

EAA BUDAPEST 2022:

FAIRly Front-loading the Archive:

Moving beyond Findable, Accessible and Interoperable to Reuse of Archaeological Data

3 Sept. 2022



**THE ROLE
OF CULTURE
IN EARLY
EXPANSIONS
OF HUMANS**

The ROCEEH Out of Africa Database: A Case Study in Reusing Archaeological Data Sensibly

Andrew Kandel, Zara Kanaeva, Chris Sommer & Miriam Haidle

The ROCEEH Team



**HEIDELBERGER AKADEMIE
DER WISSENSCHAFTEN**

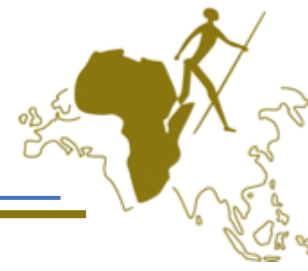
Akademie der Wissenschaften
des Landes Baden-Württemberg

SENCKENBERG
world of biodiversity

**EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN**



The ROCEEH research center



Long-term grant

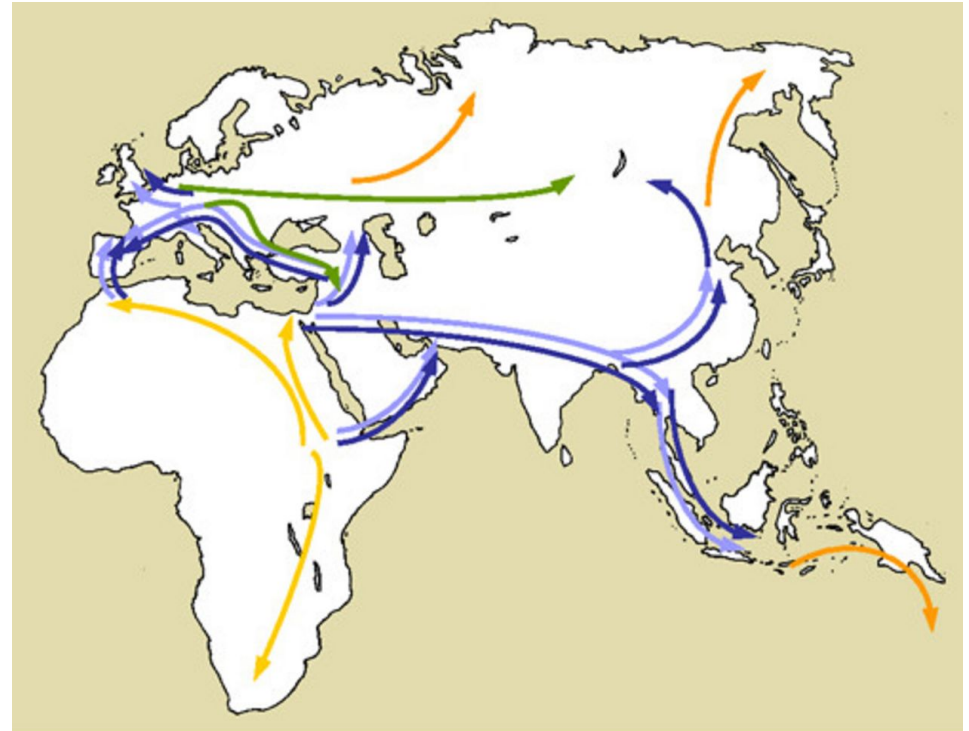
- 2008-2027

Study expansion of humans in Africa & Eurasia between 3 million & 20,000 years ago

- Culture
- Habitat
- Distribution

Interdisciplinary perspective with researchers from

- Archaeology
- Anthropology
- Paleobotany
- Geography...



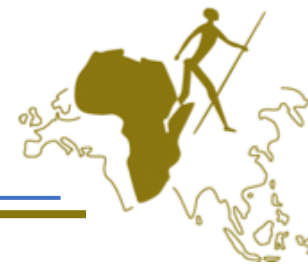
**HEIDELBERGER AKADEMIE
DER WISSENSCHAFTEN**
Akademie der Wissenschaften
des Landes Baden-Württemberg

EBERHARD KARLS
**UNIVERSITÄT
TÜBINGEN**



SENCKENBERG
world of biodiversity

The ROCEEH Out of Africa Database



- >2,200 Paleolithic sites *
- > 20,000 assemblages *
- Culture, humans, fauna, plants
- Georelational database
- Uniform taxonomies

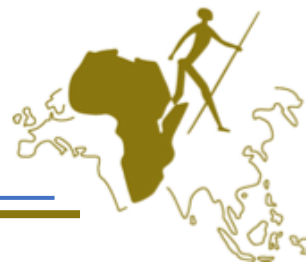
A screenshot of the ROCEEH ROAD simple search web interface. The interface is titled 'ROAD simple search (developed with Firefox, but suitable for other browsers)'. It features a search bar for 'ROAD Summary Data Sheet for locality' with a 'PDF' button. Below the search bar, there are several filter sections: 'Assemblage Category (or logic combination is used as default)' with checkboxes for 'Human remains', 'Plant remains', 'Symbolic artifacts', 'Organic tools', 'Fauna', and 'Lithics'; 'Region (for multiple selection use Ctrl-key)' with a dropdown menu listing regions like 'Africa - Central Africa', 'Asia - Caucasus', etc.; and 'Age (enter age in years or choose period)' with input fields for 'min' and 'max' and a 'Choose period' button. A 'Map' section on the right shows a world map with numerous red dots indicating site locations, primarily concentrated in Europe and Africa. The map includes a 'Satellite' view option and a 'Google' logo. At the bottom, there are 'Search' and 'Reset' buttons, and a 'Terms of Use' link.

ROAD Database:

www.roceeh.uni-tuebingen.de/roadweb/

* As of 19 July 2022

Methods



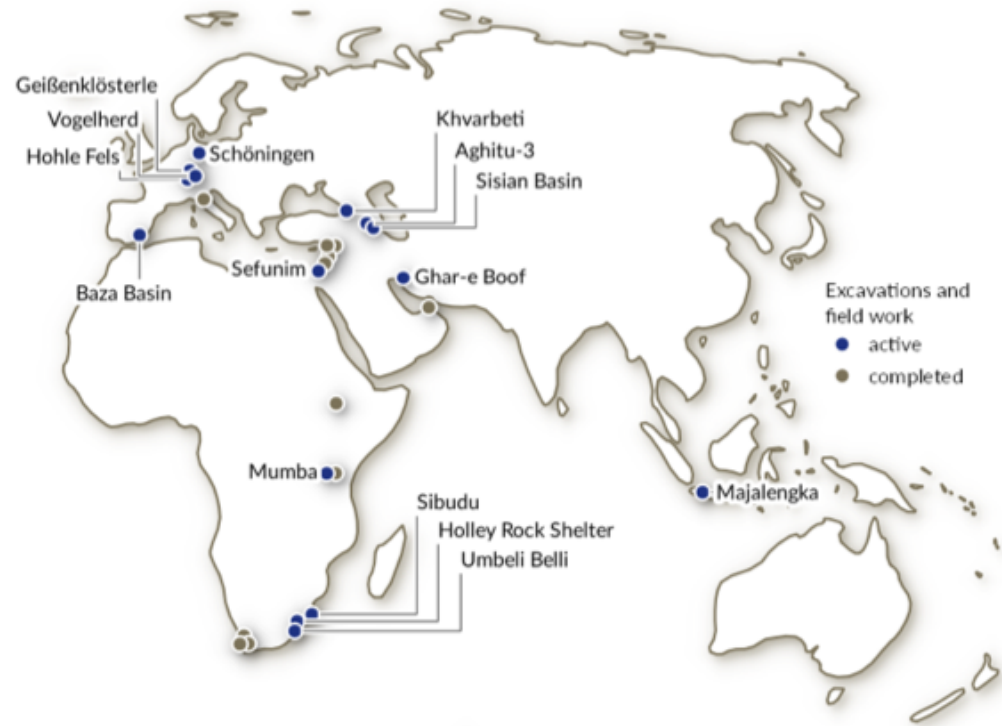
How do we distill the most informative elements using the collective experience of a multidisciplinary team of archaeologists, anthropologists, paleontologists, botanists, geographers, and data scientists?

- Overcome differences in research traditions by choosing **least common denominator**
- **Select attributes** regardless of when/where collected and published
- **Harmonize information** to ensure smooth flow of data entry

Data acquisition



- Digitalization by paid research assistants using data entry masks
- >4,600 sources * in >10 languages
 - Articles
 - Books
 - Theses
 - Excavation reports
 - Own research
- Georeferencing
- Standardization & Thesauruses
 - Raw materials, tools, etc.
 - Tool groups
 - Cultural periods



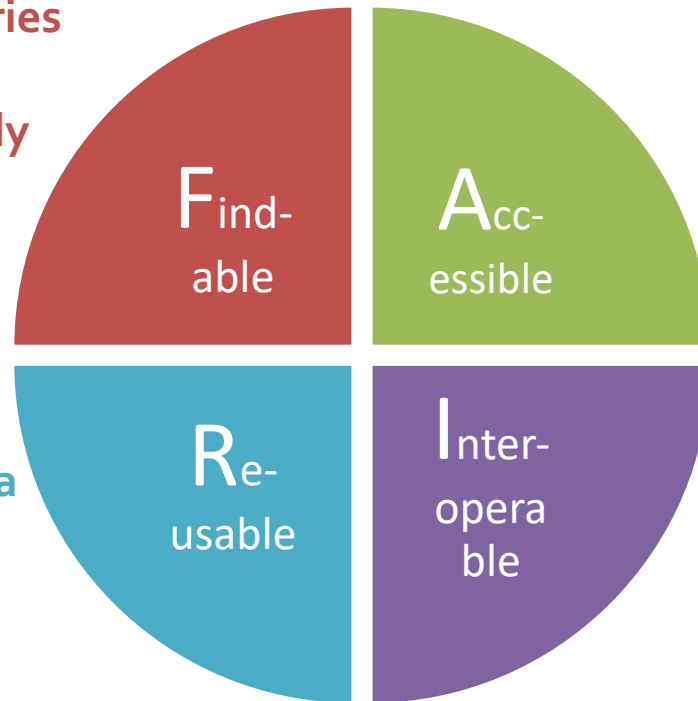
ROCEEH research & excavations

* As of 19 July 2022

ROAD – FAIRness Assessment 2018



- ✗ Not listed in repositories
- ✗ No Unique Identifier
- ✗ Internal metadata only

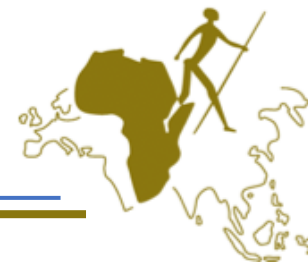


- ✓ Documentation of data provenance
- ✗ Licensing unclear
- ? Community standards

- ✓ SQL Query Builder
- ✗ Not machine accessible
- ✓ Non-parametric APIs
- ✓ 4-levels of user access
- ✗ No Long Term Storage

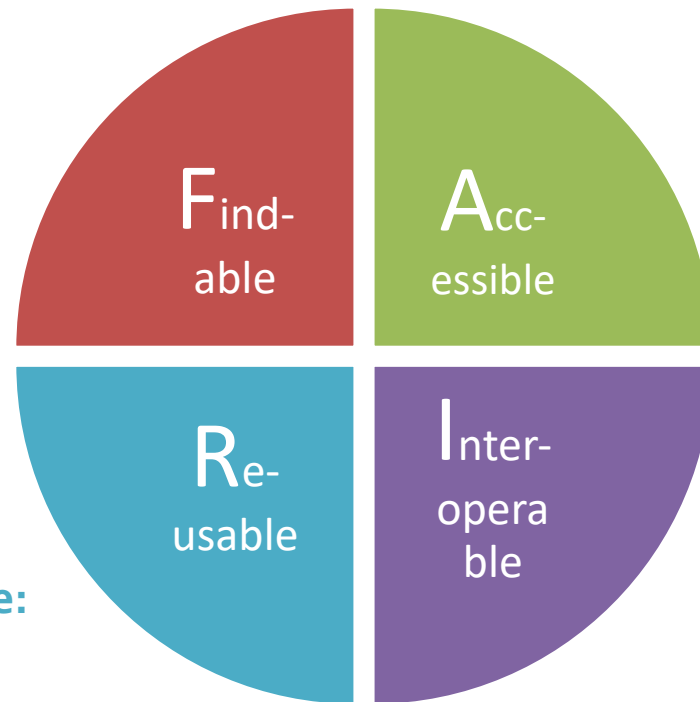
- ✓ SQL-based
- ✓ Connected to other DBs
- ✗ No standardized vocabularies

ROAD – FAIRness Assessment 2022



- ✓ Listed in repositories:
ARIADNE, re3data
- ✓ Use of URIs
- ✓ ARIADNE metadata
scheme

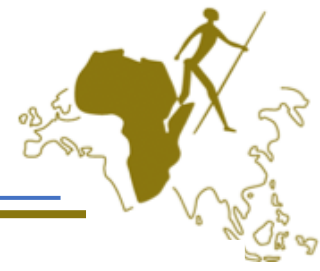
- ✓ Documentation of data
provenance
- ✓ Open Standard License:
CC BY-SA 4.0
- ? Community standards



- ✓ SQL Query Builder
- ✓ Machine-actionable
SPARQL endpoint
- ✓ Non-parametric APIs
- ✓ 4-levels of user access
- ✗ No Long Term
Storage

- ✓ SQL and SPARQL-based
- ✓ Connected to other DBs
- ✓ Semi-standardized
vocabularies

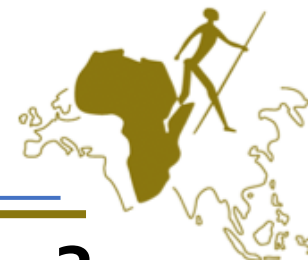
Focus of talk – reuseability



- ✓ Documentation of data provenance
- ✓ Open Standard License: CC BY-SA 4.0
- ? Community standards



Licensing

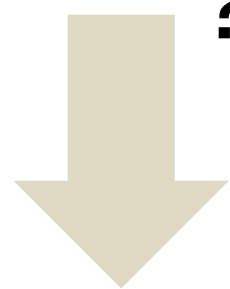


2008

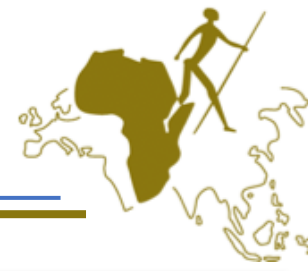
- We started with an open attitude, but...
- Own User Agreement and Data-Use Policy
- No open licenses
- Suggested guidelines for citation and co-authorship

July, 2022

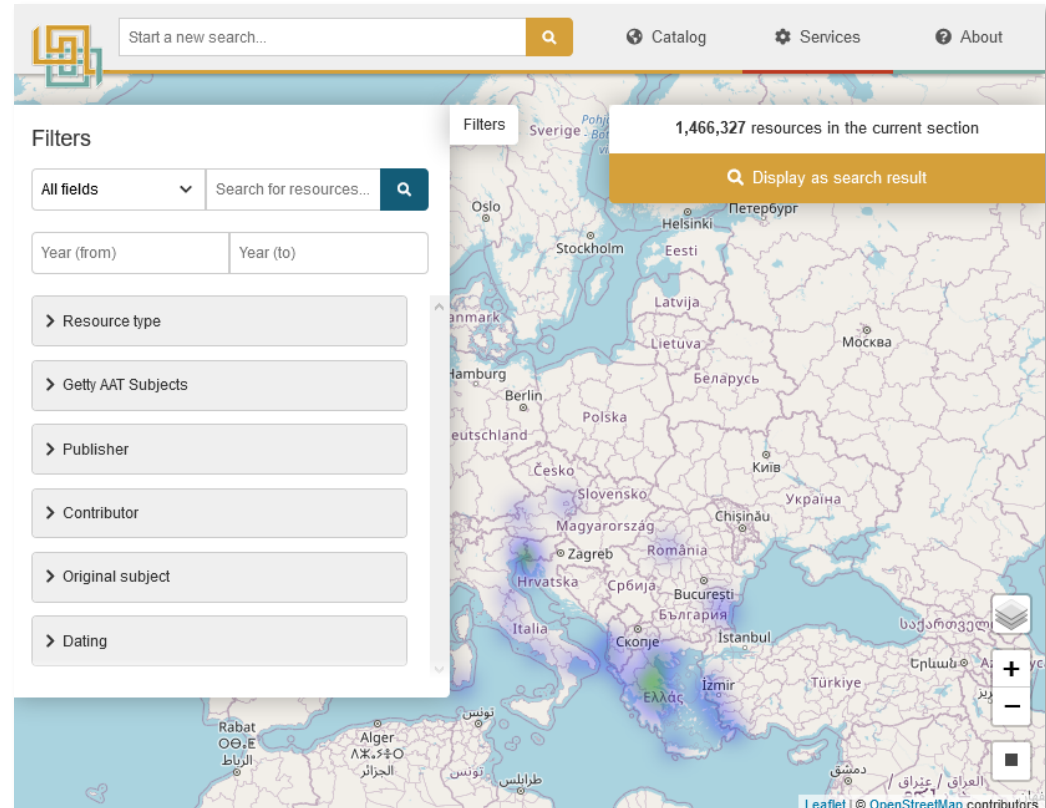
- **CC BY-SA 4.0**



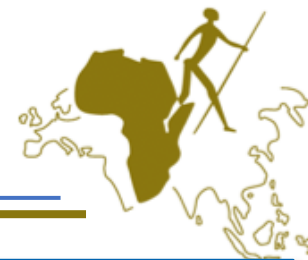
Cooperation with



- Match ROAD assemblages to **Getty AAT** objects
- Define cultural periods (space, time) in **PeriodO**
- Specify exports
- Translate ROAD Data into ARIADNE ontology **AO-Cat** with **Vocabulary Mapping Tool**
- **Regular export** through ARIADNE Aggregation Pipeline, starting in Sept, 2021



Wikipedia



- Provide geodata for interactive web maps on Wikipedia
- Target: > 30 Wikis about archaeological cultures and periods in multiple languages



ROAD

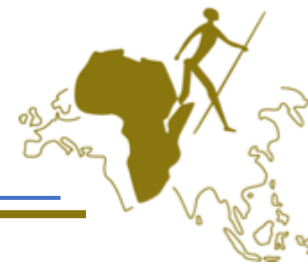
SPARQL
Endpoint

Wikibot

Wikimedia
Commons

Wikipedia

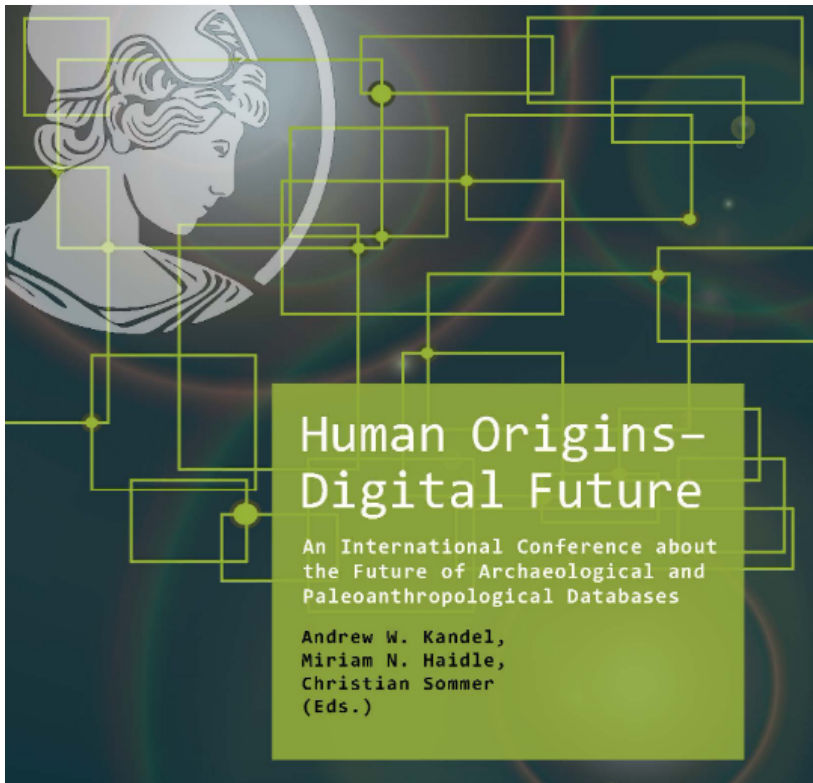
Personal and institutional frameworks



ROCEEH Conference

Human Origins—Digital Future, an international conference about the future of archeological and paleoanthropological databases

Miriam N. Haidle, Michael Bolus, Angela A. Bruch, Christine Hertler, Volker Hochschild, Zara Kanaeva, Christian Sommer, Andrew W. Kandel ✉



Haidle et al., *Evol. Anthropol.* (2020)
Propylaeum Verlag (2022)



CfAS

Coalition for Archaeological Synthesis

To understand how migrations affect human securities, look to the past

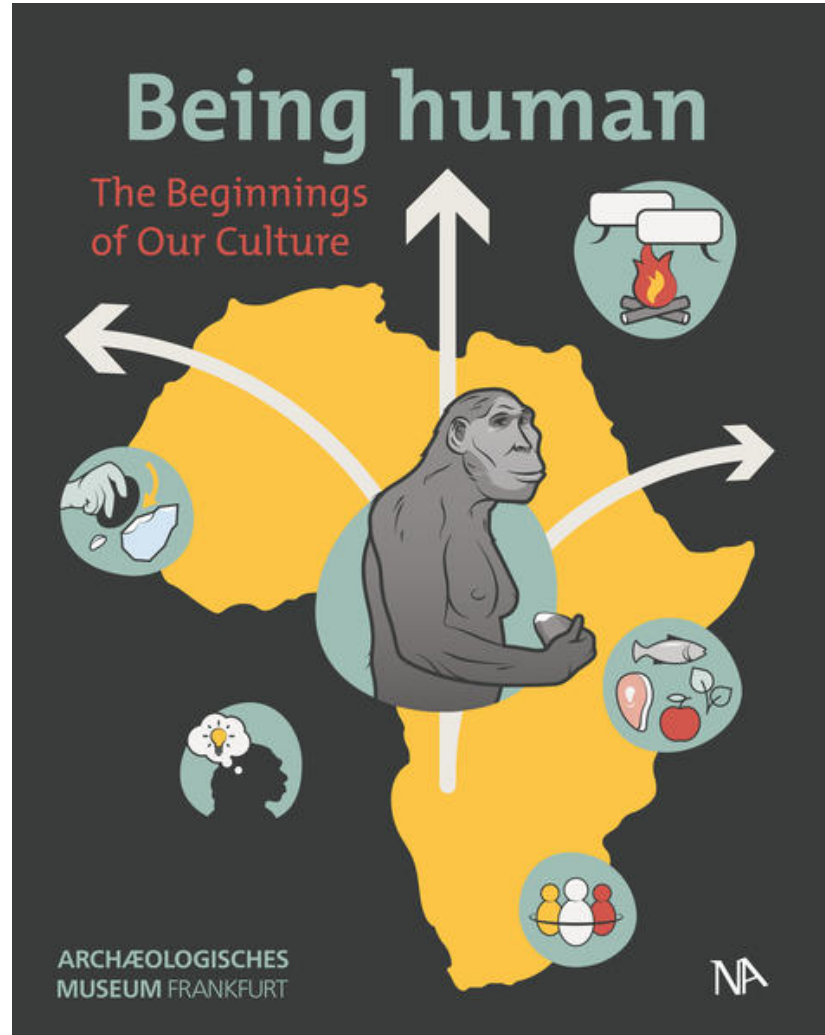
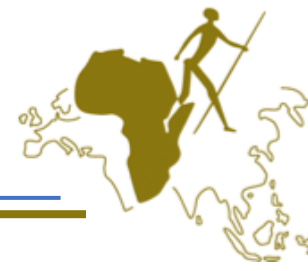
Jeffrey H. Altschul^{a,b,1}, Keith W. Kintigh^{a,c}, Mark Aldenderfer^d, Elise Alonzi^e, Ian Armit^f, Juan Antonio Barceló^g, Christopher S. Beekman^h, Penny Bickleⁱ, Douglas W. Bird^j, Scott E. Ingram^k, Elena Isayev^l, Andrew W. Kandel^m, Rachael Kiddeyⁿ, Hélène Timpoko Kienon-Kabore^o, Franco Niccolucci^p, Corey S. Ragsdale^q, Beth K. Scaffidi^r, and Scott G. Ortman^s



archsynth.org
Altschul et al., *PNAS* (2020)

Museum exhibitions

ARCHÄOLOGISCHES
MUSEUM FRANKFURT



Catalog (English & German) by NA Verlag

Coding da Vinci



{CODING DA VINCI} BADEN-WÜRTTEMBERG 2022

07.05.2022 - 24.06.2022

Germany's largest annual **Hackathon** aims at creative engagement with open cultural data

ROCEEH contributed its **spatio-temporal dataset of archaeological sites** to this year's competition hosted by the state of Baden-Württemberg

<https://www.roceeh.uni-tuebingen.de/cdv/TimeFlies/>



Conclusion



Motivation

- FAIRness as a condition for sustainable archiving and reuse

Implementation

- Engagement in ARIADNEplus, Heidelberg Academy of Sciences & Humanities, Working Group “eHumanities” of the Union of Academies
- Advanced training of key individuals (e.g. Knowledge Graphs, OpenHPI, etc.)
- FAIR development as a democratic project

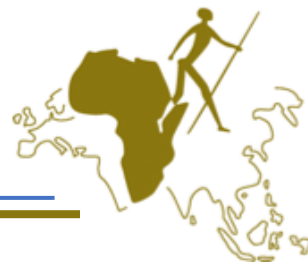
Results

- Sustainable data, personal and institutional networks, positive surprises

Unresolved questions

- Make ROAD more like Wiki – through open community platform or curated data?
- Archiving – through tDAR or similar agencies?
- Extension of ROAD after 2027 – through institution or foundation?

Thank you!



2,810,000 years ago



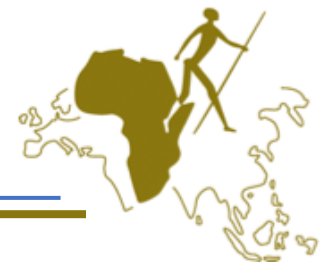
Lithic assemblages
from 3 Ma – 20 ka

- New assemblages
- Old assemblages
- Recurring assemblages

www.roceeh.net

www.roceeh.uni-tuebingen.de/roadweb/

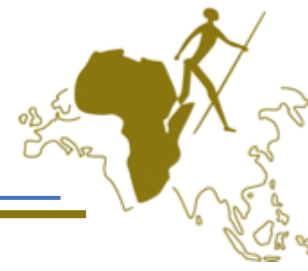
@ROCEEH_news



More Examples of Reuse

Studies reusing data for new aims

Applications to explore prehistory



<https://www.roceeh.uni-tuebingen.de/cdv/TimeFlies/>

Explore with Amelia

Archive with Otto

African beginnings

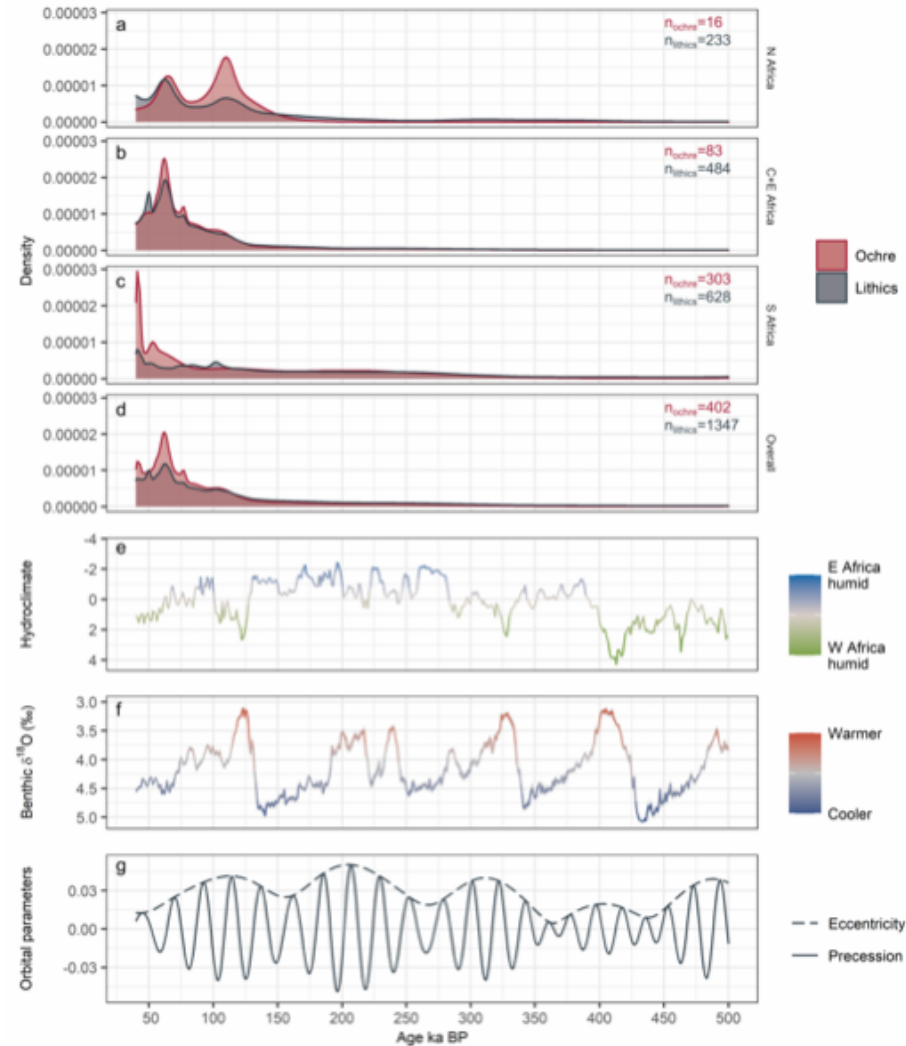
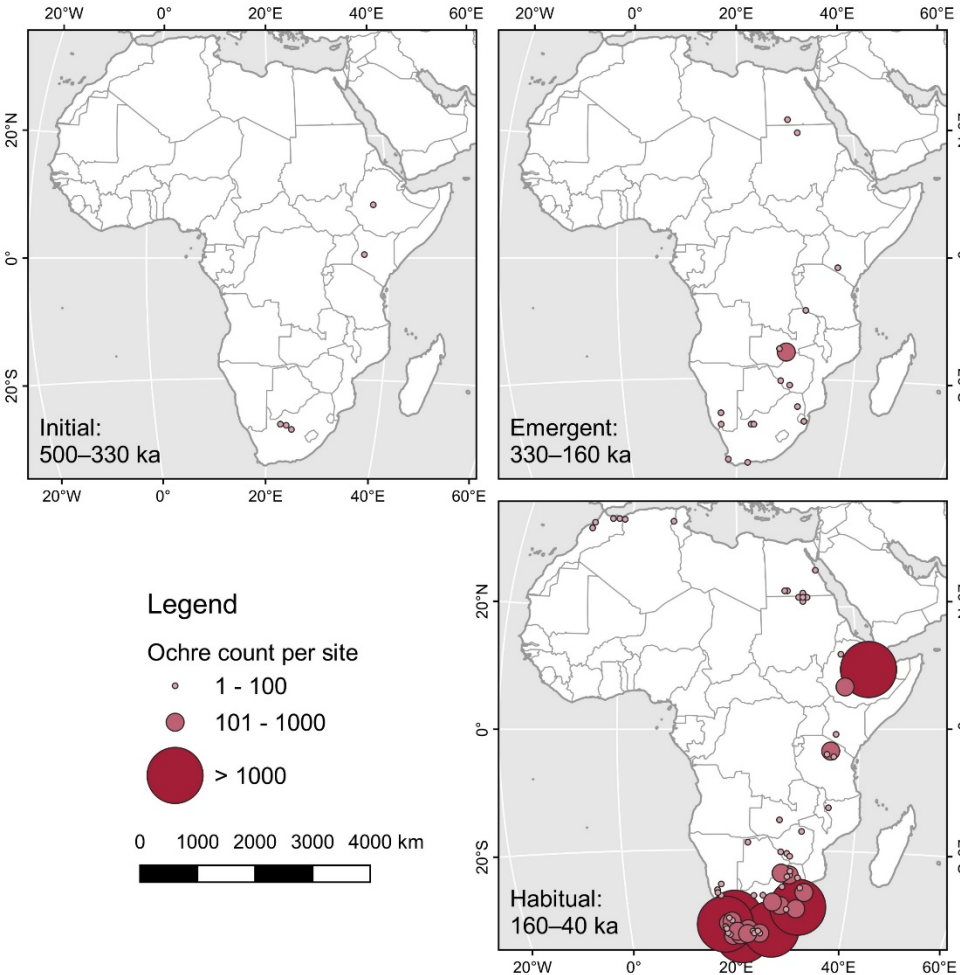
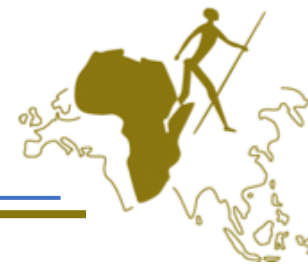
Expanding worlds

Instructions

FLY!



The emergence of habitual ochre use in Africa



The Use of Prehistoric 'Big Data' for Mapping Early Human Cultural Networks

Visualization of cultural similarities between archaeological assemblages from the ROAD Database

Middle Pleistocene & Middle Paleolithic

These maps show the technological and cultural similarities between time-overlapping archaeological assemblages throughout the Pleistocene. The lines represent the similarities as estimated with the Jaccard Index from 32 attributes comprising e.g. lithic and organic tool groups, symbolic artefacts, combustion and burial features. We used a weighted function based on Euclidean distance and the similarity measure to visualize the cultural links. Brighter colors indicate higher similarities and higher densities of connections.

The ROAD Database is one of the largest collections of archaeological data integrating sites, assemblages, chronologies, artifacts, human remains and environmental indicators from 3 Million to 20,000 years ago. This study, as of Nov. 2021 comprises ~16,000 assemblages from ~2,100 sites. However the resulting maps presented here should be treated with caution, because data collection from published literature in seven languages is still ongoing and the archaeological record grows constantly. We present two cultural periods, Middle Paleolithic and Middle Stone Age, as well as seven technocomplexes.

Middle Paleolithic



Middle Stone Age



Similarities between Middle Paleolithic and Middle Stone Age

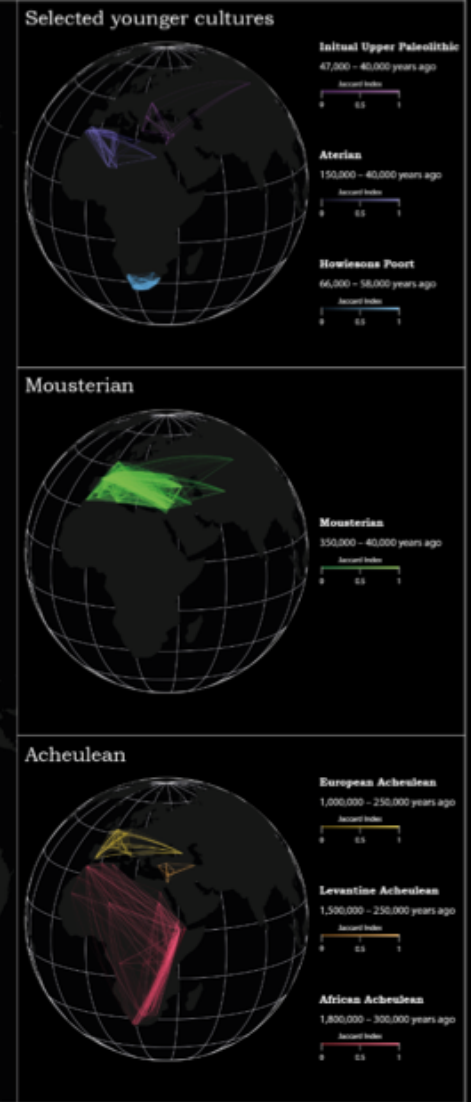
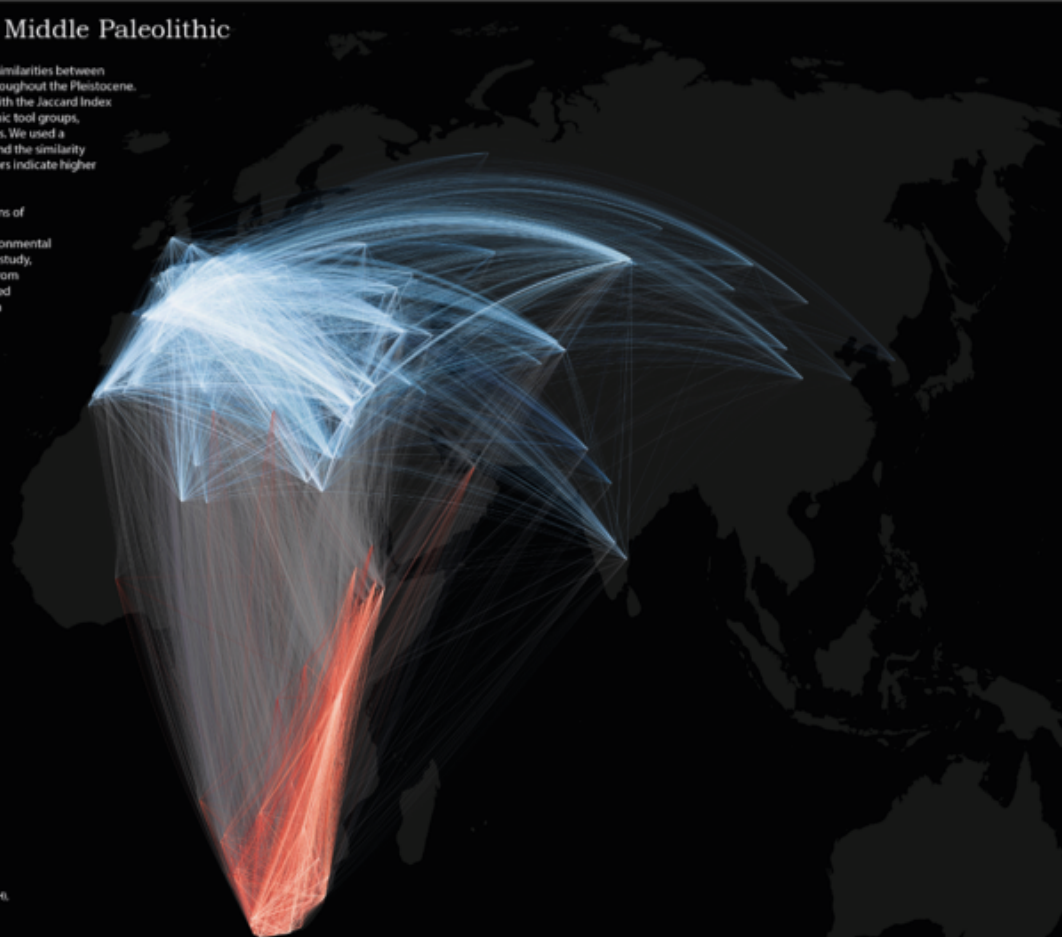


Map design: Sommer, C.^{1,2}, Kandel, A.W.¹, Hochschild, V.¹
¹The Role of Culture in Early Expansions of Humans (ROCEEH), Heidelberg Academy of Sciences and Humanities
²Department of Geography, University of Tübingen
Contact: christian.sommer@uni-tuebingen.de

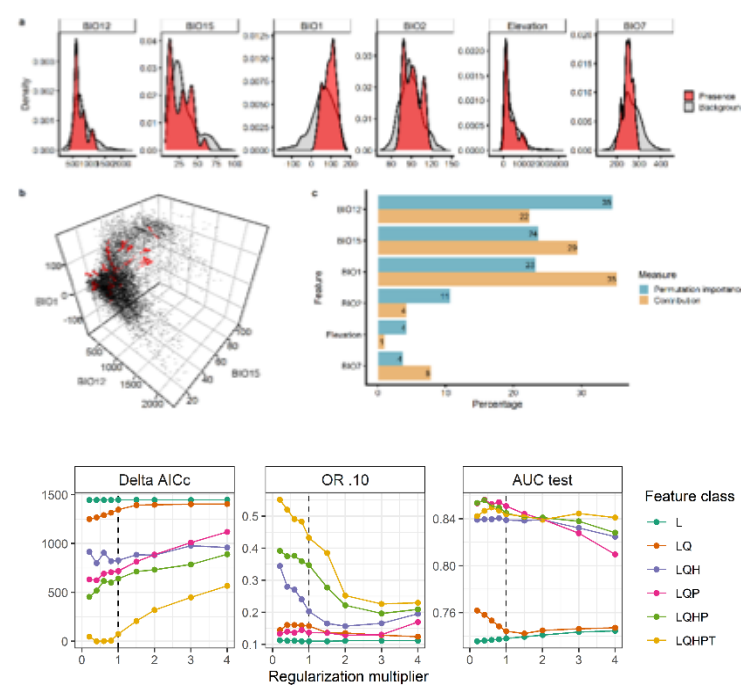
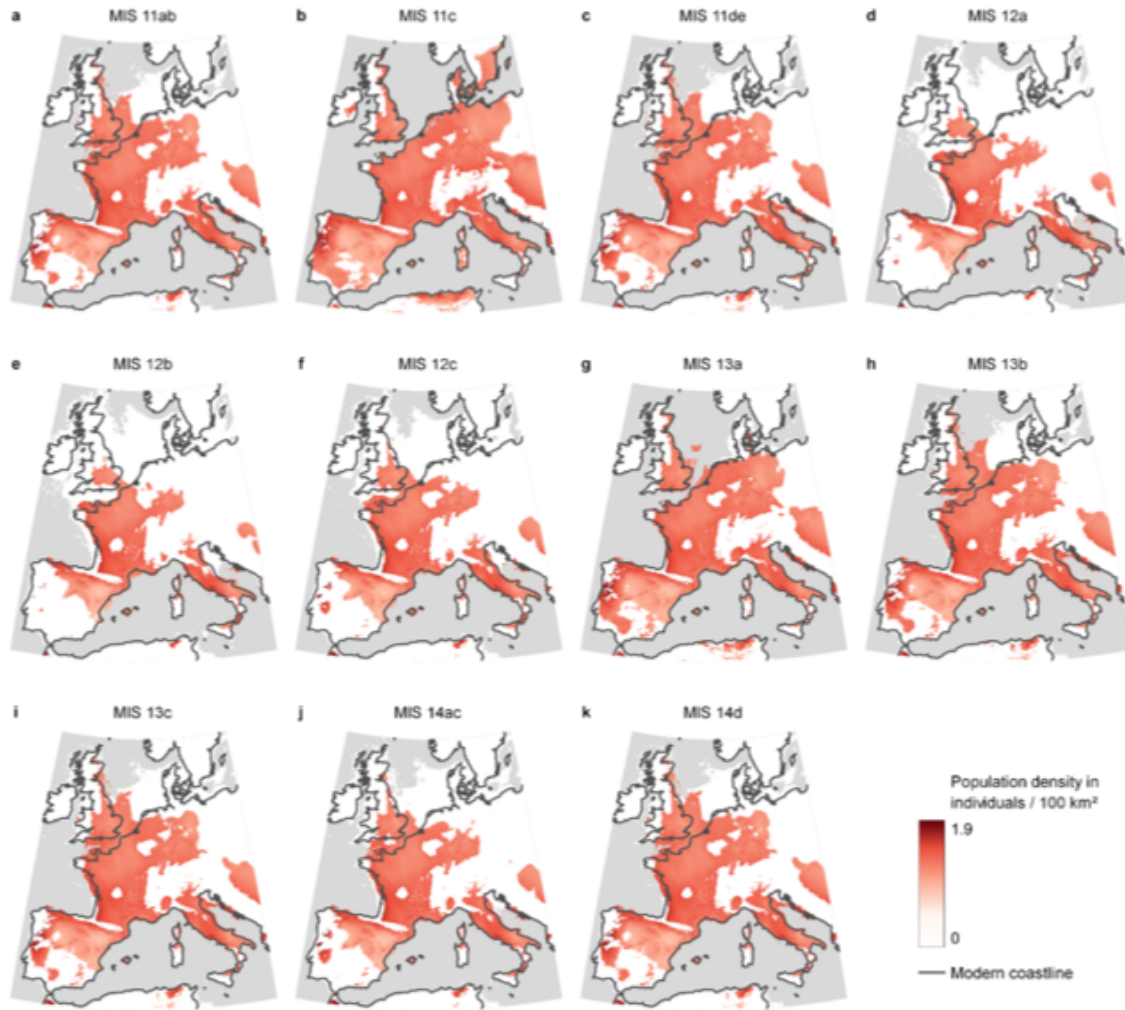
Projection: Winkel Tripel centered at 60°E
Datum: WGS 1984

Data source: ROCEEH Out of Africa Database (ROAD), www.roadh.net/ as of 2021-11-16, naturalisearthdata.com

Acknowledgements:
This research was conducted within the project "The Role of Culture in Early Expansions of Humans" (ROCEEH) funded by the Heidelberg Academy of Sciences and Humanities.



Sustainable human population density in Western Europe between 560,000 and 360,000 years ago

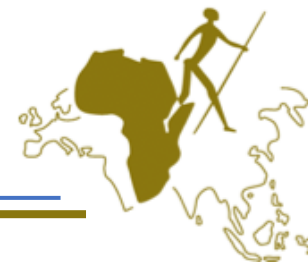


Some studies using reused data

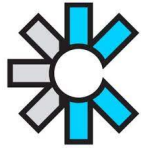


- MPI / Leipzig – Carrying capacity, demography and complexity using tool frequency as proxy for behavioral in Africa (W. Archer. 2020)
- Uni. Heidelberg – Development of ochre use during the Middle Stone Age of Africa (R. Dapschauskas, in press)
- MPI / Jena – Climate and environment vs. lithics in Africa (E. Scerri, E. Hallett, in review)
- SFB-806 / Köln – Demographic models through time using hominin and artifact data (A. Zimmermann, I. Schmidt, in prep.)

FAIR history of development



GENERAL TRENDS



2004
Open
Knowledge
Foundation



2009
Panton
Principles



Wilkinson
et al. 2016



Horizon 2020
Guidelines
DFG



2021 UNESCO
Recommendation
on Open Science



2008
Start of
ROCEEH

2009
First site entered
in ROAD

2018
Project
evaluation

2020
Workshop with
ARIADNEplus
&
ROAD Conference
HODiF 2020

2021
ARIADNE Mapping
&
Licensing

ROAD TRENDS

5 Star Open Data



General Information

Name: Aghia-Cave
 Synonyms: No synonyms available
 Country: Armenia
 Region: Gyumrik
 Coordinates: 39.51 N, 45.68 E
 Type: cave
 Summary: Aghia-Cave is situated at the base of a basalt outcrop along the Vayotsk River valley. This site yields the earliest evidence for modern human occupation in Armenia dated between 35-24 ka cal. Notable finds include shell beads, an eyed stone beads and other bone tools, plus many animal tools made from obsidian sourced up to 200 km away.

Assemblage: (for legend see test page)

Age: [Timeline]

Geological Profiles

Profile 1: [Table with 2 columns: Profile Name, Description]

```

1 assemblage.name;assemblage.category;assemblage.x;assemblage.y;assemblage.e;litic.piece.count;locality.id;locality.type;locality.x;locality.y;query_age_min;query_age_max;age_comments
2 Aar Layer 1 Archaeological Assemblage:miscellaneous finds, raw material, symbolic artifacts, technology, typology;;;4320;Aar 1;rock shelter;16.488989;-26.772622;30000;300000;Absolute date not available. Waiting for date through archaeological material;
3 Aar Layer 4-2 Archaeological Assemblage:miscellaneous finds, raw material, symbolic artifacts, technology, typology;;;1;Aar 1;rock shelter;16.488989;-26.772622;25000;40000;Absolute date not available. Waiting for date through archaeological material;
4 Aar 2 Layer 4-1 Archaeological Assemblage;raw material, miscellaneous finds, technology, typology;;;2066;Aar 2;rock shelter;16.468356;-26.728179;30000;300000;;
5 MSA hilltop rubble;raw material, technology, typology;;;709;Aar;veelkop;open air;28;-26-17;30000;300000;;
6 Abdur MSA assemblage;miscellaneous finds, raw material, technology, typology;;;1;Abdur Archaeological Site;open air;39.864369;15.151203;118000;132000;;
7 Abri Blanchard des Roches Exc. Didon Archaeological Assemblage:miscellaneous finds, raw material, symbolic artifacts, technology, typology;;;1499;Abri Blanchard des Roches;rock shelter;1.101003;45.000152;30000;430000;;
8 Abri Blanchard des Roches Sector 1 Archaeological Assemblage;raw material, technology, typology;;;1494;Abri Blanchard des Roches;rock shelter;1.101003;45.000152;30000;430000;;
9 Abri Blanchard des Roches Sector 4/5 Unit C Archaeological Assemblage;organic tools, raw material, symbolic artifacts, technology, typology;;;2566;Abri Blanchard des Roches;rock shelter;1.101003;45.000152;30000;430000;;
  
```

ARIADNE

Welcome

Browse the Catalog

houses

works (engineering w)

ROAD Query Export: CSV, static APIs - HTML

ARIADNE Mapping XML

NFDI? Other Initiatives?



(not used)

