The CDGP Repository for Geothermal Data

Mathieu Turlure\textsuperscript{1}, Marc Schaming\textsuperscript{2}, Alice Fremand\textsuperscript{1,3}, Marc Grunberg\textsuperscript{1}, and Jean Schmittbuhl\textsuperscript{2}

\textsuperscript{1}Université de Strasbourg, CNRS, EOST-UMS830, F-67000 Strasbourg, France
\textsuperscript{2}Université de Strasbourg, CNRS, IPGS-UMR7516, F-67000 Strasbourg, France
\textsuperscript{3}Now at British Antarctic Survey, UK Polar Data Centre, Cambridge, UK

The Data Center for Deep Geothermal Energy (CDGP – Centre de Données de Géothermie Profonde, https://cdgp.u-strasbg.fr) was launched in 2016 by the LabEx G-EAU-THERMIE PROFONDE (http://labex-geothermie.unistra.fr) to preserve, archive and distribute data acquired on geothermal sites in Alsace. Since the beginning of the project, specific procedures are followed to respect international requirements for data management. In particular, FAIR recommendations are used to distribute Findable, Accessible, Interoperable and Reusable data.

Data currently available on the CDGP mainly consist of seismological and hydraulic data acquired at the Soultz-sous-Forêts geothermal plant pilot project. Data on the website are gathered in episodes. Episodes 1994, 1995, 1996, and 2010 from Soultz-sous-Forêts have been recently added to the episodes already available on the CDGP (1988, 1991, 1993, 2000, 2003, 2004 and 2005). All data are described with metadata and interoperability is promoted with use of open or community-shared data formats: SEED, csv, pdf, etc. Episodes have DOIs.

To secure Intellectual Property Rights (IPR) set by data providers that partly come from Industry, an Authentication, Authorization and Accounting Infrastructure (AAAI) grants data access depending to distribution rules and user’s affiliation (i.e. academic, industrial, ...).

The CDGP is also a local node for the European Plate Observing System (EPOS) Anthropogenic Hazards platform (https://tcs.ah-epos.eu). The platform provides an environment and facilities (data, services, software) for research onto anthropogenic hazards, especially related to the exploration and exploitation of geo-resources. Some episodes from Soultz-sous-Forêts are already available and the missing-ones will be soon on the platform.

The next step for the CDGP is first to complete data from Soultz-sous-Forêts. Some data are still missing and must be recovered from the industrial partners. Then, data from the other geothermal sites in Alsace (Rittershoffen, Illkirch, Vendenheim) need to be collected in order to be distributed. Finally, with other French data centers, we are on track to apply the CoreTrustSeal certification (ANR Cedre).

The preservation of data can be very challenging and time-consuming. We had to deal with
obsolete tapes and formats, even incomplete data. Old data are frequently not well documented and the identification of owner is sometimes difficult. However, the hard work to retrieve, collect old geothermal data and make them FAIR is necessary for new analysis and the valorization of these patrimonial data. The re-use of data (e.g. Cauchie et al, 2020) demonstrates the importance of the CDGP.