NFDI4Objects
Research Data Infrastructure
for the Material Remains of
Human History

TRAIL3.3:

Workflow tool for archaeological experiments and analytics

Partner **Lead**: Alexandra Busch (Römisch-Germanisches Zentralmuseum – Leibniz Research Institute for Archaeology), Thomas Stöllner (Deutsches Bergbau-Museum Bochum, Ruhr-University Bochum)

Members: Susanne Greiff (Competence Center Archaeometry Baden-Württemberg, University of Tübingen), Ivan Calandra (Römisch-Germanisches Zentralmuseum – Leibniz Research Institute for Archaeology), Erica Hanning (Römisch-Germanisches Zentralmuseum – Leibniz Research Institute for Archaeology), João Marreiros (Römisch-Germanisches Zentralmuseum – Leibniz Research Institute for Archaeology), Geoff Carver (Römisch-Germanisches Zentralmuseum – Leibniz Research Institute for Archaeology), Roeland Paardekooper (EXARC, Netherlands; Middelaldercentret, Denmark) and Christoph Berthold (Competence Center Archaeometry Baden-Württemberg, University of Tübingen)

Contact Thomas Stöllner / thomas.stoellner@bergbaumuseum.de

Summary

A lack of standards and tools for recording and reporting on archaeological experiments and analytics hinders comparison, scientific knowledge exchange and the evaluation of data quality. The TRAIL fills this gap by developing a flexible, easy-to-use visual workflow tool that will allow researchers to record paradata: experiment design plus protocols of sample preparation procedures and/or analysis. The need for such a tool has long been recognised, but resources have been lacking. The aim is to produce a prototype of a tool that can be continually improved in line with the needs of our community.

Description

The workflow tool will document archaeological experiments in detail. Within the TRAIL, it will be tested using existing experimental and analytical data. This will

contribute towards developing a terminology and ontology structure for the various stages in an experimental and/or archaeological process. Such protocols are not well established or reported, due to the reproducibility crisis in many disciplines. Some tools for recording and/or reporting protocols (e.g. protocols.io, Open Lab Notebooks and NatureProtocols) are field-specific (e.g. chemistry/biology lab) and/or limited in functionalities or supported data formats. Therefore, a key feature of this TRAIL will be the community coordination to meet the following challenges:

- academic disdain towards experimentation by non-academics results in reluctance to share data;
- data derived from instrumental analytics is often available only in proprietary formats;
- experimental results are disseminated in a series of publications with very different levels of academic quality.

An experiment is part of a formal scientific process, testing a hypothesis. Every archaeological interpretation of an object's manufacture or use is a hypothesis to be tested. But archaeologists often lack the infrastructure, theoretical basis, resources and/or tools to do this testing or disseminate their protocols and results. Consequently, experimental archaeology urgently needs to develop ways of reporting, documenting and sharing the designs and results of experiments.

Experimental archaeology traditionally includes replication, reproduction and reconstruction of archaeological finds and features. Analytics and archaeometry are generally derived from scientific disciplines (e.g. chemical or DNA analysis) where testing and measurement methods have become routine, and only refined in response to advances in hardware and software.

Currently, there are few databases (or lists) of archaeological experiments available, and few reports are detailed enough for the procedures to be repeatable and/or results to be reproducible.

This tool should help experimenters plan better and avoid misrunning experiments, reducing the waste of time and resources (key for large-scale experiments, which often cannot be repeated). The quality of open access data needs to assessed by asking two questions.

(1) Why were the data generated? Many experiments have been run and many methods applied without clear research goals or in a way that is not appropriate to the hypothesis.

(2) How were the data generated (samples, method, equipment, and associated accuracy/precision, etc.)? Only some meta/paradata can be extracted from the data itself, and for many methods it is not yet clear how much of this is needed to assess data quality.

(3) Why they were generated in that way: methods are often employed without questioning their appropriateness for the given samples and/or hypotheses. This illustrates a reflexive process.

This TRAIL will develop a flexible tool for recording experiment design and the protocol of a sample preparation procedure or analysis. A similar tool being developed in NFDI4Chem for laboratory environments; archaeologists also need to use this sort of tool in open-air museums and in the field. EXARC's RETOLD project (in development) would complement the workflow tool, while the European Research Council funded project CAPTURE should provide guidance regarding the selection of paradata to record. The underlying ontology will build upon the Scientific Observation Model (CRMsci) and EXARC's general cyclical script for experiments in archaeology (Lammers-Keijsers 2005, EuroREA 2/2005).

The workflow tool will enable the construction of flowcharts similar to Harris Matrices. The user will construct a tree-like workflow by dragging predefined building blocks such as "hypothesis", "sample selection/preparation", "experimental/analytical steps" and "expected outcomes/data generated" from a menu pane. Other blocks will connect the steps and be used to edit a workflow: repeat (loop) steps; add (e.g. a new route), split and merge "branches"; clear "dead ends"; note unexplored routes and experience-based recommendations; and include statements regarding expected results. It will be possible to comment on every step and every connection, sharing best practice, explaining why each step was performed in a certain way, and allowing researchers to reflect on their workflow to improve the quality of the output. The results will be exportable in vector and raster formats. Versioning and version control are critical since every experiment can deviate from the original plan: changes to the workflow must be tracked (e.g. Git) and different released versions must be saved and assigned a DOI/URI (e.g. Zenodo) for citation purposes.

Data from two archaeological experiments will demonstrate the utility of this tool:

- data from mechanical, lab-based experiments at the Laboratory for Traceology and Controlled Experiments (TraCEr, RGZM) on past Hominin stone tool use to understand use-wear formation;

- data from experiments at the Laboratory for Experimental Archaeology (LEA, RGZM), in collaboration with the Competence Center Archaeometry Baden-Württemberg (CCA-BW, University of Tübingen), for quantifying the productivity of pottery shaft kilns from Mayen in a field setting, including the resulting series on ceramic raw material and objects.

The experiments have been performed and documented, and the data on the experimental samples (analytics) has been collected. However, this data is not systematically structured or connected. In this TRAIL, a prototype of the proposed workflow tool will structure this data and enable better reporting, sharing and reuse of the (meta)data.

The workflow tool will be expanded after the TRAIL phase to incorporate more types of experiments, analytics, raw materials, etc., and to store all types of data produced in experiments and analytics (images, videos, tables, graphs, 3D models, etc.) in a structured way. To this end, the prototype will be developed, refined and tested. Ultimately, every piece of data should be connected to the sample preparation/experimental/analytical workflow, and to the associated metadata. Ideally, the resulting tool will be flexible enough to augment conservation/restoration and excavation documentation.

Relevance

This TRAIL enables documentation and archiving of meta- and paradata pertaining to archaeological experiments and analytics. The workflow tool will be used to plan, prepare and conduct experiments and analyses, and provides the basis for properly publishing and long-term archiving the data. Data providers and data users researching in different fields (archaeologists, historians and archaeo-scientists), and experimental archaeologists (professionals and amateurs) will benefit from the tool. The workflow tool can be used for comparative studies, within the archaeological community and interdisciplinary. Its can be adopted to record excavation processes, decisions made when archiving material, conservation/restoration, etc., and interface with experimental studies run in other NFDI platforms, e.g. NFDI4Biodiversity, NFDI4Plants, NFDI4Cat and NFDI4Chem (especially TA2: Smart Lab). Experimental archaeology has great potential for knowledge transfer. Open-air museums and other people pursuing archaeological interests in civil society have limited access to scientific resources. The workflow tool will help them to develop, conduct and interpret their own experiments, promoting networking within the wider scientific community. The TRAIL fully complies with the FAIR principles. All participants are committed to FAIRification of the research data they contribute. The tool developed within this TRAIL should highlight the role of experimentation within the archaeological process and strengthen the epistemological basis of archaeological interpretation as a whole.

Deliverables

The TRAIL will devise, evaluate and describe the necessary disciplinary and functional requirements of the planned workflow tool. A prototype will be developed and implemented within the general work programme of TA3 (T3.1.3). Designed to be flexible and usable for all aspects of archaeological protocols, the workflow tool will contribute to a shared knowledge commons based on the FAIR principles. It can easily be integrated into existing courses and trainee programmes, contributing to a broad competence framework.

Work plan

Proposed duration: 24 months with 0.5 FTE position.

Months:	**1 - 6**	**6-12**	**12-18**	**18-24**
WP1 – Collecting the needs of the experimental and analytical communities				
WP2 – Developing and refining the ontology				
WP3 – Designing the tool (requirements, structure…)				
WP4 – Implementing/programming the prototype				
WP5 – Testing the prototype on case studies				
WP6 – Disseminating results (conferences, publications)				