructures and if they have decided on one way of doing things it takes years to change that. It is difficult to persuade people because of potential risks when things get changed and resources required. A lot of people do not really believe in automatic metadata. They say people can just go into our website and read the PDFs about the data and decide if it is interesting. However, showing what other people are doing in other European projects and what approaches are popular can sometimes be a more persuasive argument than demonstrating why it is useful. It is being investigated whether a vocabulary and a minimal data model can be specified as an extension of Schema.org, which can be used to structure (a collection of) assets to describe the results of a clinical study.

■ ■ Key message

Those that collect data tend not to have much incentive to share it and thus lose enthusiasm for new technologies and additional efforts. Automation might improve this outlook, such as providing a simple checkbox that would mark files as part of a research archive package.





@fairimpact_eu



company/fair-impact-eu-project/



fair-impact.eu

