

Research Data Management in Scientific Drilling: from the acquisition of consistent data at the drill site to FAIR data dissemination

Henning Lorenz^a, Cindy Kunkel^b and Katja Heeschen^b

^aDepartment of Earth Sciences, Uppsala University, Uppsala, Sweden, henning.lorenz@geo.uu.se; GFZ German Research Centre for Geosciences, ^bHelmholtz-Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, Geomechanics and Scientific Drilling (International Continental Scientific Drilling Program, Operational Support Group), Potsdam, Germany.

Proper Research Data Management (RDM) is a key requirement for making major scientific projects work smoothly and effectively. In general, the scientific drilling community was an early adopter of RDM and making data open and citable, but with varying success. During recent years, the demands on RDM have been increasing further through, e.g., the adoption and enforcement of the FAIR data principles (Findable, Accessible, Interoperable, Reusable) by major science programmes, funding organisations and scientific journals. This positive development significantly fosters the (re-)usability of data and sample material, resulting in a higher scientific output for the invested funding. Luckily, FAIR research data management has not to be cumbersome with the appropriate tools.

The mobile Drilling Information System (mDIS): consistent data at the drill site and in the core repository

The mDIS is a data management system designed for the acquisition of data and scientific documentation at the drill site. It can be easily adapted to the individual projects' needs and is able to handle all common tasks and workflows at a drill site. Its core task is to document the sample material from the borehole in a hierarchical structure ("borehole – core – section – sample") and assign globally unique and resolvable identifiers to all items (IGSNs, see below). The mDIS development was initiated and has been significantly funded by the International Continental Scientific Drilling Program (ICDP, www.icdp-online.org), of which the majority of the Nordic Countries are members. The system is also used for curation at major core repositories.

The International Generic Sample Number (IGSN): a unique identifier for physical samples

The International Generic (formerly "Geo") Sample Number is a globally unique, persistent and resolvable identifier for physical samples (www.igsn.org). Beyond a basic, for all disciplines common data schema, extended schemata are available for various disciplines, such as for (scientific) drilling. IGSNs are created automatically by mDIS. Every little piece from a drill core that has an IGSN assigned can be resolved easily via the Handle System, and data including the full sample hierarchy up to the borehole level retrieved. IGSN is "FAIR samples" and provides an easy way to unambiguously cite samples and resolve associated data.

The European Plate Observing System (EPOS): FAIR solid Earth sciences data

EPOS is the pan-European distributed e-infrastructure for solid Earth sciences (www.epos-eric.eu). Its main purpose is making FAIR solid Earth sciences data available to the scientific community, either through its data portal or by directly interfacing the data services. To achieve pan-European FAIR data provision, metadata and data have to be harmonized both within scientific communities and across borders, and the metadata/data transfer has to follow open standards. For "boreholes", such standards were defined in and data services set up by the "Geological Information and Modelling" thematic community (TCS GIM). Scientific drilling data are part of EPOS FAIR data provision. While for practical reasons only discovery data are provided for pre-mDIS boreholes, which include links to the respective data repositories (e.g. https://www.geodata.rocks/BoreholeView/SD-DSDP_12-114), fully Linked Data services are supported by the information derived from mDIS, providing rich information directly to the user (e.g. https://www.geodata.rocks/Boreholes/SD-5054_2_A).

In summary, the combination of mDIS, IGSN and EPOS provides an appropriate and user-friendly solution for consistent data acquisition and curation, FAIR handling of drill core/sample material and FAIR (scientific) drilling data provision, with the COSC-2 (Collisional Orogeny in the Scandinavian Caledonides) ICDP project being the first example to utilise all of its components.