Exploration of Open Science Practices within the Geosciences

GAGE
National Science Foundation's Geodetic Facility for the Advancement of Geoscience

Allison Sowers[1] and Jimmy Swift[2]
Front Range Community College[1] and San Juan College[2]

UNAVCO
USGS science for a changing world

Project Overview

Our project goal is to highlight the importance of open science by utilizing open science principles to establish a digital archive of interoperable geochemical data for the US Geological Survey (USGS).

Background

The USGS is responsible for **collecting** and **sharing** many different types of **data** characterizing Earth systems.

Our primary focus was the **GeoArchive project** where we are formally documenting an information management protocol for dealing with data extraction from **National Instrument (NI) 43-101 technical reports** and the archival process for these reports.

Importance of Open Science

Open Science makes data available and interoperable. The goal of open science is to increase the reproducibility and integrity of academic research across all sectors.

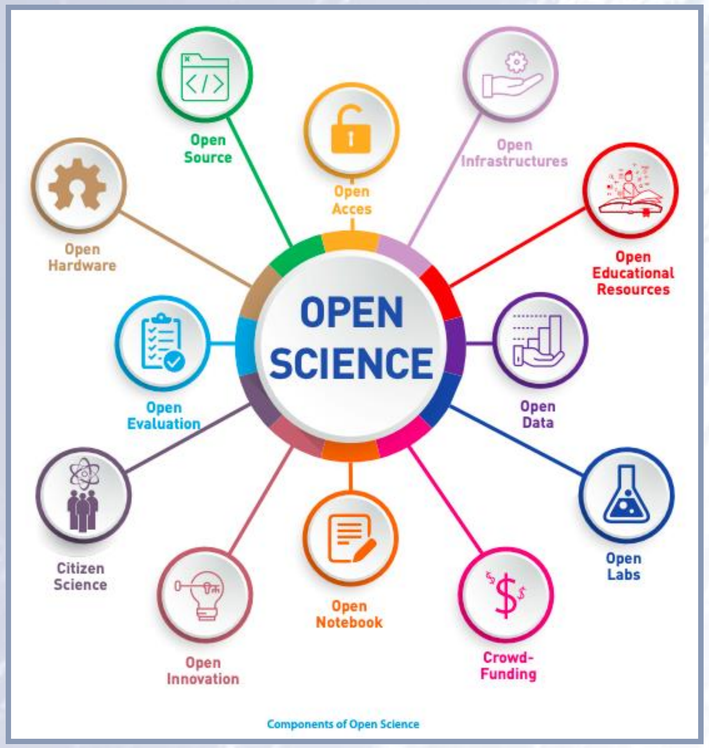

Figure 1: UNESCO Open Science Brochure, 2021.

GeoArchive

- The **GeoArchive** project goal is to promote **open science principles** through creating **standardized operating procedures** (SOP's) for geochemical data and making these standardized data reports more widely accessible.
- These **technical reports document a summary of scientific and location information about geochemical samples** from mineral deposit sites around the world.
- We **published a SOP** on the open science platform Protocols.io **to extract and archive the data and metadata** from each report.

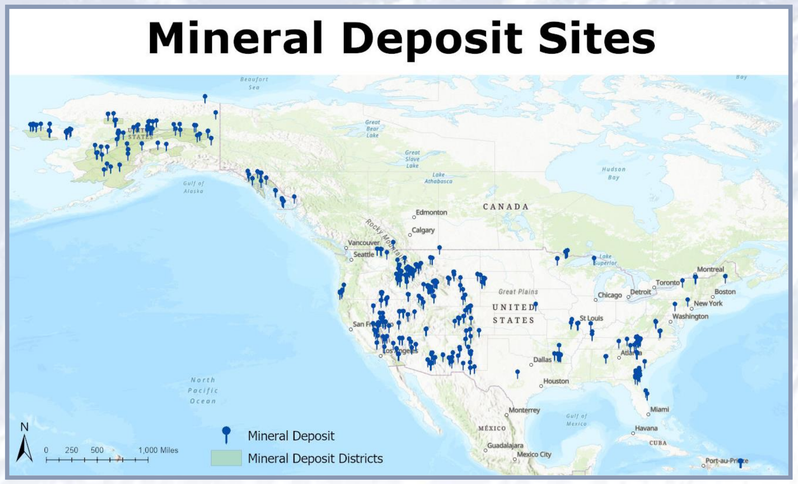

Figure 2: This map shows the mineral deposit locations located at mining sites across North America, this location data comes from the USMIN Mineral Deposit Database. The GeoArchive project is updating the data and metadata for these mineral deposits.

Process

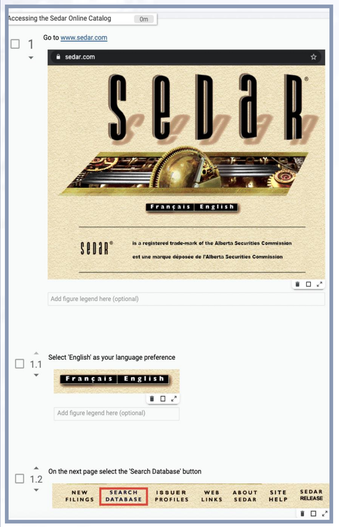

Figure 3: Samples from the SOP on Protocols.io

1. **Analyzed** the problem of data extraction with the **NI 43-101 reports**.
2. **Pulled** these data files from **SEDAR.com** then recorded steps.
3. **Designed** a second protocol used to **publish reports** on Sciencebase.Gov identifying metadata in the report.
4. **Combined** the two protocols used for **extraction and publication** of the reports.

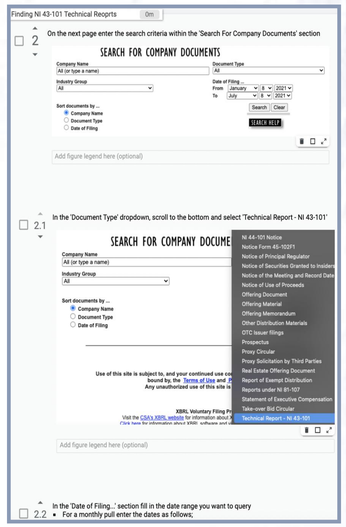

Figure 4: Samples from the SOP on Protocols.io

Open Science Tools

protocols.io

Protocols.io was a useful tool for designing protocols that were used to retrieve and record formalized scientific research data.

zoterobib

For our bibliography organization we used Zotero to manage bibliographic data or related research materials.

OSF

We used Open Science Framework (OSF.io) as an open shared archive for this project.

Trello

We used Trello as a task manager for this project. Team members were able to manage all tasks in the workspace collectively.

.

GitHub

We used GitHub to publish our Python code that we made to create an archive of our Trello board, this was done to keep our work transparent.

Products

OSF Work Space

Protocol.io SOP for Zotero

Protocol.io SOP for data extraction from SEDAR

Acknowledgements

We would like to thank our project mentor from the USGS Sky Bristol for sharing his time and knowledge with us. We would also like to thank our mentors from UNAVCO, Kelsey Russo-Nixon and Anika Knight for providing us with the opportunity to participate in this academic and professional development program.

We also want to thank the Geological Society of America for providing us with the On To the Future (OTF) award which will fund our travel and enable us to present at the GSA Connects conference this fall.

Land Acknowledgement:
We feel it is important to acknowledge that we are on the traditional territory of the Cheyenne, Núu-agha-tuvu-pu (Ute), Arapahoe, Očhéthi Šakówiŋ (Sioux), and Jicarilla Apache peoples.

This material is based upon work supported by the National Science Foundation under Grant No. 1724794. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

SAN JUAN COLLEGE

References:
Bristol, S. (2021, May 24). *Google Colab*. Google Colab.
Bristol Sky. (2021, June 14). *Trello.com* [Is a web-based list-making application]. Trello.Com/Summer Intern Task.
CERN. (2021, July 22). *Zendo* [Library]. Zenodo.Org.
Swift, J., Sowers, A., & Bristol, S. (2021, May 24). *OSF (EMMA Summer of R&D)* [Open source web application that connects and supports the research workflow]. OSF.Io. https://osf.io/t6azp/wiki/home/
United Nations Educational Scientific and Cultural Organization (UNESCO). (2021). *Towards a UNESCO Recommendation on Open Science*. https://en.unesco.org/sites/default/files/open_science_brochure_en.pdf

FRONT RANGE COMMUNITY COLLEGE