ROCEEH Out of Africa Database (ROAD)

HEIDELBERGER AKADEMIE DER WISSENSCHAFTEN
Akademie der Wissenschaften des Landes Baden-Württemberg

THE ROLE OF CULTURE IN EARLY EXPANSIONS OF HUMANS

Foreword

A database is a valuable tool for managing the ever increasing amounts of digital data used in science and everyday life. Structure, quality assessment and accessibility of data require innovative approaches to maintain the life expectancy of a database. The challenge to build up a comprehensive system for the analysis of the early expansions of humans was triggered by our need to integrate interdisciplinary work from the fields of archeology, paleontology, paleoanthropology and paleogeography, as well as computer science. The ROAD database presented here is an open, web-based information system providing maps, view graphs and text explanations. It offers users the ability to carry out SQL-queries and spatial analysis about many aspects of early human expansions. Please enjoy this newsletter which explains the ROAD database and highlights the latest developments.

Prof. Dr. Volker Hochschild
Principal Investigator for GIS and ROAD database

Project background

In 2008 the Heidelberg Academy of Sciences and Humanities initiated a project entitled "The Role of Culture in Early Expansions of Humans" (ROCEEH). The ROCEEH research center is projected to continue for 20 years with the aim of investigating the interplay among cultural evolution, biology and changing environmental conditions. These factors allowed humans to evolve from the behavioural niche of a great ape and move into a culturally defined behavioural niche characterized by a wide range of technological innovations. ROCEEH seeks to reconstruct the spatial and temporal patterns of the expansion of hominins between three million and 20,000 years ago in Africa and Eurasia. The project explores the routes of expansion, the biological mechanisms and the cultural potential with which our ancestors sustained themselves in new environments. The main aim of the project is to decipher and explain the reasons for different hominin expansions.

In order to achieve these goals we assembled meaningful data from all of the relevant disciplines. We integrated paleoanthropological and cultural information, vegetation history, paleontological and paleoecological data, climatic records, stratigraphical settings, age models, and geophysical and geomorphological information. In addition, we included the underlying geographical data such as digital elevation models, and topographic and thematic maps in standardized, homogenous formats.

Since ROCEEH required a specific data infrastructure to store, manage, manipulate and visualize the different data types and formats, we developed our own central, web-based georelational database. This database is known as the ROCEEH Out of Africa Database (ROAD). ROAD allows for the assessment of prehistoric habitats, as well as the dynamics of early human expansion, from archaeological and paleoanthropological perspectives. This web-based, georelational database enables the various disciplines to access and query all data and visualize them in a geographical framework. With its GIS functionalities, ROAD is an essential prerequisite for illustrating spatio-temporal information at different scales, leading to the visualization of complex expansion dynamics. In the following sections we focus on technical details of the database system, its GIS functionalities and the provided interfaces. Moreover, we describe the development process including the specific requests from the different users and highlight the functionalities that allow for the spatio-temporal assessment of data.

Development of ROAD

The ROCEEH research center screened existing databases that focus on topics of human evolution and expansions (e.g., Semal et al. 2004, Weniger et al. 2005, Foertsch 2006, Kondo et al. 2011). However, none of these databases were able to visualize and analyze geographic data with respect to human expansion processes in a suitable manner. For this reason, we decided to develop ROAD.

Aims and functionalities of ROAD

One of the first tasks involved in building ROAD was to define the aims and functionalities of the entire system. We designed ROAD to: 1) be easily accessible; 2) integrate and homogenize the data using internationally accepted data standards; 3) import data; 4) store and access data through data queries; 5) perform simple spatial analysis; 6) visualize data; and 7) exchange data. These general requirements implied specific functionalities. ROAD needed to be easily accessible through the internet with a simple, operable and manageable interface for those entering and using the data. Moreover, ROAD was based on open source software, common standards and related formats (e.g., Dublin Core standards, Open GIS Consortium (OGC) standards, ISO/TC 211, ISO 19115). To conclude, we conceived the entire ROAD system as a web-based interface solution with WebGIS functionalities.

User groups and data sharing protocols

Prior to the design of ROAD the ROCEEH team (Fig. 1) conducted a thorough assessment of the potential user groups, the types of data they required, and proposed protocols for data sharing. Basic information and maps will be provided to the general public, researchers and academics alike in a virtual atlas that summarizes the research results and presents thematic maps. Associated researchers can access selected parts of ROAD in order to enter, query, visualize and analyze data within the framework of their research questions; however, full access to ROAD is restricted to the ROCEEH team. As indicated in Fig. 2, users can be broadly categorized into four classes: 1) general public interested in early human studies, as well as researchers and academics interested in basic information and maps; 2) associated researchers involved in analysis and visualization; 3) associated researchers involved in database and GIS analysis; and 4) associated researchers involved in data input, as well as the ROCEEH team.

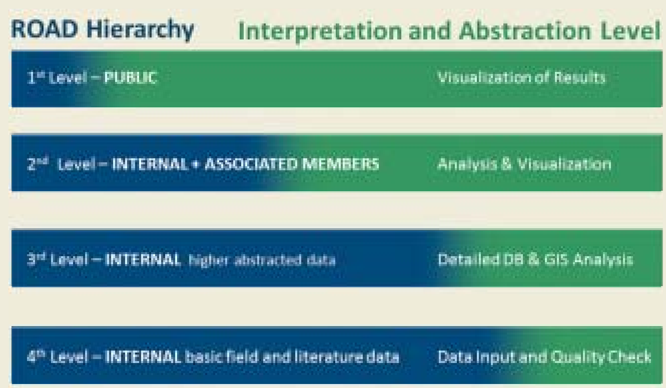


Fig. 2: ROAD hierarchy showing different users, as well as interpretation and abstraction levels.

Fig. 1: ROCEEH team planning the development of ROAD.

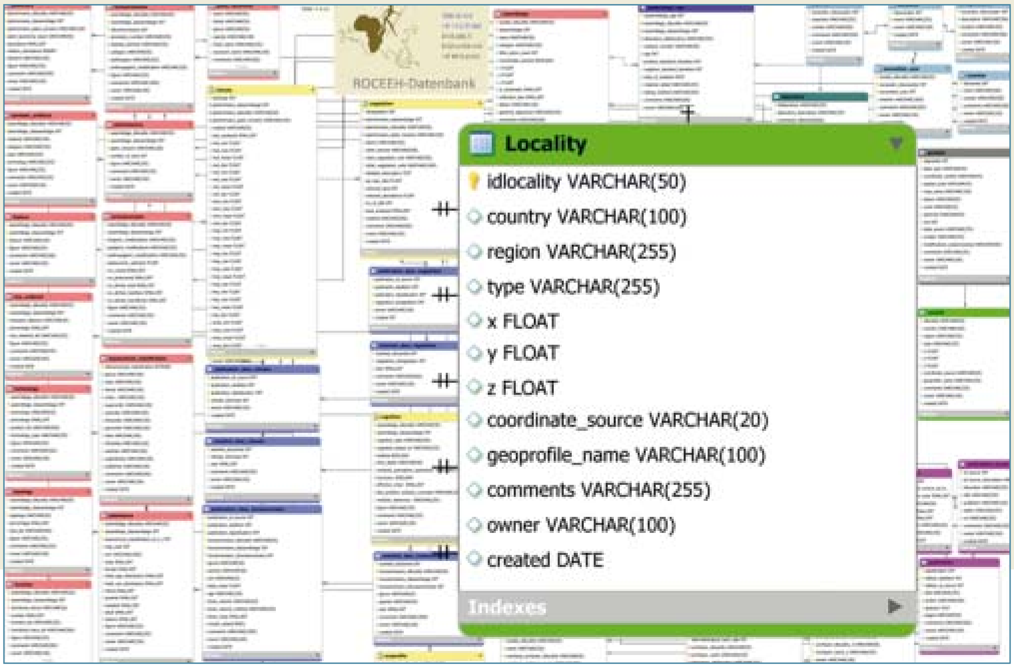


Fig. 3: This "Entity Relationship Diagram" depicts the database structure on the back-end, which is where data are stored. The table "Locality" is enlarged to illustrate an example of structure. Different colors group tables of similar type, for example, pink tables pertain to archaeological finds, while yellow tables contains the results of our own analysis of climate and environment.

The relational structure of ROAD

We approached the design of ROAD by thinking in terms of entities, their descriptive attributes, and the relationships between them. Entities are real-world objects or things which are distinguishable from other objects. The task of designing a database to store the dataset depends on assigning attributes to particular entities and then defining the relationship between these entities. Once a set of entities and their attributes are decided upon, the database design consists of a separate table for each entity and a separate row within each table for each attribute. Rather than storing the dataset as a single large table of information, separate tables with information about different entities are used. We developed tables containing key characteristics, such as assemblage, human remains or geological stratigraphy, and incorporated their descriptive characteristics. These features are illustrated and linked in the "Entity Relationship Diagram" (Fig. 3).

What kind of information is contained in ROAD?

ROAD connects data about sites and assemblages of finds with related datasets containing chronological and geographical data. This functionality enables chrono-spatial analyses based on those datasets. ROAD has the ability to generate distribution maps of various finds and allows various aspects about the finds to be queried across different time ranges.

The structure of ROAD is divided into five main parts:

1. **Localities and assemblages:** General information about archaeology, paleoanthropology, paleontology and geography, gathered from localities (sites) and assemblages (inventories of finds). The tables that yield such information include: Locality, Assemblage, Geodata, Locality has geodata.

2. **Dating:** Chronological assessment of sites, geological layers, archaeological layers and assemblages using radiometric and other methods of dating. Tables needed for the chronological description include: Geological layer, Geological layer age, Geological stratigraphy, Archaeological layer, Archaeological layer age, Archaeological stratigraphy, Assemblage age, Assemblage lies in archaeological layer, Analytical laboratory.

3. **Analysis and interpretation:** Assemblages can be analyzed and further described according to the types of finds, e.g., archaeological finds, faunal remains, human remains and plant remains. We use the following tables to conatain such data: Raw material, Typology, Technology, Function, Feature, Organic tools, Symbolic artifacts, Miscellaneous finds, Cognition, Human remains, Paleofauna, Animal remains, Paleoflora, Plant remains, Plant taxonomy, Vegetation, Scientist describes vegetation, Climate, Scientist describes climate.

4. **Publication data:** Bibliographic information linking publications to localities and assemblages. Tables characterizing publication data include: Publication source, Edition, Publication, Publication describes locality, Publication describes assemblage, Publication describes geological layer, Publication describes geological stratigraphy, Publication describes archaeological layer, Publication describes archaeological stratigraphy, Publication describes human remains, Publication describes climate, Publication describes vegetation, Publication describes cognition.

5. **Researchers:** Information about scientists, researchers, excavators and institutions such as: Scientist, Excavator, Person, Organization, Excavation year.

Spatial vs. non spatial databases

The ROAD system stores data in terms of entities, attributes and relationships. Data that do not contain explicit geographic information are considered non-spatial data. Information that describes geographic data, such as map name, creator and scale, are non-spatial data, or so called metadata.

While non-spatial data can be used to understand numeric and character types of data, additional functionality is required in a database in order to process spatial data types. A spatial database is optimized to store and query data related to objects in space, including points, lines, polygons and raster data. These objects are typically called geometry or feature. The spatial data stored in the database can be subdivided in two types: vector data and raster data (images). Vector data can be stored directly in ROAD tables and consists of point type, line type and polygon type data, for example, find locations, cities, river network, roads, country borders or lakes. Raster data can also be imported into ROAD but are treated differently. In its simplest form, a raster consists of a matrix of cells (or pixels) organized into rows and columns (or a grid) where each cell contains a value representing information such as temperature or elevation. Raster data include digital aerial photographs, digital elevation models, imagery from satellites, digital pictures, or scanned maps. In ROAD the spatial data are supported using a PostGIS enhancement. PostGIS is a spatial database extender for a PostgreSQL object-relational database. PostGIS adds support for geographic objects allowing location queries to be run in SQL.

ROAD was generally designed assuming two data sources, that is, primary and secondary. Primary data are produced by the ROCEEH researchers during field work, laboratory and computer analysis or museum collection work (Fig. 4). However, the lion's share of the data are derived from publications (secondary data). We review the available literature concerning early human find locations and related data. The geographical emphasis in the first years was concentrated on Africa, and has now shifted towards Eurasia. We analyse publications and enter the relevant information according to a specific data entry procedure. Since data quality is of paramount importance, we keep track of all information sources. For our case studies this information is entered into the database at a higher synthetic level.

The database architecture and technology

ROAD on the back-end

The ROAD system is easily accessible, allowing users to communicate with the system through an internet browser to: 1) use and manage data; 2) import and export data; 3) query and analyze data; and 4) visualize data. Fig. 5 illustrates the ROAD architecture on the back-end.

The entire ROAD system is implemented on a server running with a FreeBSD operating system and an Apache Webserver. Through an internet browser the user operates with the PostgreSQL database via a WebGIS interface based on PostGIS 2.0 (www.postgis.net) and MapServer (www.mapserver.org). The MapServer client (www.mapserver.org), OpenLayers (www.OpenLayers.org) and QGIS Server systems (www.QGIS.org) together with Google's application programming interface (API) and Java Scripts allow data visualization and manipulation on different levels using the services offered by the MapServer. Services like Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS) or Styled Layer Descriptor (SLD) guarantee the visualization and manipulation of the spatial data using Open GIS Consortium (OGC) standards and protocols.

Fig. 4: Some aspects of primary data production, from data collection in the field to laboratory analysis and computer processing.

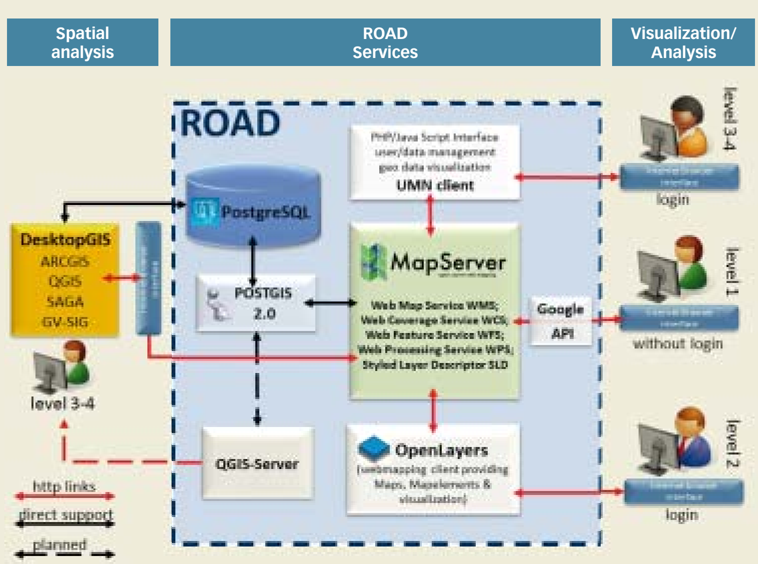


Fig. 5: ROAD back-end structure showing a variety of implemented software packages.

ROAD on the front-end

The ROAD systems are accessible in four different ways through a Web browser using:

1. **Open Access Interface** (Fig. 6). This is the easiest way to access the ROAD system through the implementation of Java scripts and a Google map API (Fig. 7). This portal allows for a semi-dynamic visualization of the data, accessing the system without the need to log in.
2. **WebGIS Interface** (Fig. 8). At hierarchy levels 3 and 4, authorized researchers have full access to the analytical and visualization functions of ROAD (Fig. 2). On these levels, data can be accessed using an OpenLayers interface as well as respective Java and php scripts. The internal GIS functions are provided by the Web Map Service (WMS).
3. **Developer-WebGIS Interface.** Full access is provided for the ROCEEH team (hierarchy level 4). The Developer-WebGIS allows the generation of maps and provides full spatial data management functionalities. This interface is realized using UMN MapServer, Java and php scripts.
4. **Alternatively, the PostgreSQL database** may be linked directly to an external desktop GIS system (e.g., SAGA-GIS, ArcGIS, QGIS, GRASS, GV-SIG) or indirectly through the MapServer.

After elaboration of the data, the resulting information can be re-imported into the system (see Fig. 5). Moreover, spatial data can be exchanged via a URL link provided by the MapServer or media, such as flash drives, CDs or DVDs.

Exploring the system at a glance

Access to the ROAD system is provided through the ROCEEH website (www.roceeh.net). Within the folder "ROAD" you will find a link to ROAD as illustrated in Fig. 6. As shown in Fig. 5 the ROAD system has three front-end portals. The simplest access to ROCEEH data is via the Open Access Interface for the public and automatically opens when you click on ROCEEH's website link to ROAD (Fig. 7). The front-end portal permits simple querying of information contained in ROAD. You can select a certain time period by entering numeric values or using the global chonostratigraphical correlation table. Subsequently, a query of available sites for a selected time period can be performed. This request delivers a map showing the spatial distribution of human remains, fauna, plant remains, lithics, symbolic artifacts or organic tools.

Fig. 7 shows the spatial distribution of all selected assemblages contained in ROAD between 300,000 and 500,000 years ago. In the left corner of the ROAD window there is a login button. By clicking on this button, you arrive at the internal user interfaces: i) WebGIS based on Open Layers; and ii) Developer-WebGIS using the UMN MapServer. To explore the WebGIS interface you can use a public login with limited access (userid: public; password: roceeh) (Fig. 7). Here you can explore selected maps that are available to the public user only. The Developer-WebGIS (Fig. 8) that provides you with simple GIS functionalities requires user registration.

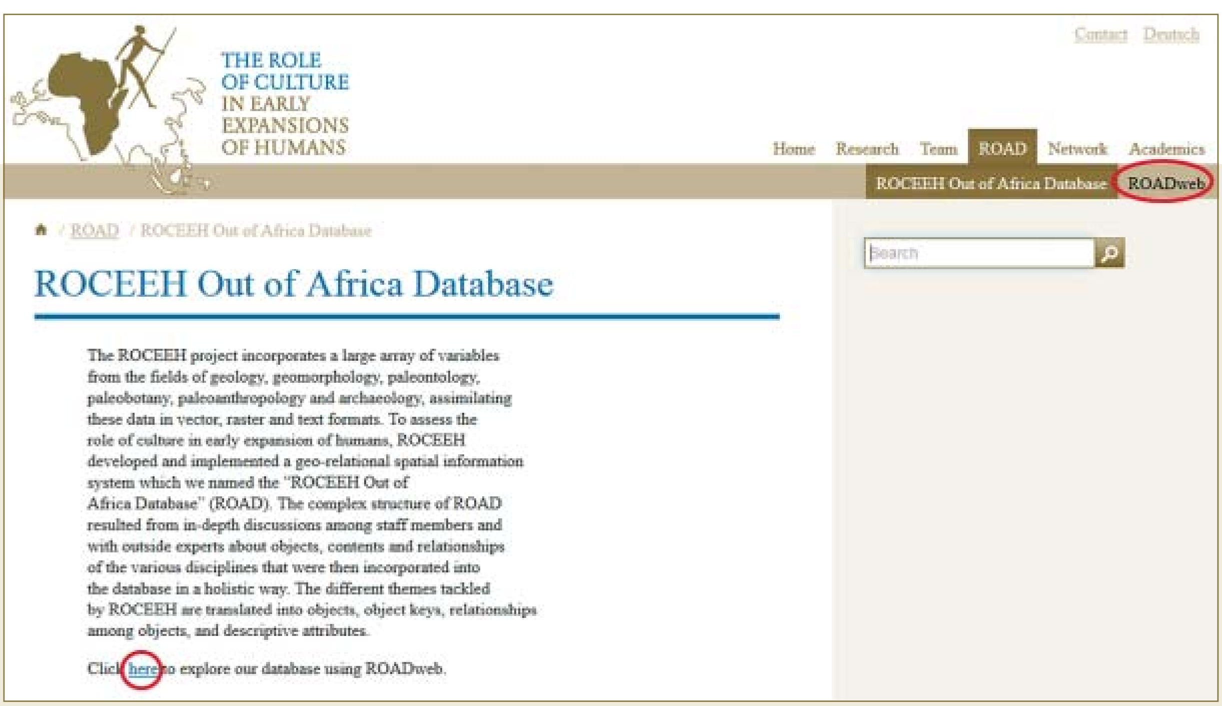


Fig. 6: ROAD access from the ROCEEH website. Click one of the links marked with red circle to get access

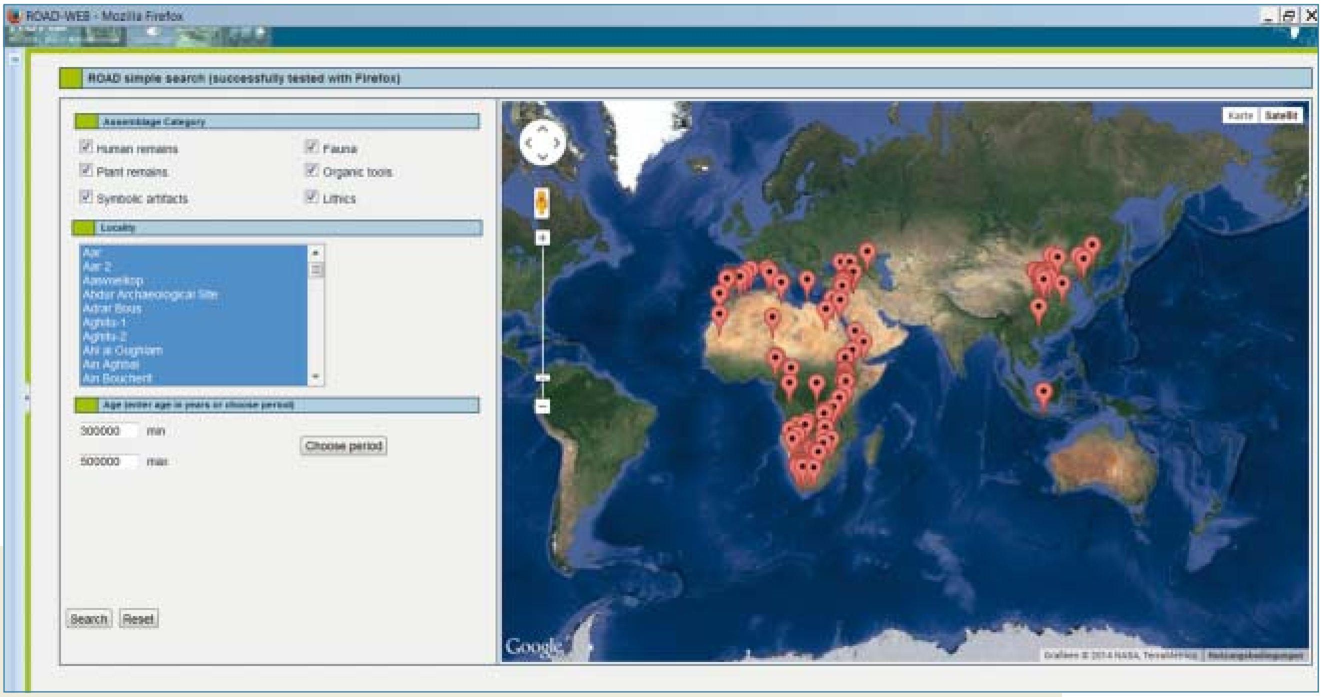


Fig. 7: Open access front-end portal using Java Scripting, Google Maps application programming interface.

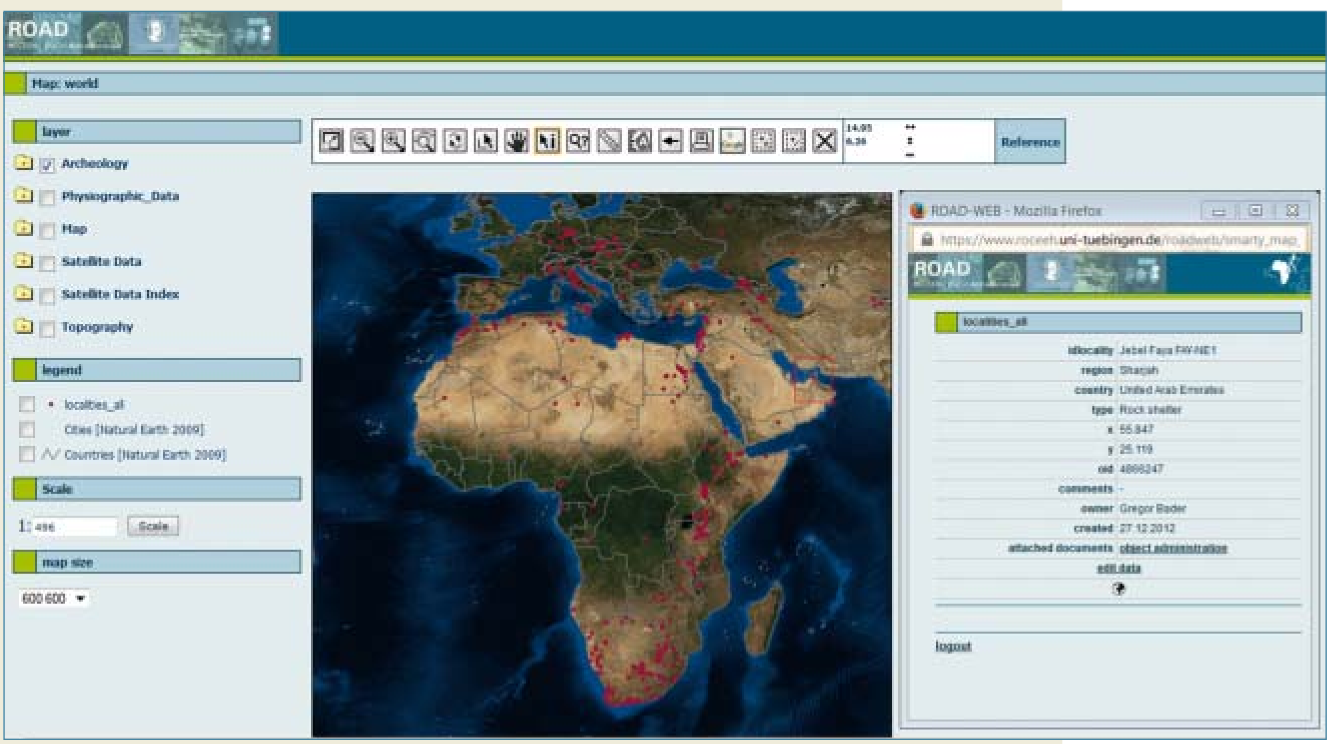


Fig. 8: The Developer-WebGIS front-end portal with limited access (login).

With the Developer-WebGIS Interface, the data management menu offers additional options as presented in Fig. 9. Apart from WebGIS functionalities, spatial data can be imported in different vector and raster formats. Moreover, the spatial data can be visualized and edited. To enter information about archaeology, paleofauna, paleobotany, paleoanthropology or literature, we provide workflows that guide the user through data input and editing.

A key functionality in the main menu is the utility to query data. We offer users a list of precompiled queries, but you can also code your own query in Structured Query Language (SQL). Fig. 10 illustrates a precompiled query to find all hominids between 1.2 and 1.6 Ma. The tables involved in this query are locality (attributes: id locality, country, region, x,y) and publication describes human remains (attributes: genus, species). The results of the query can be visualized on the WebGIS or stored as csv files (Fig. 10).

Fig. 9: Main data management menu available through the Developer-WebGIS Interface.

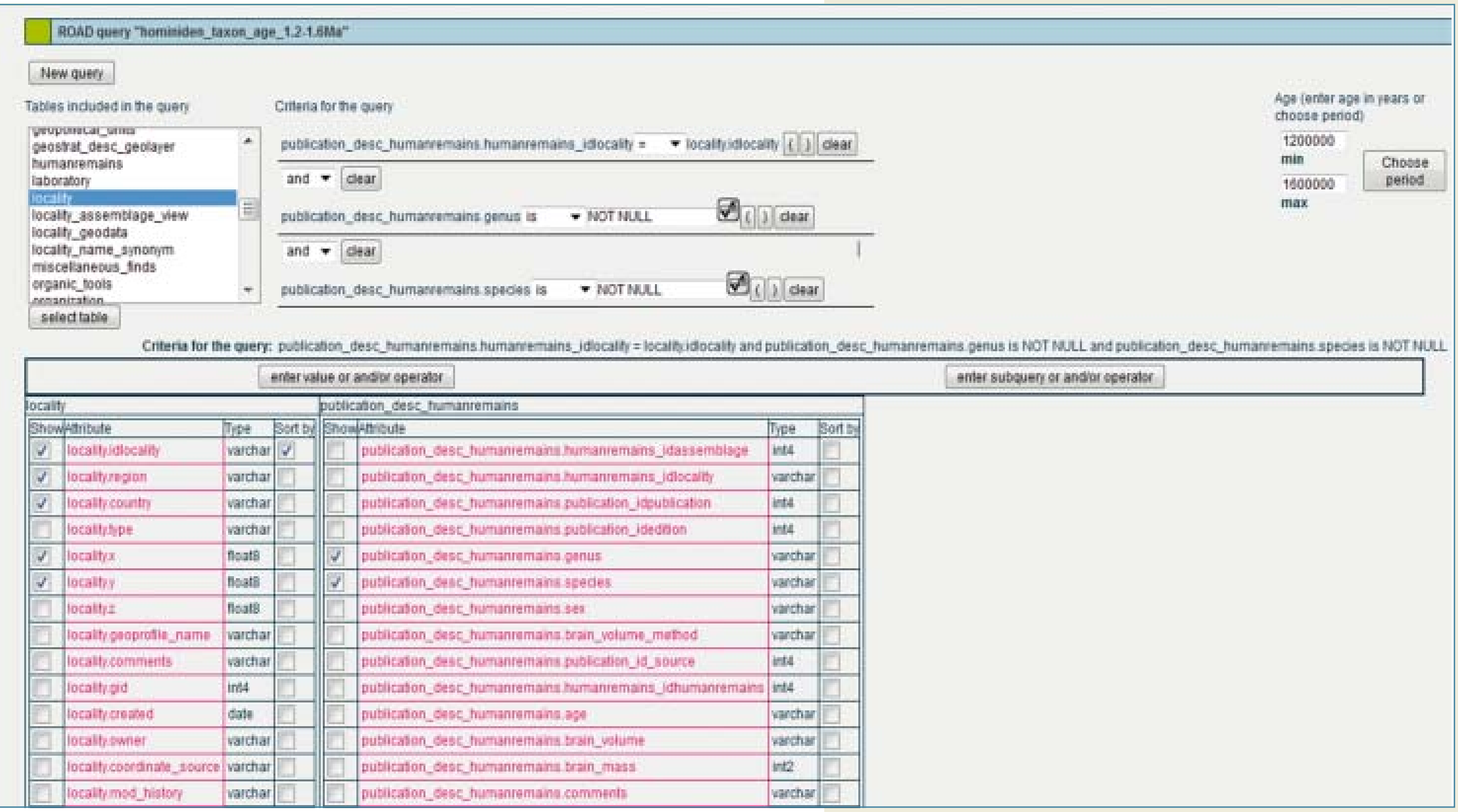
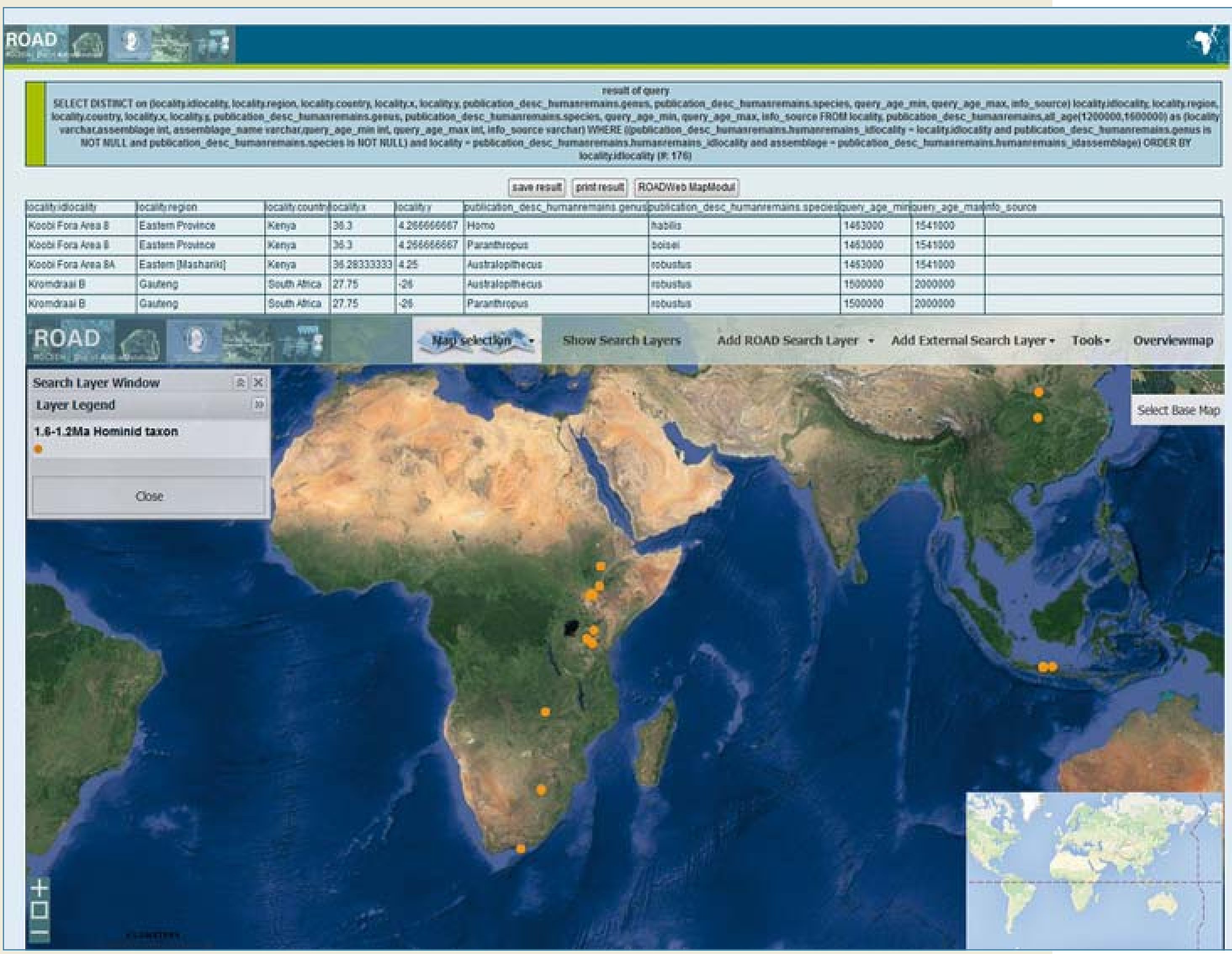


Fig. 10: Data query tool showing a precompiled query to find all hominids between 1.2 and 1.6 Ma.

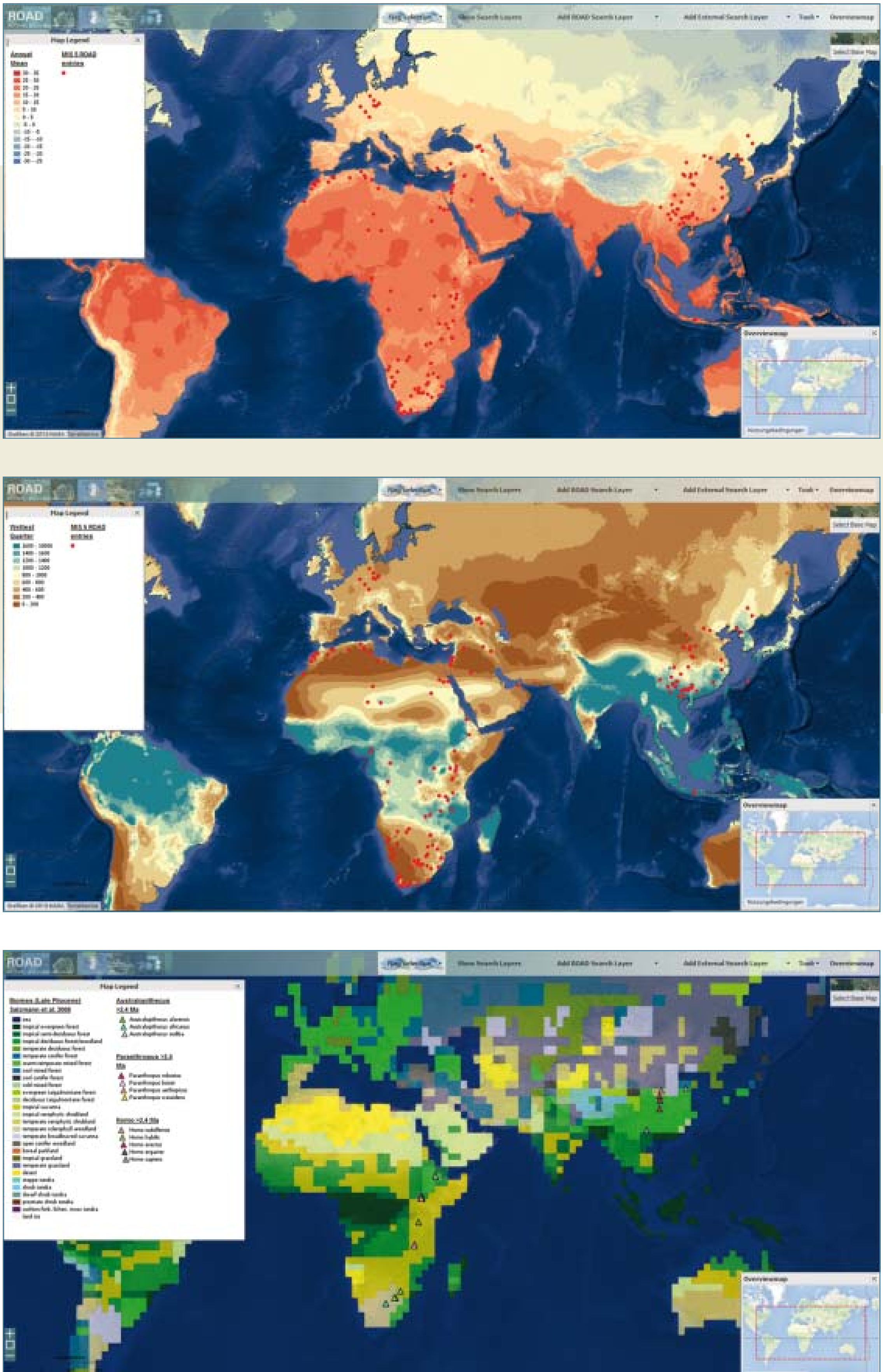
Visualization of additional environmental data

Figures 12 a, 12 b and 12 c show some examples of data visualization within the WebGIS front-end of ROAD. Archeological find locations can be displayed with different background information such as topography, bathymetry, vegetation, geological settings or climatic aspects, such as mean annual temperature of MIS 5 or precipitation of the wettest quarter (Fig. 12 a, b). Thus, we can place archeological and paleontological information in an environmental context (Fig 12 c). The relevant physiographic information on topography, bathymetry, geology, soils, vegetation and climate is stored in ROAD and available at different spatial resolutions and extents, as well as for different time slices through the WebGIS front-end.

Fig. 11: Results of the query for hominids between 1.2 and 1.6 Ma shown as a table (above) and on WebGIS map (below).



a)



b)

c)

Fig. 12: WebGIS front-end showing: **a)** archeological sites and mean annual temperature for MIS 5; **b)** precipitation of the wettest quarter for MIS 5 (source www.worldclim.org); and **c)** find locations of the genuses *Homo, Paranthropus* and *Australopithecus* showing the mega-biomes of the later Pliocene.

Fig. 13: Linking external NEOTOMA data to ROAD via the WebGIS front-end. Explore this tool yourself at: www.roceeh.uni-tuebingen.de/roadweb/map_modul/index.html.

Security in the ROAD system

Security in ROAD covers a broad range of information and concerns the protection of the database system, the stored data, ROAD applications and functions, the server and associated network links against compromises of confidentiality, integrity and availability. It involves various types of technical, procedural, administrative and physical controls.

The user management system allows the attribution of specific rights in terms of data access and procedures. Hence, the ROAD team specifies who has access to data and what kind of processes the user is allowed to apply. As a consequence, user and property rights are guaranteed. Access to higher levels of data editing and data analysis is strictly related to authentication procedures and administrative control. Every user action is logged and can be traced. We conduct backups of the entire ROAD system on a regular basis so that even in the event of damage, the system can be restored.

Linking ROAD with other networks

The above described WebMap interface is created in a flexible way, so that anyone can visualize their own data within the ROAD context. Data in ROAD can be combined with data from other sources, for example, from other databases, for analytical purposes (Fig. 13). In this case, the map surface at the front-end functions as an access tool. Such links are permanently established by queries sent to remote databases. Retrieved data are then displayed on a map surface. Data from multiple sources may be combined by creating several layers.

References

Foertsch, R., 2006. ARACHNE – Datenbank und kulturelle Archive des Forschungsarchivs für Antike Plastik der Universität Köln und des Deutschen Archäologischen Instituts. URL: http://arachne.uni-koeln.de. Accessed 4 May 2015.

Kondo Y., Kadowaki, S., Kato, H., Naganuma, M., Ono, A., Sano, K., Nishiaki, Y., 2011. Network computing for archaeology: A case study from the 'Replacement of Neanderthals by Modern Humans' Database Project. Computer Applications in Archaeology Conference 2011- Revive the Past. Proceedings of the 39th Computer Applications and Quantitative Methods in Archaeology Conference, Beijing, China, 12–16, April 2011.

Semal, P., Cornelissen, E., Cauwe, N., 2004. MARS: Multimedia Archaeological Research System. Notae Praehist. 24, 203–208.

Weniger, G.-C., Döllner, J., Macchiarelli, R., Mandel, M., Mayer, P., Radovcic, J., 2005. TNT. The Neanderthal Tools and Nespos. Proceedings of the 33rd Computer Applications and Quantitative Methods in Archaeology Conference, Tomar, Portugal, 21–24 March 2005. IPT, Tomar.

ROCEEH Out of Africa Database (ROAD)

EDITORIAL Prof. Dr. Volker Hochschild
Dr. Michael Märker

CONTACT ROCEEH Research Center Tübingen
Rümelinstr. 23,
D–72070 Tübingen, Germany
www.roceeh.net
michael.maerker@ggi.uni-tuebingen.de

THE ROLE
OF CULTURE
IN EARLY
EXPANSIONS
OF HUMANS

JULY 2015

THE ROLE
OF CULTURE
IN EARLY
EXPANSIONS
OF HUMANS