

The Aggregation of ROAD Data in the ARIADNE Pipeline: Pitfalls and Successes

Andrew W. KANDEL, Universität Tübingen, Germany

Michael BOLUS, Universität Tübingen, Germany

Angela BRUCH, Senckenberg Gesellschaft für Naturforschung, Germany

Claudia GROTH, Senckenberg Gesellschaft für Naturforschung, Germany

Miriam N. HAIDLE, Senckenberg Gesellschaft für Naturforschung, Germany

Christine HERTLER, Senckenberg Gesellschaft für Naturforschung, Germany

Julia HEß, Senckenberg Gesellschaft für Naturforschung, Germany

Volker HOCHSCHILD, Universität Tübingen, Germany

Zara KANAIEVA, Universität Tübingen, Germany

Maria MALINA, Universität Tübingen, Germany

Christian SOMMER, Universität Tübingen, Germany

Nicholas J. CONARD, Universität Tübingen, Germany

Keywords: *Prehistory, Paleoanthropology, FAIR principles, ARIADNE database, ROAD database*

CHNT Reference: Kandel, A., Bolus M., Bruch A., Groth C., Haidle M., Hertler C., Heß J., Hochschild V., Kanaeva Z., Malina M., Sommer C. and Conrad N. (2022). 'The Aggregation of ROAD Data in the ARIADNE Pipeline: Pitfalls and Successes, in CHNT Editorial board. *Proceedings of the 27th International Conference on Cultural Heritage and New Technologies, November 2022*. Heidelberg: Propylaeum.

The ROCEEH Out of Africa Database (ROAD; www.roceeh.uni-tuebingen.de/roadweb/) contains data about archaeological, paleoanthropological, paleontological and paleobotanical localities in Africa and Eurasia spanning from three million to 20,000 years ago. The database was conceived in 2008 as the ROCEEH project (www.roceeh.net/) began, and data entry started in 2009. Since then, the multidisciplinary team has integrated over 2,200 localities containing more than 20,000 assemblages collected from over 4,700 publications written in English, French, German, Italian, Spanish, Portuguese, Russian and Chinese, among others. ROAD serves as a valuable resource for archaeologists and other paleoscientists because it contains vast amounts of information that can be explored using innovative methods in data science.

ROAD is a relational database managed with a PostgreSQL database management system. The database allows user interaction through its application called ROADWeb, which is a web-based application written in .php, javascript and .html (Fig. 1). ROAD and its applications are hosted on a server located at the University of Tübingen. The ROCEEH team purposely chose to use open access software with the intention of increasing the database's longevity.

To make ROAD data more FAIR in the future, the research team is working to incorporate its data into the Semantic Web and Linked Data. Almost all data in the Semantic Web are distributed using Resource Description Framework (RDF), a highly interoperable standard developed by the World

Wide Web Consortium (W3C) to describe data or metadata. In 2021, the ROCEEH team completed the development of an RDF data model (i.e. ontology) and the RDF export of ROAD data.

In accordance with overriding developments towards open science, ROCEEH registered ROAD with the repository re3data (www.re3data.org/), and published it under an open Creative Commons license (CC BY-SA 4.0). Based on our experience with data models, thesauri and data synthesis, we worked to promote sustainability of the database by developing standardized practices. Our work was complemented by networks of collaboration with ARIADNE, the Coalition for Archaeological Synthesis, and the German National Research Data Infrastructure (NFDI4Objects), among other agencies.

ROCEEH first met with the ARIADNEplus team in Prato in January 2020, to plan out a timeline for data integration. After this, ROCEEH began to use ARIADNE's data infrastructure (portal.ariadne-infrastructure.eu/) in order to map the data contained in ROAD onto ARIADNE's scheme. With the help of standardized vocabularies such as the Getty Art & Architecture Thesaurus (AAT) and PeriodO, which stores our defined chrono-cultural entities, ROCEEH successfully completed the first round of data integration in September 2021 (Fig. 2). Since then, users are able to search ARIADNE to find the prehistoric data contained in ROAD, a function which enhances the use of both databases. The first update occurred in March 2022, and additional updates are planned every six months.

In this presentation we report on some of the pitfalls and successes our team encountered as we tried to make ROAD data available in the ARIADNE portal. For example, one setback occurred when we tried to map ROAD attributes to those of ARIADNE using their 3M tool (Mapping Memory Manager). We could not bring the geological ages of finds in ROAD into ARIADNE's graph database. The issue was that the model which describes the datasets contained in the ARIADNE catalog (AO-Cat), offered no appropriate resource class for establishing the geological age of the finds, while this feature was present in ROAD. Another setback occurred during the mapping phase, when we discovered that the Getty AAT lacked certain entries better suited for prehistoric artifacts and cultures. We had to homogenize ROAD data to overcome this. Another issue was the regionalization of ROAD's cultural entities, as these did not conform well with those in PeriodO. We used alternative labels to solve this. Despite these setbacks, we succeeded in integrating ROAD data and continue to update ARIADNE periodically.

We also highlight our ongoing efforts to make the data FAIR (findable, accessible, interoperable, reusable), a philosophy that has become increasingly important in securing the future of Big Data in science. This last topic dovetails nicely into another of ROCEEH's successes, namely in making ROAD data findable through ARIADNE. Finally, we touch upon some of the recent advances the research team made with regard to the database, and expound briefly on the way in which the team innovated methods, designed applications, developed products and gained perspectives, as these issues may have relevance for the other partners of ARIADNE.

To explore the full potential of ROAD and ARIADNE, we encourage you to visit our respective websites (www.roceeh.uni-tuebingen.de/roadweb/ and portal.ariadne-infrastructure.eu/) to discover what else these databases have to offer. Should you wish to explore ROAD further, ROCEEH provides expanded access for anyone interested.

References

- Bolus, M., A. Bruch, M. Haidle, C. Hertler, J. Heß, Z. Kanaeva, A. Kandel, M. Malina & C. Sommer (2020). Explore the History of Humanity with the new ROAD Summary Data Sheet / Durch die Menschheitsgeschichte mit dem neuen ROAD Summary Data Sheet. *Mitteilungen der Gesellschaft für Urgeschichte* 29, 145-147. <https://doi.org/10.51315/mgfu.2020.29008>
- Haidle, M.N., M. Bolus, A.A. Bruch, C. Hertler, V. Hochschild, Z. Kanaeva, C. Sommer & A.W. Kandel (2020). Human Origins – Digital Future, an International Conference about the Future of Archaeological and Paleoanthropological Databases (Summary). *Evolutionary Anthropology* 29, 289-292. <https://doi.org/10.1002/evan.21870>
- Kandel, A.W., M.H. Haidle & C. Sommer (Eds.) In press. *Human Origins – Digital Future: An International Conference about the Future of Archaeological and Paleoanthropological Databases*, 91 pp. Heidelberg, Propylaeum. <https://doi.org/10.11588/propylaeum.882>

The screenshot displays the ROAD website interface. On the left is a blue sidebar with navigation links: 'login', 'User Agreement', and 'Data-Use Policy'. The main content area features search filters:

- Assemblage Category** (or logic combination is used as default):
 - AND (check this box for the 'and' logic combination)
 - Human remains Plant remains Fauna
 - Symbolic artifacts Organic tools Lithics
- Region** (for multiple selection use Ctrl-key):
 - Africa - Central Africa
 - Africa - East Africa
 - Africa - Island Africa
 - Africa - North Africa
 - Africa - Southern Africa
 - Africa - West Africa
 - Asia - Caucasus
 - Asia - Central Asia
 - Asia - East Asia
- Age** (enter age in years or choose period):
 - If age input is incomplete or min age is greater than max age, the age condition will not be considered in the query.
 - min:
 - OR Choose period
 - max:

On the right, a world map shows numerous red circular markers indicating query results. The map includes 'Map' and 'Satellite' tabs, a full-screen icon, a person icon, and zoom controls. At the bottom of the map, it says 'Keyboard shortcuts', 'Map data ©2022', and 'Terms of Use'.

Figure 1. View of the entry page of the ROAD website showing the results of a simple query for localities containing both human remains and stone artifacts. By clicking on a site, a user generates a Site Summary Data Sheet for that locality as a PDF.

ARIADNE PORTAL

🔍 Catalogue

🌐 Browse

⚙️ Services

❓ About

ROAD - locality: Aghitu-3 Cave, finds: archaeology, culture: Upper Paleolithic - Eurasia

📘 Description

Aghitu-3 Cave is located at the base of a basalt outcrop along the Vorotan River valley. This UP site yields the earliest evidence for modern human behavior in Armenia dated between 39-24 ka cal. Notable finds include shell beads, an eyed bone needle and other bone tools, plus many laminar tools made from obsidian sourced up to 250 km away.

📄 Metadata

Original ID: Aghitu-3 Cave, 2, age: 25000-25700

Landing page: <http://www.roceeh.uni-tuebingen.de/roadweb/sdsPDF.php?localityName=Aghitu-3 Cave>

Language: English

Resource type: [Site/monument](#)

Subject - Original:

[Archaeology](#)

[Obsidian](#)

Dating: [Upper paleolithic: -41000 to -20000](#)

Place: [Armenia](#)

Type: [Dataset \(Provided record\)](#)

🔗 Resource links

[🔗 View resource at provider](#)

[📄 Json](#)
[📄 Xml](#)
[📄 Rdf](#)
[🔗 Cite](#)
[✉ Report an issue](#)

📁 Resource is part of

[ROCEEH Out of Africa Database](#)

🎯 Thematically similar

Thematically similar resources based on terms in common of:

Subject & Time period ▾

- [📄 ROAD - locality: En Roche, finds: archaeology, culture: Upper Paleolithic - Eurasia](#)
- [📄 ROAD - locality: En Roche, finds: archaeology, culture: Upper Paleolithic - Eurasia](#)
- [📄 ROAD - locality: Jinsitai Cave, finds: archaeology, culture: Upper Paleolithic - E Asia](#)
- [📄 ROAD - locality: Jinsitai Cave, finds: archaeology, culture: Upper Paleolithic - E Asia](#)
- [📄 ROAD - locality: Jinsitai Cave, finds: archaeology, culture: Upper Paleolithic - E Asia](#)
- [📄 ROAD - locality: Jinsitai Cave, finds: archaeology, culture: Upper Paleolithic - E Asia](#)

Figure 2. Screenshot taken from the ARIADNEplus website showing the results of a search for the Upper Paleolithic site of Aghitu-3 Cave in Armenia. By clicking on the URL (landing page), a user can download a Site Summary Data Sheet, which is a PDF summarizing the results of Aghitu-3 Cave, directly from ROAD without the need to log in.