

# How to deal with old and industrial data?

## The example of the Data Center for Deep Geothermal Energy (CDGP)

Marc Schaming<sup>1</sup>, Alice Frémand<sup>2</sup>, Nicolas Cuenot<sup>3</sup>, Eléonore Dalmais<sup>3</sup>, Jean-François Girard<sup>1</sup>, Clément Grellier<sup>2</sup>, Marc Grunberg<sup>2</sup>, Jean Schmittbuhl<sup>1</sup>

<sup>1</sup>Université de Strasbourg, CNRS, IPGS UMR7516, F-67000 Strasbourg, France / <sup>2</sup>Université de Strasbourg, CNRS, EOST UMS830, F-67000 Strasbourg, France / <sup>3</sup>ES-Géothermie, 5 rue de Lisbonne F-67300 Schiltigheim, France

✉ cdgp@eost.unistra.fr

### INTRODUCTION

Old and industrial data are useful datasets, but because of their specific status, they bring new challenges for their diffusion: recovering them and dealing with Intellectual Property Rights (IPR) are often big difficulties. How to deal with these data? The example of the CDGP (Centre de Données de Géothermie Profonde, deep geothermal data center) is presented in this poster.

### THE CDGP



The CDGP has been set up by the LabEx G-EAU-THERMIE PROFONDE to preserve, archive and diffuse legacy data acquired on the geothermal sites of the Upper Rhine Graben. It is now a local node for the European Plate Observing System (EPOS) Anthropogenic Hazard platform.

### DATA DISTRIBUTED BY THE CDGP

Data processed at CDGP mainly consist of seismological and hydraulic data that have been acquired during stimulation or circulation phases at Soultz-sous-Forêts pilot plant. They are gathered into "episodes": time-correlated collections of geophysical, technological and other relevant geo-data over a geothermal area. Data are available on the CDGP platform.

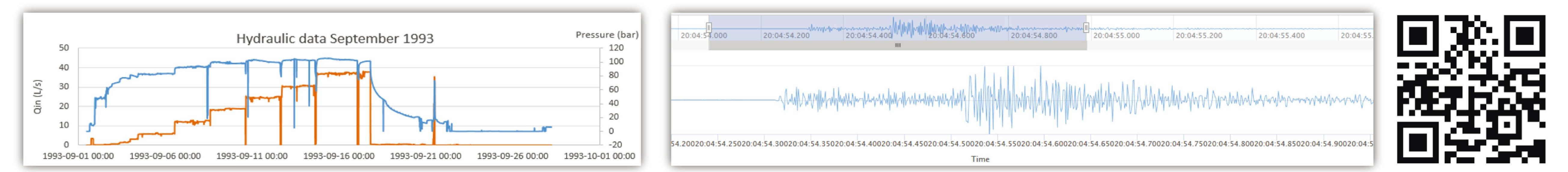


Figure 2: Example of data distributed by the CDGP infrastructure. (<https://cdgp.u-strasbg.fr/>)

### OLD DATA

Dealing with old data becomes an issue if datasets are not regularly processed. With time, media, formats, softwares become obsolete and information may be lost. A specific workflow has been set up at the CDGP to process legacy data (figure 1). Now new users give interesting feedbacks, contributing to the completeness of the datasets.

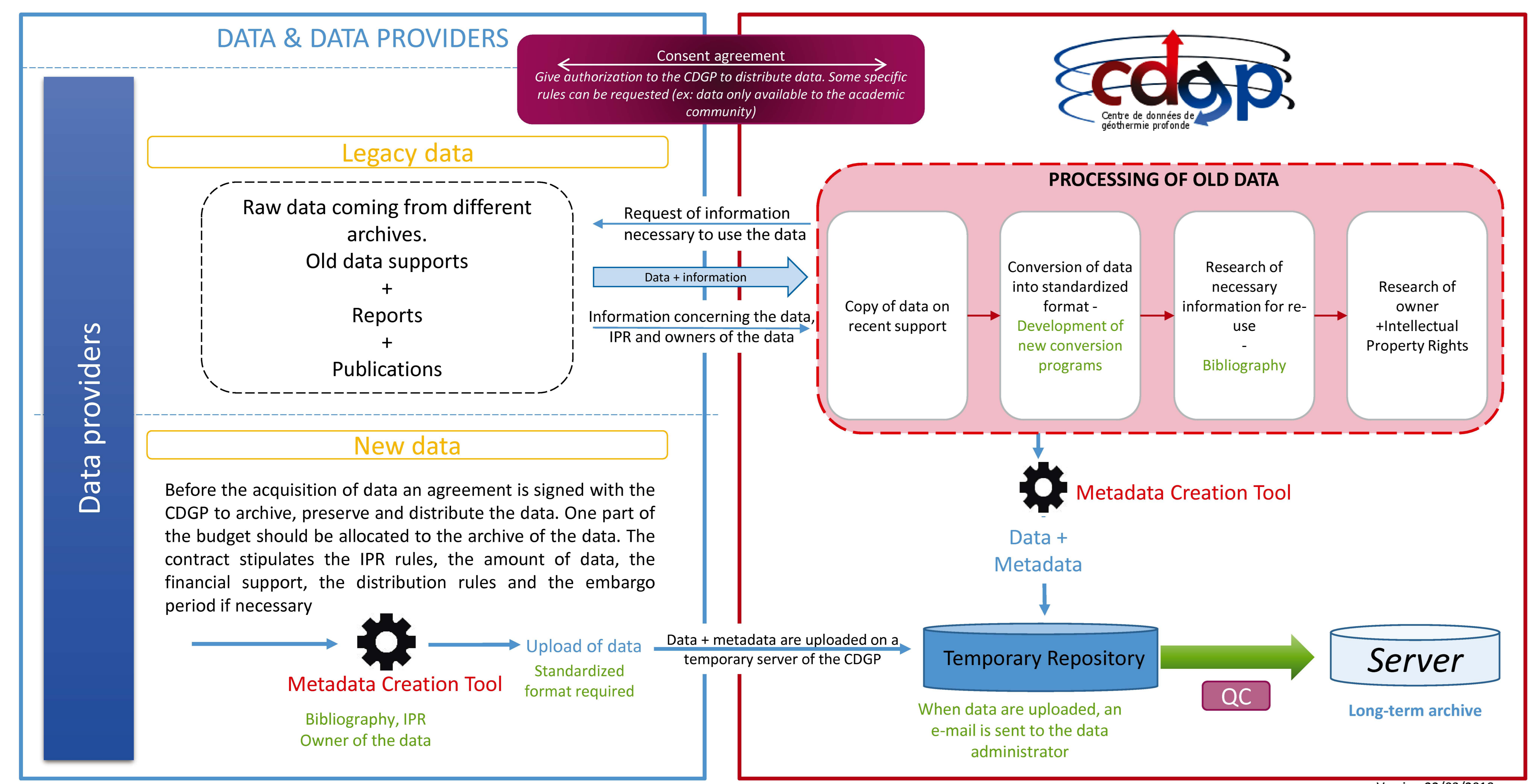


Figure 1: Functional diagram showing the processing of old and new data

### INDUSTRIAL DATA

Industrial data bring new challenges in terms of IPR. Indeed, information about the geothermal resources and specific expertise in the geothermal field may need certain confidentiality for business reasons. The CDGP set up a new workflow to deal with industrial data (figure 3).

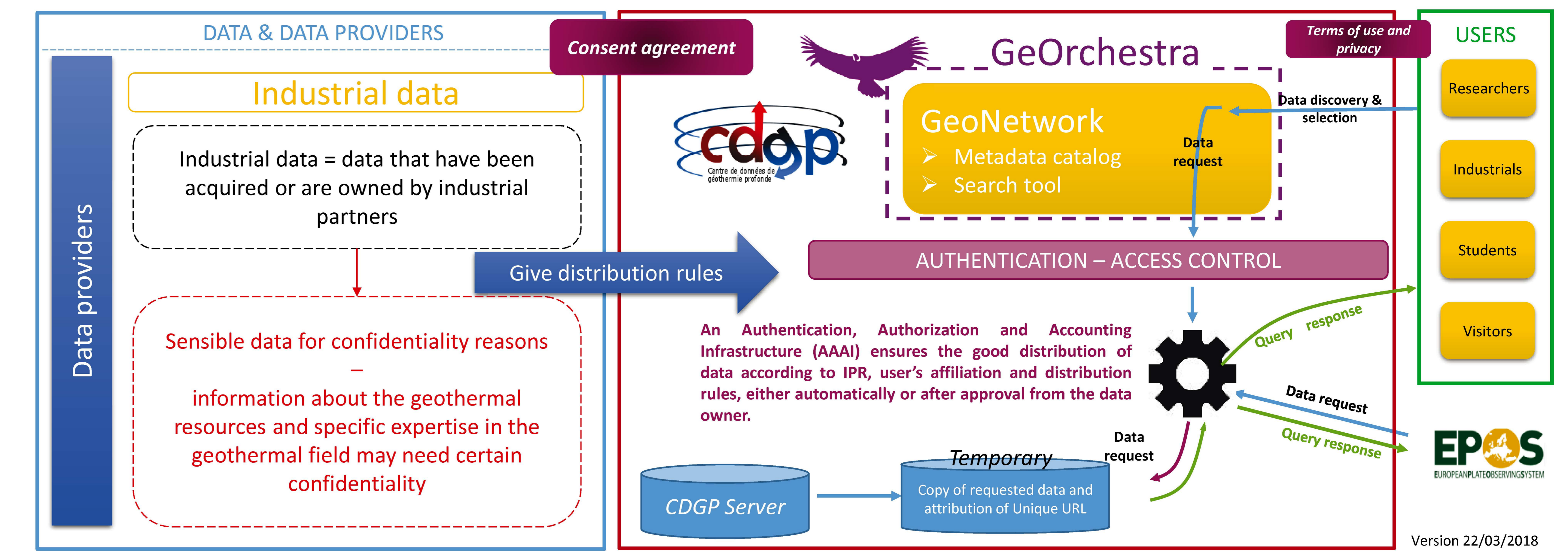


Figure 3: Functional diagram showing the processing of industrial data

### LESSON LEARNT IN TERM OF DATA MANAGEMENT

The key to preservation is distribution by keeping datasets alive. Defining the status of data, the owners, the embargo period and the distribution rules before the acquisition can save a lot of time and avoid heavy processing of obsolete data. Documentation, use of standard formats and data description in metadata are crucial for data re-use and preservation. Procedures to process new datasets at CDGP have been set up following these simple recommendations and are gathered in a Data Management Plan (DMP).