

FAIR data to accelerate scientific discovery at national scattering facilities

Thomas Proffen^a

^a Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830,
tproffen@ornl.gov

Rapid developments in experimental instrumentation and detectors allow researchers to collect experimental data in unprecedented amounts. At the same time advanced in computing, data analytics and machine learning open the door to accelerating scientific discovery from experimental data and offer the possibility to ‘learn’ from large amounts of experimental data available in the user community. However, it is still common practice that experiments, data analysis and modeling are confined to a small team of collaborating researchers and results are published in the form of scientific papers. One notable exception is crystallographic structural results being published and shared via databases such as the PDB or CSD.

Using the full potential of experimental data for scientific discovery requires data to be FAIR [1]: Findable, Accessible, Interoperable and Reusable. In this contribution, these four aspects will be discussed from the point of view of Synchrotron and Neutron User Facilities. Both individual researchers in need of individual data sets to e.g. validate a theoretical model as well as the needs required for large scale data mining, data analytics and machine learning will be discussed.

References

[1] Wilkinson et al, *Scientific Data* **3**, 160018 (2016) <https://doi.org/10.1038/sdata.2016.18>