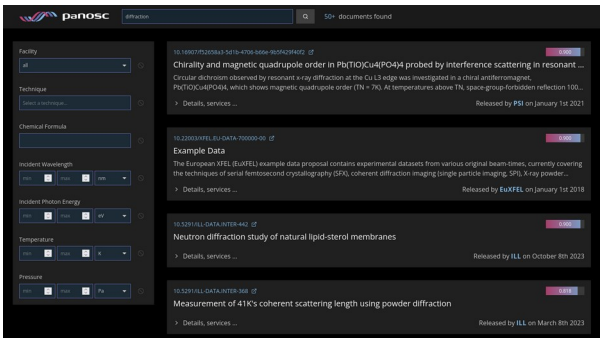
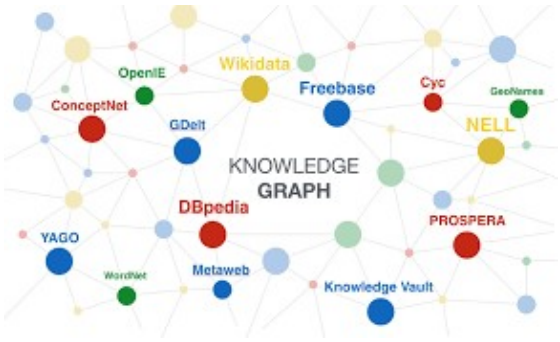
Student project
European Spallation Source
Data Management and Software Centre

Creating knowledge graph and estimating biases in data from photon and neutron open science cloud data portal





Supervisor	?	Desired coding skills	★★★★☆
Co supervisor	Max Novelli	Desired data science skills	★★☆☆☆

DESCRIPTION

The European Spallation Source (ESS) has embraced the FAIR data principles. Its data policy mandate that all publicly founded research data must be made public after three years. In this context, ESS has participated in the european PaNOSC project and the creation of the PaNOSC data portal. The data portal allows users to search public data available at multiple research facilities across Europe.
The PaNOSC data portal is online and its full potentials are still unexplored. We would also like to improve its functionalities. In order to focus our effort, we would like explore better the behavior and estimate the biases introduced as in type of data returned and distribution of the results across facilities.
This project will be a multi-step process: data collection, data analysis and data visualization. First we will prepare queries to be submitted to the data portal, collect the results and estimate metrics and biases. The process will be iterative and the results from the previous iteration will guide the creation of the queries submitted in the following one. Later a well defined analysis will be performed and different types of visualization will be explored in order to present the data collected and findings in the most effective way.

REQUIREMENTS

Experience with Python, Jupyter Notebook, web API, git and github.
Willing to learn about: data catalog, data management, information retrieval (IR), natural language processing (NLP) and machine learning techniques.

LINKS

https://data.panosc.eu
https://www.go-fair.org/fair-principles/
https://www.panosc.eu/
https://ess.eu

Contact Info

Max Novelli (max.novelli@ess.eu)

Matchmaking Day 2023 Engineering physics and companies meet