Yiu-Kang Hsu
Research Associate
German Mining Museum Bochum
yiu-kang.hsu@bergbaumuseum.de

Report on the atRium Training School in Brno, 16-20 September 2024

I am trained as an archaeometallurgist and currently work at the German Mining Museum in Bochum, where I manage analytical data of archaeological metals as a data curator. My daily tasks involve data management and interpretation, which has sparked my interest in expanding my computational skills to improve my handling of data. The atRium Training School in Brno provided me with a valuable opportunity to learn new methods for querying and analyzing archaeological data. Below is an overview of what I learned and accomplished during the five-day workshop held from 16 to 20 September 2024.

On the first day, we were introduced to the R programming language, one of the most widely used languages in academic research. Thanks to a well-structured lecture, I was able to query, reorganize, and summarize datasets related to archaeological sites, artifacts, and stable isotope data. By the end of the session, we were taught how to visualize results and create custom plots such as scatter plots, boxplots, histograms, and kernel density plots.

The second day focused on semantic technology and linked open data, emphasizing the importance of establishing standards and facilitating data sharing within archaeological communities. A standout example was the ARIADNE project, which hosts a vast archaeological database containing over 4 million online resources, organized in a large knowledge base. This database utilizes RDF triple stores, and we learned how to use SPARQL queries to retrieve specific information based on our interests. Rather than diving into complex query languages, the instructor simplified the process by using R and related packages. By the end of the day, we successfully extracted artifact data from the ARIADNE database and plotted them spatially on a topological map using straightforward R scripts.

During the third and fourth days, we tackled more advanced spatial analysis in R. This was challenging, as it required learning different processing techniques for vector and raster data. For me, the most difficult part was understanding how each command functioned, as they often require different parameters and have specific data type and coordinate system requirements. In addition to the theoretical lessons, the instructors provided a hands-on exercise with a "point pattern analysis" using real archaeological data from Jordan. We analyzed 90 archaeological sites in the region, testing whether their distribution was random or correlated with climate and terrain. This case study proved highly valuable, as it allowed us to test other variables that might influence site placement.

On the final day, we had the opportunity to work on our own datasets and apply what we had learned throughout the week. I chose to focus on deepening my knowledge of SPARQL queries and explored both the Wikidata and ARIADNE knowledge bases. I successfully extracted detailed information about bronze swords from prehistoric Europe and mapped them spatially. I then performed a point pattern analysis to assess whether their distribution was influenced by climatic factors.

Overall, I found the summer school highly enlightening and gained valuable skills for handling archaeological data. The sections on the semantic web and linked open data were particularly beneficial for my daily work managing the large and diverse database at the museum.