



Latest developments at GEOFON Data Centre

Javier Quinteros, Susanne Hemmleb, Peter Evans, Andres Heinloo, and Angelo Strollo
Deutsches GeoForschungsZentrum, 2.4 Seismology, Potsdam, Germany (geofon@gfz-potsdam.de)

Data management at GEOFON has evolved in the last decade as our focus has shifted from the main challenge of acquiring high quality seismic data in standard formats towards the development of services which allow the automatic execution of complex workflows on big volumes of data.

Nowadays, the scientific community needs access to as much data as possible in the easiest possible way and thus a big part of our efforts of the last years was focused on building a federation, which provides a uniform interface across the different data centers, so that users are freed from which data center holds a particular station's data. This is the spirit behind European Integrated Data Archive (EIDA), ORFEUS's distributed data center, in which we have played an important role since its inception. Moreover, earth observations are becoming more and more integrated, and there is a need of high-quality general purpose services rather than specific solutions for each discipline. In recent years we have evolved our offering towards standard web services. This will open our data to new cross-discipline users and to citizen science. Researchers may also want to easily stage their data into computational facilities, produce interesting results, share and publish them ensuring that results are also citable and reproducible.

Any type of data archived at GEOFON follows clear and predefined automatic workflows, which ensure not only its preservation and replication, but also discoverability and proper citation/attribution. Our solution is based on state-of-the-art technology (e.g. Moore, 2008), developed and deployed at the most important research data centers in the world (f.i. RENCI, CERN). Advanced Authentication technologies using European initiatives like eduGAIN are also being implemented at GEOFON for embargoed data.

We've been one of the driving forces to establish the attribution of a DOI (Digital Object Identifiers) to every seismic network as an international standard, so that data can be properly cited (Evans et al, 2015).

GEOFON assigns Persistent Identifiers (PID) to every data file to have an unambiguous specification of datasets. We also evaluate solutions to support versioning of data to be able to regenerate them as they were at any point in time.

Attribution of PID will facilitate data discovery and stage of data directly into the computational facility, where these need to be processed.

A modern scientific data center is not just a place to store data, but a complex and powerful infrastructure that should support almost every step in the activities of a scientist, facilitating interdisciplinary research.

- Moore, R. Towards a Theory of Digital Preservation, International Journal of Digital Curation, 2008, Vol. 3, No. 1, pp. 63-75, doi:10.2218/ijdc.v3i1.42

- Evans, P., Strollo, A., Clark, A., Ahern, T., Newman, R., Clinton, J. F., Pedersen, H., Pequegnat, C. (2015): Why Seismic Networks Need Digital Object Identifiers. - Eos, Transactions American Geophysical Union, 96. doi:10.1029/2015EO036971