

# Data Management Plan (DMP) – Research project C14/21/056

## ADMIN DETAILS

### Project Name:

Multiple mineralization and fluid flow phases in the Central African Copperbelt: implications for the age dating of the ore-forming processes.

**Principal Investigator:** Philippe Muchez

**Researcher:** Jessica Torres

## 1. GENERAL INFORMATION

### Name applicants

*Philippe Muchez - Jessica Torres*

### Project Number & Title

C14/21/056 - Multiple mineralization and fluid flow phases in the Central African Copperbelt: implications for the age dating of the ore-forming processes.

### Affiliation

*KU Leuven, Department of Earth and Environmental Sciences, Geology Division*

## 2. DATA DESCRIPTION

### Will you generate/collect new data and/or make use of existing data?

*This project will reuse existing data and generate new data.*

**Describe the origin, type, and format of the data (per dataset) and its (estimated) volume, ideally per objective or WP of the project. You might consider using the table in the guidance.**

*The existing data to be reused are thick and thin sections previously produced and stored at the Geology laboratories, Celestijnenlaan 200C (building 493-11), lab rooms 00.76 and 00.77. The inventory of the rock samples and sections is compiled in an Access file, Access 2003.mdb (60Mb).*

*The primary data that will be generated consists of*

*1) Analogue data:*

*1.1) Samples: thick sections produced from rock samples stored in the geology storage facility, Celestijnenlaan 200C (building 493-11), lab rooms 00.76 and 00.77. The inventory of the rock samples and sections is compiled in an Access file, Access 2003.mdb (60Mb).*

1.2) Researchers' notebooks.

2) Digital data (total estimated volume of tens to hundreds of GB/year):

2.1) Experimental data (as directly measured): ASCII data files containing numerical data (measurement output) and text metadata (describing analytical/instrumental parameters).

2.2) Geochemical datasets initially produced as raw data files with different formats (.txt, .dat, .csv, .l6s), depending on the analysis technique (FEG-EPMA, LA-ICP-MS, Raman spectroscopy).

2.3) Processed and analyzed experimental data: Data related to post-processing will be dominantly stored as dedicated analytical file formats (e.g., Labspec files), text documents, Microsoft Excel files or codes/scripts/output from the R statistical programming software environment.

2.4) Electronic laboratory logbooks.

2.5) Maps and (micro-)photographs related to the geological framework and petrography, as well as diagrams and other graphical representation of data, used for (or resulting from) data analysis, will be produced in the .jpg, .png, .tiff, and .pdf formats.

2.6) Reports, manuscripts, and presentations will be dominantly created in Microsoft Word and PowerPoint. 2.

*The kind of data that is generated is similar for all work packages of the project.*

**The answers to this section were checked by:**

*Philippe Muchez*

### **3. LEGAL & ETHICAL ISSUES**

**Will you use personal data? If so, shortly describe the kind of personal data you will use (add the reference to your file in your host institution's privacy register - not relevant yet)**

*No*

**Are there any ethical issues concerning the creation and/or use of the data (e.g., experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)**

*No*

**Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?**

*Although tech transfer and valorisation are not direct goals of this project, if such an opportunity presents itself, it will be discussed among the staff involved in the project. The conclusions of that discussion will be appended to this data management plan.*

**Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?**

No

**The answers to this section were checked by:**

*Philippe Muchez*

## **4. DOCUMENTATION & METADATA**

**What documentation will be provided to enable reuse of the data collected/generated in this project?**

*During each experiment, a detailed logbook will be used (different logbooks for the different analysis setups). These logbooks will contain the date, a brief description of the performed analysis, the parameters used for each measurement, as well as the names of all the saved files.*

*The names of the files will be structured in a comprehensible way: sample studied/date/main parameters used.*

*In addition, data will be stored in a folder per analysis setup, the type of investigated sample, and the corresponding date. In this way, by tracking the corresponding logbook notes, each file can be easily found on the local computers controlling the setup and on the server of the laboratory.*

*The analysis files will contain notes describing the analysis procedure and mention which original data files are included. A readme file describing the goal of the experiment and the analysis procedure will be stored in the folder where the data is saved.*

*An Access database will also be produced with the results obtained from each analytical method. Thus, all information about the samples will be easily and quickly displayed whenever needed.*

**Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.**

*In the relevant research field, there is no formal metadata standard. However, the standardized steps described in section 4.1 will ensure that the data is easy to find and reuse.*

**The answers to this section were checked by:**

*Philippe Muchez*

## **5. DATA STORAGE & BACK UP DURING THE PROJECT**

**Where will the data be stored?**

*Rock samples are stored for at least 10 years in the geology storage room 01.12, Celestijnenlaan 200C (building 493-11), Heverlee and all sections prepared from the rocks in lab rooms 00.76 and 00.77, Celestijnenlaan 200C, Heverlee.*

*During the project, all physical logbooks and the researcher notebooks will be stored in the room 02.202 of the Geo-institute, Celestijnenlaan 200E. At the end of the project they are stored in room 02.206, same address.*

*The electronic data, including the logbook copies, raw data, and analysis files, as well as reports, will be saved on local computers. The data stored on these computers is synchronized with the data server of the laboratory. A duplicate of this data is made overnight on the backup server. The servers of the laboratory (vsm-datasrv1 and vsm-datasrv2) will be used in this respect. All researchers involved in the project will have access to the corresponding folder.*

*Moreover, the electronic data will be saved on the researcher's KU Leuven One Drive account as an additional personal backup.*

#### **How is back up of the data provided?**

*All data at the server are automatically backed-up on a daily basis (overnight routine) as implemented by the IT responsible.*

**Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.**

*Yes. On the local computer, the researcher will have 50 GB available at the Personal (I:) folder, 0.99 TB (K:) at the Archive folder, and 0.99 TB at the LSV (L:) folder.*

*The amount of data that will be generated in this project should not exceed a few hundred GBs, which is small enough to be stored on local computers and on the data server of the laboratory.*

**What are the expected costs for data storage and back up during the project? How will these costs be covered?**

*The laboratory has its own data server and backup system. The cost for replacements of outdated infrastructure is shared by the staff members of the laboratory. Assuming a minimum lifetime of 4 years, the costs budgeted on this project are a few hundred euros per year.*

**Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?**

*The data generated during the project will be systematically transferred to the VSM server, with restricted access (managed by the IT responsible). Only the promotor and involved researchers have access to the shared folders where the data, analysis files and reports will be stored. Also, credentials are required to log in to local computers in the lab.*

**The answers to this section were checked by:**

*Philippe Muchez*

## **6. DATA PRESERVATION AFTER THE PROJECT**

**Which data will be retained for the expected 5 years period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).**

*All the generated data will be stored at the VSM server for at least 5 years after the project ends.*

**Where will the data be archived (= stored for the longer term)?**

*Data on work PCs are automatically backed-up daily with the Microsoft OneDrive backup-archive client. After completion of a work-package, the resulting research data will be more permanently stored on the KU Leuven ICTS "Large Volume Storage" infrastructure (<https://icts.kuleuven.be/sc/english/storage/large-volume-storage>).*

**What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?**

*The system and cost allocation mentioned under 5.4 will also be used for data preservation after the project ends. Costs are expected to be small (few hundred euros/year) and will be covered by the projects in the group (in the worst case scenario of no running projects, accumulated reserves will be used).*

**The answers to this section were checked by:**

*Philippe Muchez*

## **7. DATA SHARING AND REUSE**

**Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?**

*There are no legal restrictions or restriction related to IP potential.*

**Which data will be made available after the end of the project?**

*All relevant data can be made available on an open repository, for example if requested by the editor or publisher of a scientific journal or via restricted access upon request of an individual (e.g., a researcher who intends to reproduce an experiment).*

**Where/how will the data be made available for reuse?**

*Upon request and after the agreement of the project promotor, all data can be made available on a repository.*

**When will the data be made available?**

*After the research results have been published and after agreement of the promotor.*

**Who will be able to access the data and under what conditions?**

*All involved researchers will have access to all the data. Upon request and after agreement of the promotor, access to the data can be granted to other individuals/parties.*

**What are the expected costs for data sharing? How will the costs be covered?**

*Since the volume of most files is not too large, freeware can be used to transfer and share the files.*

**The answers to this section were checked by:**

*Philippe Muchez*

## **8. RESPONSIBILITIES**

**Who will be responsible for data documentation & metadata?**

*The promotor of the project*

**Who will be responsible for data storage & back up during the project?**

*The promotor of the project, together with Hilde Vandenhoeck, the IT responsible of the research group, who takes care of the implementation of the storage and regular back up on the VSM server: Information and communication technologies unit, Science, Engineering and Technology Group, KU Leuven (hilde.vandenhoeck@kuleuven.be)*

**Who will be responsible for ensuring data preservation and reuse?**

*The promotor of the project*

**Who bears the end responsibility for updating & implementing this DMP?**

*The promotor of the project*