

## New Perspectives for Data Exploration in ROAD

Christian Sommer<sup>1,2</sup> , Volker Hochschild<sup>1,2</sup> & the ROCEEH Team

<sup>1</sup> The Role of Culture in Early Expansions of Humans (ROCEEH), Heidelberg Academy of Sciences and Humanities, Hölderlinstr. 12, 72074 Tübingen, Germany

<sup>2</sup> Inst. of Geography, Univ. of Tübingen, Rümelinstr. 19-23, 72070 Tübingen, Germany

[Watch the Video](#) 

**Keywords:** Data mining, machine learning, digital humanities, GIS, spatial analysis

Since 2009, the [ROCEEH](#) project has accumulated information on more than 20,000 assemblages in the [ROCEEH Out of Africa Database](#) (ROAD) from the fields of archaeology, paleoanthropology, paleontology, paleobotany as well as geographical information with landscape reconstructions and its diverse elements. Linked over space and time, these data allow a broader picture of the deep human past to be drawn. In parallel to the development of ROAD, [Artificial Intelligence](#) (AI) analytical methods such as [machine learning](#) and [data mining](#) have become increasingly efficient and user-friendly in recent years. Supplementary to ‘classic’ analyses commonly used in the database (A. Kandel–Databases) and modelling approaches (C. Hertler–Applications), we will draw attention to further analytical methods, providing new opportunities to confirm existing hypotheses and discover patterns in the data that lead to the formulation of new hypotheses.

One example highlighted in this presentation investigates spatial human-environment interactions through the use of [unsupervised classification](#) to categorize archaeological sites by paleo-ecological conditions, and furthermore, predict potential habits through supervised classification. The derived suitability maps help to identify overlapping habitats, and thus potential migration corridors, as well as environmental barriers that help to understand (dis-)continuities in the Paleolithic record.

Another example focuses on methods to discover patterns of co-occurrence or obstruction of cultural items in archaeological assemblages. Association rule mining (Agrawal et al. 1993, 1996) is widely used in online marketing and applied to ROAD, offering the possibility to explore relations between e.g. the combination of lithic tools and site types on a quantitative basis. This allows us to identify potentially meaningful relationships between objects in large datasets.

While these methods allow us to gain new insights into prehistoric life from the perspective of quantitative data, the specific characteristics of archaeological data, such as incompleteness and biases, must also be taken into account.

**ORCID®**

Christian Sommer  <https://orcid.org/0000-0001-9062-0876>

**REFERENCES**

Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *SIGMOD Record*, 22(2), 207–216. <https://doi.org/10.1145/170036.170072>

Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., & Vermako, A. I. (1996). Fast discovery of association rules. In U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, & R. Uthurusamy (Eds.), *Advances in Knowledge Discovery and Data Mining* (pp. 307–328). Menlo Park: American Association for Artificial Intelligence. <https://dl.acm.org/doi/10.5555/257938.257975>