



FAIR and open data: state of affairs for seismological networks and infrastructures globally

Helle Pedersen¹, Elizabetta D'Anastasio², Jerry Carter³, Rob Casey³, Jonathan Hanson², Florian Haslinger⁴, Javier Quinteros⁵, and Lesley Wyborn⁶

¹University of Grenoble and CNRS-INSU, ISTERRE, Observatoire de Grenoble, Grenoble Cedex, France (helle.pedersen@univ-grenoble-alpes.fr)

²GNS Science, New Zealand

³Earthscope Consortium, USA

⁴Swiss Seismological Service (SED) at ETH, Zurich, Switzerland

⁵Deutsches GeoForschungsZentrum (GFZ), Potsdam, Germany

⁶National Computational Infrastructure, Australian National University, Australia

Over the past years the awareness of the importance and usefulness of Globally Unique Persistent Resolvable identifiers (GUPRIs), appropriate licenses, further standardisation of metadata, and general adherence to the FAIR principles has increased significantly in the international seismological community. One important milestone was the introduction of seismological network identifiers as an FDSN recommendation in 2014 and recently updated at the end of 2023. Further advances were made, or are in development, in connection with the establishment of more formalised research infrastructures like EPOS in Europe, Auscope in Australia, and the reorganisation of IRIS and UNAVCO to Earthscope in the U.S, as well as national and international initiatives such as for example COPDESS, and RDA. In Europe the developments in seismology have taken place within or with close links to projects such as Geo-INQUIRE, ChEESE, Digital Twin of GEophysical extremes (DT-GEO), and building on achievements and tools from projects (e.g. FAIRsFAIR), and in general as part of the European Research Infrastructure environment.

In this contribution we reflect on the current state of the use of identifiers, application of licenses and other improvements in the FAIRness of seismological data, products and services, focusing on FDSN and ORFEUS/EIDA, Earthscope, Auscope, and the GFZ/Geofon and RESIF data centers.

Experience gained with DOIs as seismological network identifiers is conclusive in terms of acceptability of associating a DOI at network level: more than 70% of FDSN registered networks now have a DOI. On the contrary, the correct citation based on the DOI is only gradually gaining traction in scientific publications, due to a combination of slow uptake by researchers and the difficulty of the scientific journals to set up automatic or semi-automatic checking procedures. Additional challenges remain e.g. when trying to implement identifiers for data collections and/or downstream products that properly support reproducibility of scientific workflows. A simple collection of DOIs would not be enough to describe a user defined dataset, that is characterized by a much finer granularity. Therefore, the evaluation of other alternatives, like the inclusion of time-

stamped query related to a new DOI describing such a dataset could be needed. This could be used for small datasets with data from different sources, or even for ML/AI training sets, defined as a collection of networks.

Some standardisation and best practices have emerged with regard to licensing of seismological data and products, in particular the use of by attribution licenses like CC-BY. A common and harmonised understanding of legal implications, intellectual property, and consequences of specific licenses, however, seems still quite a bit away.

Implementing FAIRness, and then measuring it or even reporting the level of FAIRness to funding agencies has met with some success at least in specific initiatives or through specific projects. One noteworthy development is the introduction of FAIR Implementation Profiles (FIPS) that allow a quantitative assessment of the achieved FAIRness.