CDGP, a data center for deep geothermal data from Alsace

Marc Schaming[1,3], Marc Grunberg[2,3], Markus Jahn[3], Jean Schmittbuhl[1,3], Nicolas Cuenot[4], Albert Genter[5,3], Eléonore Dalmais[5,3]

[1] Institut de Physique du Globe (IPGS, CNRS/UNISTRA), 5 rue René Descartes F-67084 Strasbourg Cedex, France

[2] Ecole et Observatoire des Sciences de la Terre (EOST), Université de Strasbourg (UNISTRA), 5 rue René Descartes F-67084 Strasbourg Cedex, France

[3] Labex G-Eau-Thermie Profonde, UNISTRA, 5 rue René Descartes F-67084 Strasbourg Cedex, France

[4] GEIE "Exploitation Minière de la Chaleur", route de Soultz, F-67250 Kutzenhausen, France

[5] ES-Géothermie, 3A chemin du gaz F-67500 Haguenau, France

Marc.Schaming@unistra.fr

Keywords: geothermal data, metadata, spatial data infrastructure, Upper Rhine Graben, Soultz-sous-Forêts, EPOS-Anthropogenic Hazards, Episodes.

ABSTRACT

CDGP is the data centre for LabEx G-EAU-THERMIE PROFONDE. Its objectives are to archive and distribute the data collected on deep geothermal sites in Alsace, in respect with intellectual property rights. Collected data cover the whole life of geothermal projects, from drilling to production. Metadata are defined to allow discovery and selection of data; access to data will be granted after registration and validation to respect distribution rights, data will be distributed in open and community-defined formats. A case study is used to depict the challenges posed by the publication.

1. INTRODUCTION

The LabEx G-EAU-THERMIE PROFONDE is a research program on deep geothermal energy founded by the French Ministry of Research and Education in the framework of the «Laboratories of Excellence» initiative.

It contributes to the development of the use of deep geothermal energy, a source of renewable energy, thanks to a better knowledge of deep geothermal reservoirs and new technologies for exploitation. Since 2012, it is a joint project industry/university, headed by the University of Strasbourg. It has a French ministry funding of 3 million for 8 years.

The project aims to develop knowledge in the geothermal field by bringing together academic skills (EOST/ICUBE) and industrial expertise (ES/GEIE) to study the structure of deep geothermal reservoirs of the Upper Rhine Graben.

It follows 3 main objectives: scientific, education and observatory. This latter one is filled by the CDGP (Centre de données de géothermie profonde, deep geothermal data centre, http://cdgp.u-strasbg.fr) that has to archive the data collected in the Upper Rhine Graben geothermal sites and to distribute them to the scientific community for R&D activities. The CDGP main tasks are to collect/centralize, (long-term) archive, and distribute the high quality collected datasets according to IPR (Intellectual Property Rights).

2. DATA

Collected datasets cover the whole life of geothermal projects, from exploration to drilling, stimulation, circulation and production (Fig. 1). They originate from Soultz-sous-Forêts site and later from more recent surveys like Rittershoffen, Illkirch, etc.

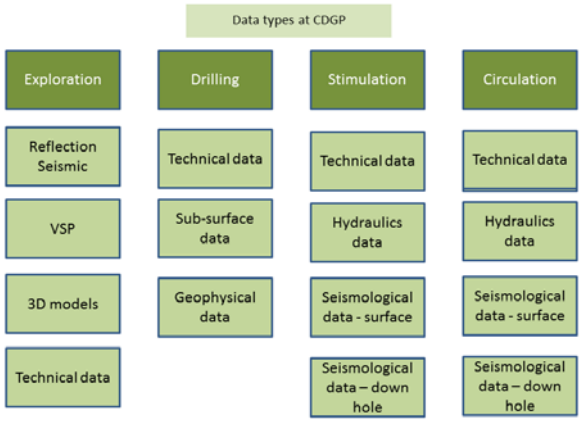

Figure 1: Data types related to the phases of a geothermal project.

Geophysical datasets from Soultz-sous-Forêts project were mainly stored on office's shelves and old digital

media. Old digital files were retrieved from disks and Jazz media.

Geologic and well logging data are described in several reports (e.g. Genter et al., 1997; Genter, 1999; Dezayes et al., 2003, 2005), and consists mainly on excel or text files, images, and paper documents.

Other geophysical data (geomagnetic, gravimetric, seismic) are also available in the vicinity of Soultz-sous-Forêts, and have also to be retrieved.

Datasets are also described in many publications (more than 1700 references on Soultz-sous-Forêts experiment between 1984 and today).

A first step of integration of reservoir data into a PostgreSQL/postGIS database (ISO 19107 compatible) was performed (Jahn, 2014). The database links depths, temperatures, pressures, flows, for periods (times) and locations (geometries) (Fig. 2).

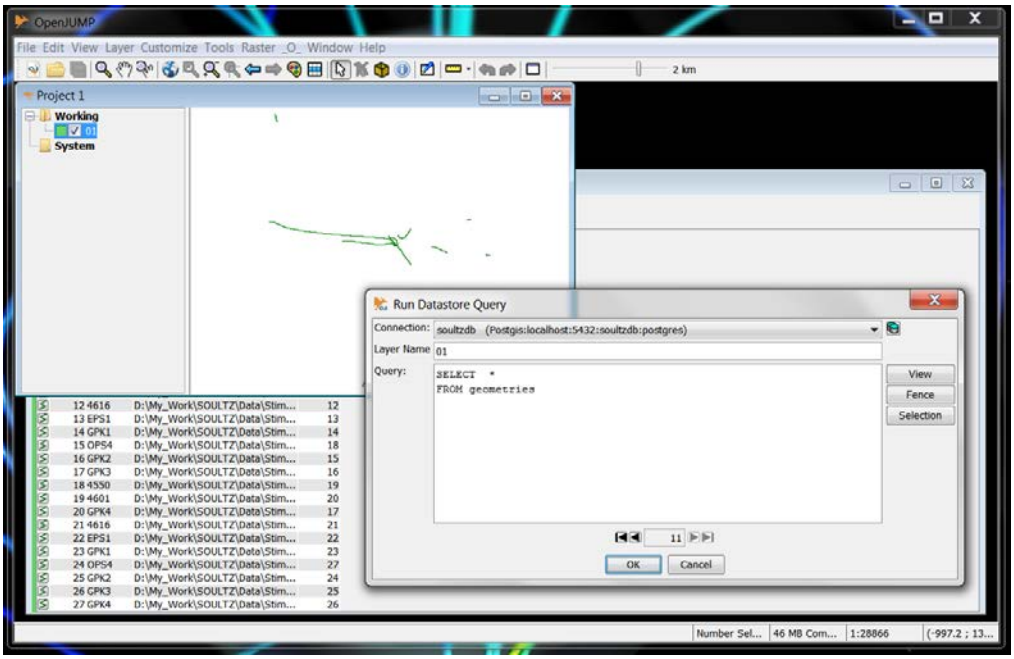

Figure 2: Query example from reservoir database.

Other geophysical data are stored in structured directories.

Seismological datasets are of two kinds: the seismological waveform ("event oriented" or "continuous time series") and the seismicity bulletin.

Recent (>2009) waveform data (Fig. 3) are stored in a standardized way both in format (miniSEED) and in files and directories structures (SDS) following international standard (FDSN); earlier waveforms, were converted from original format to the current standards.

Since 2013, seismicity bulletin (parametric dataset of induced events recorded by a seismic network during the different phases of the geothermal project) are stored in a database following the open standard QuakeML. Older ones are text files that need to be converted to QuakeML.

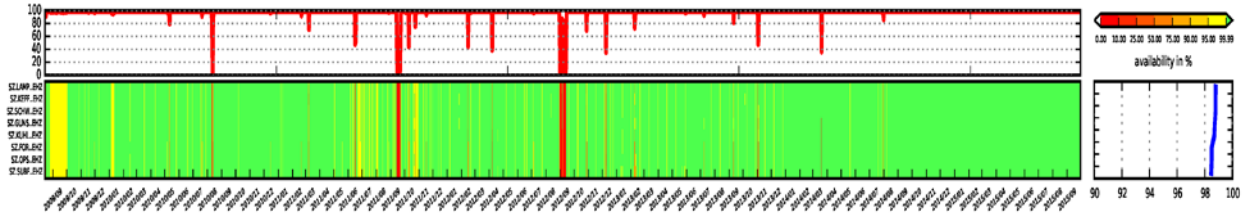

Figure 3: Availability map of seismological data from Soultz-sous-Forêts network (2009-2015).

3. METADATA

Metadata are data that describe the datasets: when, where, what, who, how, where to find the data, etc. Describing the datasets is necessary to organize the discovery of the datasets, and their selection.

CDGP chooses *GeoNetwork* catalog application to manage the resources. It provides metadata editing and search functions as well as a web map viewer. The metadata editor supports ISO19115, ISO19119, ISO19110 and ISO19139 standards used for spatial resources. A demonstrator is under deployment to verify that this application fulfils requirements, one of them being the description of proprietary and access rules.

A step forward will be to add specific metadata records as defined by the Open Geospatial Consortium to provide geophysical/geologic/reservoir information: Observations and Measurements (O&M) to describe the acquisition of information from a primary source, and SensorML to describe the sensors.

Seismological metadata, which describe all the instrumental response, use also the dateless SEED standard.

A major task is to correctly define the data owner(s), associated distribution rules, and terms of use. This is a key point to gain confidence of partners, but is not a basic task. Soultz-sous-Forêts data were acquired in a research project involving many people for a more or less long period, but funding came from industrial and/or national partners, and/or academic institutions and/or EC projects. Few people are now aware of this more than 25 years long story and can help defining the intellectual property of datasets.

For some data, distribution rules depend on the French "Code minier" regulations: the raw underground data (seismic, logging of exploration wells), acquired during an exclusive research license (PER), are open after a 10 years' period of embargo. But a big bunch of the data are outside this scope, e.g. stimulation data, circulation data, additional geophysical data, etc. The distribution rules are to be negotiated with the data owners to respect their Intellectual Property Rights (IPR). The objective is to have some simple rules stating which datasets are available to which community (academic researchers, LabEx partners, collaborating institutions, industrial partners, European research infrastructure like EPOS, citizens, etc.).

And finally the terms of use have to be defined: by citing the dataset (bibliographic reference, digital object identifier DOI), by signing a (commercial) agreement with the data owner, etc.

4. ACCESS TO DATA

Access to data will be granted after registration and validation and will respect defined distribution rights.

We are implementing *geOrchestra*, a free, modular and interoperable spatial data Infrastructure (SDI) that meet the requirements of the INSPIRE directive. The software provides all necessary tools: in addition to the GeoNetwork metadata catalog, it provides a map and features server, an advanced web map viewer, a data extractor. A security proxy and a single-sign-on authentication system for access control.

Access to the datasets will also be granted via EPOS-IP Anthropogenic Hazards project (https://tcs.ah-epos.eu/). Access to episodes (time-correlated collections of geophysical, technological and other relevant geo-data over a geothermal area) and application of analysis (time- and technology-dependent probabilistic seismic hazard analysis, multi-hazard and multi-risk assessment) will be services accessible via a web portal and will require AAAI (Authentication, Authorization, Accounting and Identification).

5. EXAMPLE OF SOULTZ-SOUS-FORÊTS 1993 STIMULATION

We'll use the Soultz 1993 stimulation experiment as a case study (Fig. 4), because of its use in Cornet (2015). A microseismic catalog restored from a DVDROM is available online (http://labex-geothermie.unistra.fr/article420.html?lang=en).

	A	B	C	D	E	F	G	H
1	52.2307	-132.0115	-2882.2935	M	25	sep93003	02-sept-93	16:19:11:72
2	54.8949	-133.9915	-2984.4690	M	26	sep93003	02-sept-93	16:23:33:33
3	185.6788	-99.0925	-2883.4912	M	27	sep93003	02-sept-93	18:27:13:58
4	122.1373	-114.6369	-2987.7437	M	28	sep93003	02-sept-93	18:29:34:87
5	118.6575	-153.2012	-2947.1094	M	29	sep93003	02-sept-93	18:29:42:97
6	63.7802	-147.0027	-2901.5874	M	30	sep93003	02-sept-93	18:37:06:82
7	264.4823	-100.9207	-2956.1807	M	31	sep93003	02-sept-93	19:03:31:27
8	70.3564	-103.1797	-2988.6089	M	32	sep93003	02-sept-93	19:13:46:17
9	153.2398	-156.0894	-2920.5388	M	33	sep93003	02-sept-93	19:15:01:01
10	98.2670	-142.2277	-2912.2327	M	34	sep93003	02-sept-93	19:17:55:18
11	83.6266	-143.1648	-2935.7996	M	35	sep93003	02-sept-93	19:18:25:64
12	106.6186	-153.9764	-2957.0894	M	38	sep93003	02-sept-93	22:27:35:04
13	47.0185	-115.9020	-2941.2256	M	39	sep93003	02-sept-93	22:40:17:88
14	83.9763	-143.8768	-2932.2437	M	40	sep93003	02-sept-93	22:42:34:41
15	111.8301	-103.7903	-3007.7327	M	41	sep93003	02-sept-93	22:42:38:01
16	74.2460	-131.0387	-2942.3486	M	42	sep93003	02-sept-93	22:45:27:22

Figure 4: First lines of the microseismic catalog related to 1993 stimulation at Soultz-sous-Forêts.

To fully describe the file, we had to investigate into PhD thesis legacy codes, to find out the signification of each column. X (A) and Y (B) are related to injection well (GPK1), Y toward Magnetic North, Z (C) is altitude. We have to define position of GPK1, the value of magnetic declination at acquisition period, and round Z values to the meter. We still have to check if Z is related to the top of the well or is real altitude, and if time is given in local or GMT time (including changes related to Daylight saving time).

Associated waveforms were retrieved from the same CDROM and converted in miniSEED format (Fig. 5).

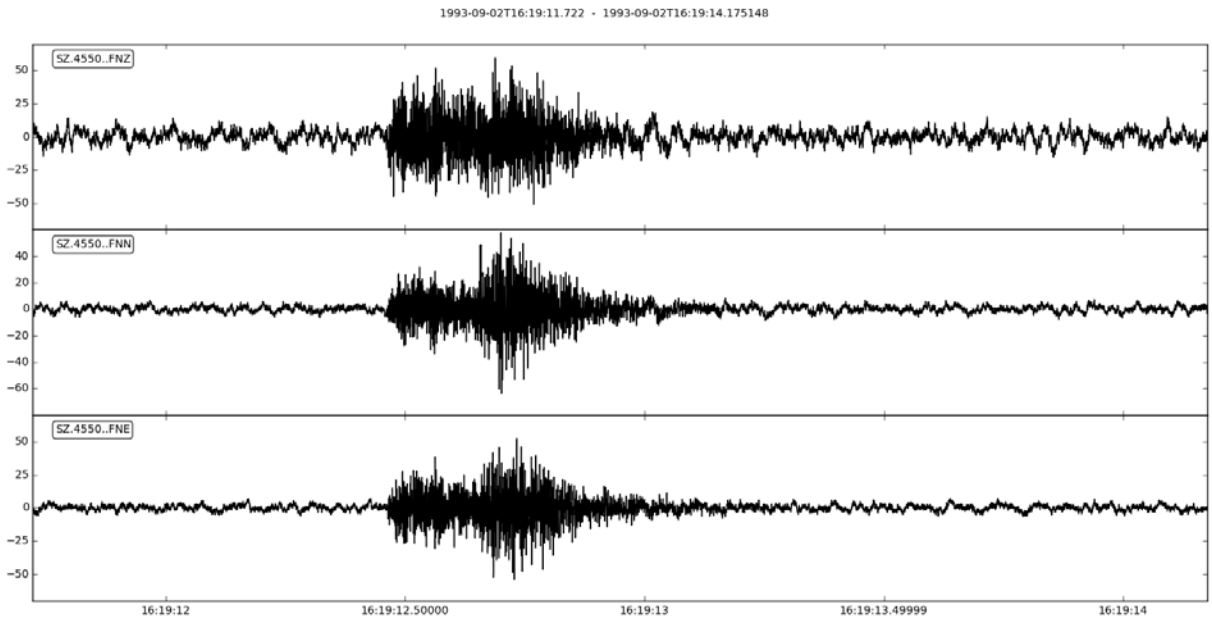

Figure 5: Recovered seismogram from 1993.

Hydraulic datasets (volumes, flow, pressure vs time) are chased in old digital backups.

Metadata are populated with first all mandatory information, and later with possible additional information.

6. CONCLUSION

CDGP, the data centre for deep geothermal data from Alsace of LabEx G-Eau-Thermie-Profonde, started challenging activities to retrieve and collect data from geothermal projects, to (long term) archive them, to describe them in metadata allowing discovery and selection, and distribute them. CDGP is setting an open and interoperable spatial data infrastructure. It is not a trivial issue to define correctly data owners, distribution rules and terms of use, but it's crucial to protect intellectual property rights and to earn the trust of academic and industrial partners. It is only if this happens that CDGP can complete its mission.

REFERENCES

Cornet, F.H.: Elements of Crustal Geomechanics, *Cambridge University Press* (2015).

Dezayes, C., Genter, A., Homeier, G., Degouy, M., Stein, G.: Geological study of GPK3 HFR borehole (Soultz-sous-Forêts, France), *Rapport BRGM RP-52311* (2003), http://infoterre.brgm.fr/rapports/RP-52311-FR.pdf

Dezayes, C., Chevremont, P., Tourliere, B., Homeier, G., Genter, A.: Geological study of the GPK4 HFR borehole and correlation with the GPK3 borehole (Soultz-sous-Forêts, France) *Rapport BRGM/RP-53697-FR* (2005), http://infoterre.brgm.fr/rapports/RP-53697-FR.pdf

EPOS-IP, Integrating European Research Infrastructures for solid Earth Science: *European Union's Horizon 2020 research and innovation*, http://www.epos-ip.org/

Genter, A., Traineau, H., Artigan, D.: Synthesis of geological and geophysical data at Soultz-sous-Forêts (France). *Rapport BRGM R39440* (1997), http://infoterre.brgm.fr/rapports/RR-39440-FR.pdf

Genter, A.: Geological and well-logging data collected from 1987 to 1999 at the HDR site Soultz-sous-Forêts. *Rapport BRGM R40795* (1999), http://infoterre.brgm.fr/rapports/RR-40795-FR.pdf

GeoNetwork: http://geonetwork-opensource.org/.

geOrchestra: http://www.georchestra.org/.

Jahn, M.: Guide to the Soultz PostGIS Database & OpenJUMP-Client(s). *Internal report* (2014).

Acknowledgements

We thank LabEx G-Eau-Thermie-Profonde, COGEOS Consortium, GEIE, EPOS IP, Ademe EGS Alsace for their support.