

Making fair data a reality for photons and neutrons

Andy Götz, the coordinator of PaNOSC for ESRF, looks at FAIR data – an initiative to deliver research data to the scientific community to use

The scientific world is undergoing a major revolution with the widespread adoption of the so-called FAIR principles for research data. FAIR stands for Findable, Accessible, Interoperable and Reusable and was first published in a paper in Nature in 2016⁽¹⁾. The catchy mnemonic was originally proposed at a Lorentz Workshop in Leiden, The Netherlands, in January 2014, named: Jointly designing a Data FAIRport. In a nutshell, the FAIR principles were proposed to ensure research data are made available to the scientific community in such a way that they can be found, downloaded, understood and reused.

The goal is to make data used in scientific publications available to the community so they can verify the results and reproduce them and eventually derive new results from them. Applying the FAIR principles systematically to research data will address the reproducibility and replicability crisis in science and make scientific data used beyond their original purpose.

How do we make data FAIR?

The major European, international, and national scientific organisations, funders and governments have recognised the huge potential of making FAIR data the new norm in science and have adopted them. The leader of the pack has been the European Commission who have published a report and action plan on Turning FAIR data into reality in 2018⁽²⁾. The report was motivated inter-alia by the results of a study on the cost of not having FAIR data which estimated the total annual cost at roughly 10 billion euros, half of which is due to time lost when trying to analyse data.

FAIR data is one of the motivations and cornerstones for the ambitious European Open Science Cloud (EOSC). The EOSC main goal is to provide FAIR data for all publicly funded research and help scientists adopt best practices for Open Science. Open Science is the new way of doing science based on Open i.e., FAIR Data, Open Source Software, Open Access Publications, Open Hardware, and Open Review, in short making all steps required for doing science open. The Open Science movement has been endorsed by all major scientific learned societies, Universities, Governments and cultural organisations like UNESCO.

Making all steps required for doing science open

To include Photon and Neutrons in the game, Europe and the UK are lucky to have a rich set of so-called photon and neutron sources made up of synchrotrons, free electron lasers, research reactors, spallation sources etc.

These exotic-sounding devices are large expensive facilities which provide unique tools to study matter in all its states, like giant microscopes. The PaN facilities produce petabytes of data for applied research the majority of which are financed by taxpayer's money. The data can be of very high value see for example <https://human-organ-atlas.esrf.eu> an atlas of human organs down to the cellular level.

But they can also be of very limited value depending on the quality of the samples.

FAIR data can help citations and publication reusability

In both these cases making data FAIR brings advantages to the scientists who produced the data saving them time by having higher quality data which they can cite in their publications afterwards and which can eventually be reused. The EOSC offers a perfect opportunity for the PaN facilities to make FAIR data reality for the PaN community.

Consequently, two major projects financed by the EC as part of the EOSC H2020 calls have been working on FAIR data for 4 years. The 2 projects, PaNOSC⁽³⁾ with 6 ESFRI and ERIC partners and ExPaNDS⁽⁴⁾ with 10 national Research Institutes (RIs), have just recently been completed (refer to the websites for more information). Making data FAIR means updating data policies, metadata frameworks, data management plans, persistent identifiers, data catalogues, data formats, data analysis, and training.

The two projects worked together on all these topics to make them FAIR. They also participated in many events around FAIR organised with the scientific and EOSC community. The FAIRness of the outcomes was self-evaluated. All 16 PaN facilities involved made huge progress towards FAIR in adopting the outcomes of the projects.

FAIR data has had a number of impacts on the PaN community already

The most direct impact in making data FAIR has been on improving the quality of the metadata and data management. This has a direct impact on the originators of the data by saving them time in their data analysis. The second impact is for the scientific community at large. A common data portal with a federated search over all PaN facilities allows searching for Open Data (see <https://data.panosc.eu>). Close collaboration with the International Union of Crystallography (IUCr) on a new journal for raw data (Raw Data Letters) will ensure the PaN community can contribute FAIR data to the journal. Over 500 datasets have been downloaded from the Human Organ Atlas.

This is only the beginning of the journey of the PaN community towards FAIR data, EOSC and Open Science. The next steps will be carried out in collaboration under the LEAPS and LENS initiatives, partly sponsored by EOSC projects but mostly sustained by the facilities themselves as FAIR data becomes the new norm in science. The FAIRness of

data will increase as more data are made open. Multiple data spaces will be created for high-impact research data in other areas. Some of these are already identified e.g. life sciences, materials science, palaeontology, batteries, quantum materials, etc.

References

1. Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016).
<https://doi.org/10.1038/sdata.2016.18>
2. European Commission, Directorate-General for Research and Innovation, Turning FAIR into reality : final report and action plan from the European Commission expert group on FAIR data, Publications Office, 2018,
<https://data.europa.eu/doi/10.2777/1524>
3. <https://panosc.eu> 4. <https://expands.eu>

Please Note: This is a Commercial Profile



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).