

A DSpace Search Portal for Laboratory Services in Archaeometry and Heritage Science

Task Area	TA3: Analytics and Experiments
Partner Institutions	Lead: Deutsches Bergbau-Museum Bochum Co-Applicants: Participants: Leibniz-Zentrum für Archäologie, Rathgen-Forschungslabor der Staatlichen Museen zu Berlin External members: Hochschule für Bildende Künste Dresden, Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen
Contact	Yiu-Kang Hsu (yu-kang.hsu@bergbaumuseum.de)
TRAILS	TRAIL 3.3: Workflow-Tool für archäologische Experimente und Analytik

1. Summary

- Please specify the specific challenges for which solutions are to be developed within the framework of the TRAIL and explain the approaches you are pursuing to achieve this.
- Please briefly summarize the objective of the TRAIL.

Researchers in archaeometry and heritage science communities face significant barriers when seeking appropriate analytical facilities for the analysis of their samples. Information about laboratory services, instrument capabilities, experimental protocols, and data quality is often fragmented and untransparent across institutional websites, personal networks, and informal communication channels. This lack of a structured platform hinders effective planning of analytical activities, discourages cross-institutional collaboration, and leads to sub-optimal selection of methods. On the other hand, laboratories lack consistent and plain-language documentation to communicate their analytical offerings, which reduces visibility of laboratory expertise, limits comparability of analytical data from different sources, and poses a serious challenge in ensuring data quality and reproducibility.

Our TRAIL addresses these issues by designing an online search portal to manage, store, and openly share essential information about laboratory services and their associated digital assets. Leveraging community-driven metadata profiles and open-source DSpace software, our portal serves as a central hub where researchers can locate and compare analytical facilities suited to their research questions. The portal enhances the findability and reliability of analytical activities, supports the consistent documentation of laboratory workflows, and promotes reproducible, high-quality scientific practice.

2. Relevance

- Please briefly outline which section(s) of the research data cycle are primarily addressed within the TRAIL.
- Which elements of FAIR are particularly relevant? To what extent are the FAIR principles feasible?
- To what extent does TRAIL make a useful contribution to the NFDI as a whole? Please explain the extent to which TRAIL contributes to the goals, measures, tasks, and KPIs from the DFG application of NFDI4Objects.
- Please explain how the process and content of the preparatory work for Task Area 6 and Task Area 7 are designed.
- Does TRAIL address a real, possibly proven need in the specialist community and show realistic prospects for impact? If so, please explain.

The TRAIL aligns with several key stages of the research data lifecycle. It begins with the **Collect and Analyse** phase, where services and instruments are assigned unique and referable persistent identifiers (PIDs) via the built-in DSpace Handle ID system. In parallel, resource metadata is captured using rich, community-driven schemas. During the phase of **Store and Evaluate**, data entries are curated and managed within a DSpace repository maintained by a dedicated repository manager and the DBM IT support team. Finally, in the **Share and Publish** stage, the portal offers a user-friendly platform to disseminate and publish laboratory information through the user-friendly interface and search features. The item metadata can be exported in XML format or harvested through OAI-PMH API endpoints.

The laboratory search portal plays a crucial role in ensuring the **findability** and **accessibility** of analytical activities. Each lab-related digital asset is given a unique persistent identifier (PID) minted via the Handle system. The PID resolves to the corresponding item page in the DSpace instance through standardised communication protocols (HTTPs).

The TRAIL will transform unorganised laboratory information into a catalogue of analytical services with the help of open-source tool and established metadata standards. The entire setup can be readily applied to the construction of any online service across NFDI consortia. Specifically, the DSpace-based platform can turn the current NFDI4O Commons into a more structured, user-friendly, and interactive resource by leveraging the power of PIDs, customised schemas, and search functionalities.

Within the NFDI4O, the portal contributes specifically to Goals 3.1 and 3.3 of Task Area 3 in terms of establishing protocols and guidelines and publishing outcomes via a web interface. It also directly supports Measure 3.1 and related Tasks 3.1.2, 3.1.3, and 3.1.6 in which generic structures for laboratory information, analytical tools, and experimental protocols are planned to be developed, implemented, and published through an online data service.

Throughout the TRAIL, we will document the DSpace configuration process in detail, enabling other NFDI4O institutions to test, adapt, and implement it to manage their own research outputs. This documentation will be published as an open educational resource (OER) regulated by TA6. Meanwhile, the search portal will be registered as a NFDI4O service under the coordination of TA7.

Above all, the TRIAL responds to a genuine need for better management and visibility of laboratory resources supporting analytical work in archaeology and cultural heritage, as shown by previous community survey [1]. Researchers require a comprehensive portfolio of available laboratory facilities for their sample-based studies, while laboratory practitioners long for consistent documentation of experimental workflows to ensure data reproducibility.

3. (Further) development of standards and guidelines

- Is the development and harmonization of metadata standards and ontologies planned within the framework of TRAIL? If so, what specifically is planned? Will relevant standards and reference data (e.g., GND, Wikidata) be used for this purpose?
- Are there plans to develop reusable workflows (e.g., for data integration) within the framework of TRAIL?

The TRAIL will incorporate several existing metadata profiles to manage corresponding resource collections within the DSpace portal. These include:

1. **Service Collection:** Information about laboratory services, based on the [E-RHIS schema](#).
2. **Device Collection:** Descriptions of scientific instruments used to generate analytical results, following the [PIDINST standard](#).
3. **Software Collection:** Digital tools that control the operation of devices or process analytical data, structured according to [CodeMeta](#).
4. **Method Collection:** Measurement parameters or analytical protocols that guide researchers through the analytical workflows based on TWG “Lab Standards” [2].
5. **Reference Material Collection:** Standard reference materials used for device calibration and data validation, based on [GeoRem](#).
6. **Person, Laboratory, and Publication Collections:** Complementary resources that provide contextual information about contributors, service providers, and references, according to established generic schemas such as FOAF/OCRID, ROR, and Dublin Core.

The main task is to transform these metadata schemas into XML formats compatible with DSpace. Each schema can be formulated as an item/resource template for the reusability within the portal. Different resources can be linked through the Handle ID system. Meanwhile, controlled vocabularies of the targeted metadata fields can also be integrated into DSpace-compliant XML. We plan to implement Getty Art & Architecture Thesaurus, GeoNames, Wikidata, [Chemical Methods Ontology](#), and European Research Infrastructure for Heritage Science ([E-RHIS](#)) vocabulary.

4. Development of digital tools and sustainable provision of services and data

- Are there plans to provide tools for open-source standards as part of TRAIL, and if so, which tools need to be developed or further developed? To what extent will these be included in tool and service catalogues?
- Does TRAIL continue existing activities of NFDI4Objects? If so, which ones?
- Are the results of TRAIL reusable, and is there concrete potential for continuity? If so, please explain.

The search portal is primarily built upon the [DSpace](#) software architecture—an open-source platform widely adopted by universities and research institutions for managing, preserving, and disseminating digital research assets. DSpace supports the implementation of the Handle PID system, the customisation of metadata profiles and controlled vocabularies, and interoperability through OAI-PMH endpoints. DSpace's strong community and active development make it a highly attractive and advantageous solution for research data management.

Since early 2025, the DBM has established a DSpace prototype in a development environment tailored specifically for digital resources related to laboratory information. The test instance contains resource collections outlined above, with customised item submission and search functionalities (Figure 1 and Figure 2). The prototype demonstrates a strong potential to evolve into a fully operational platform. Within this

TRAIL, we aim to further develop this prototype into a production-ready service with the following features:

- Minting of persistent identifiers for services, instruments, and associated digital resources.
- Reusable metadata schemas based on community-driven standards.
- A user-friendly search interface with customisable facets.
- Rich metadata description of item page and linked external resources.
- OAI-PMH support for interoperable metadata harvesting.

For the long-term sustainability, we will request the Handle service from Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) to implement PID inside the DSpace instance. This setup allows us to mint up to 50.000 PIDs per year free of charge during the funding period of PID4NFDI. The DBM is ready to host the laboratory search portal with the support of the in-house FAIR infrastructure “Georessourcen und museumsbezogener Transfer” (GermuT). This will ensure necessary resources and IT expertise to continue the maintenance and further development of the portal.

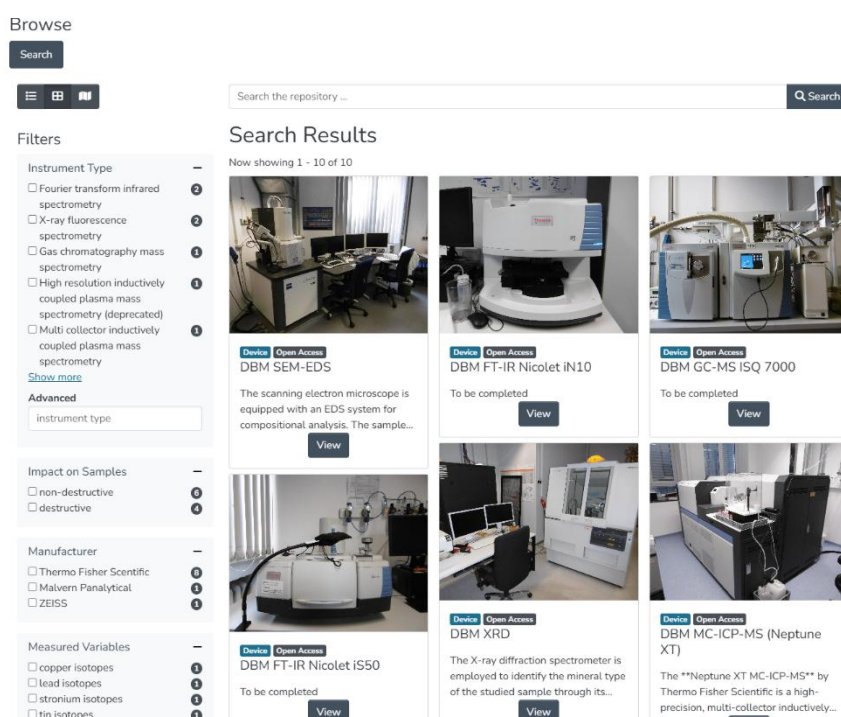


Figure 1. The search results and features based on the DBM scientific instruments. The DSpace test instance is deployed via docker containers.

 Drop files to attach them to the item, or [browse](#)

Collection

2. Device Collection ▼

Owner Information

Owner Name ▼
Owner Information

A list of institution(s) responsible for the management of the instrument. This may include the legal owner, the operator, or an institute providing access to the instrument. The default persistent identifier is ROR

[+ Add more](#)

Manufacturer Information *

Manufacturer Name ▼
Manufacturer Information

A list of the instrument's manufacturer(s) or developer. This may also include the owner for the custom built instruments

[+ Add more](#)

Model Information *

Model Name ▼
Model Information

Name of the model and identifier (normally serial number) as attributed by the manufacturer

[+ Add more](#)

Instrument Type *

Instrument Type ▼

List of classifications of the type of the instrument (also known as analytical techniques defined in other schemas) as well as its corresponding vocabulary PID

[+ Add more](#)

Measured Variables

Impact on Samples *

Figure 2. The submission form for scientific instruments. The PIDINST schema is integrated into a customised template in DSpace.

5. Visibility and empowerment of the professional community

- What measures are planned within the framework of TRAIL to support and train the specialist community (e.g., OER, training courses, onboarding formats, workshops, e-learning, communication measures for tools and standards, etc.)?

The search portal offers an excellent opportunity to introduce DSpace as a powerful open-source tool to the wider research communities of the consortium. We aim to develop an open educational resource (OER) that provides step-by-step guidance for researchers through clear, plain-language documentation. Detailed video tutorials are planned to cover all stages of repository creation, including installation, customisation, deployment, and maintenance.

6. Strategic opening, networking, and international connectivity

- To what extent is cooperation, networking, or collaboration with other NFDI4Objects task areas planned within the framework of the planned TRAIL?
- In terms of content, at which points is networking with other NFDI consortia planned?

- To what extent is cooperation, networking, or collaboration with national institutions outside the NFDI planned within the framework of the planned TRAIL?
- Does the TRAIL address new technologies, perspectives, communities, or topics that have not yet been covered? If so, please explain.
- To what extent is international cooperation, networking, or collaboration planned in the context of the planned TRAIL?

The TRAIL plans to collaborate closely with the archaeometry laboratories that previously participated in the TWG “Lab Standards”, where we have begun gathering laboratory-related information. This collaboration includes both NFDI (Leibniz-Zentrum für Archäologie, Rathgen-Forschungslabor, Curt-Engelhorn-Zentrum Archäometrie) and non-NFDI Institutions (e.g., Hochschule für Bildende Künste Dresden and Universität Tübingen). Their analytical services and instruments serve as the core database of our portal, which can be further extended by harvesting other data sources. In addition, we will establish direct cooperation with Task Area 4, whose research also relies on accurate and up-to-date laboratory information for conservation-related analyses.

Networking with PID4NFDI is also explicitly planned, particularly through the collaboration with Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG). GWDG is currently operating the [B2Inst](#) service for registering instrument identifiers. These external PIDs can be potentially linked to the digital resources stored in our DSpace search portal.

7. Deliverables

Specify the deliverables, provide names and specifications for the technical implementation of the project (e.g., shared software, developed interface, implementation and dissemination of knowledge/data, etc.), and state the expected results.

What contributions to the commons and to central offerings of NFDI4Objects (e.g., open educational resources) will be developed as part of the TRAIL? Relate this to section 5 (see above).

- WP1 DSpace Installation and Configuration

- Description and technical implementation

The WP1 sets up a DSpace instance as the underlying infrastructure to store, manage, and publish laboratory information. Yiu-Kang Hsu and the DBM IT team will install the latest version (9.1) on a dedicated DBM server. Once ready, a Handle ID minting system will be established within DSpace through the GWDG PID service. Meanwhile, metadata profiles and controlled vocabularies for each digital collection will be configured to render the item submission forms and search functionalities.

- Expected Outcomes

The work package lays the foundation for the overall data structure for laboratory-related digital resources. This setup enables us to begin the data collection with the customised submission forms and Handle ID connections.

- WP2 Use Cases and Core Database

- Description and technical implementation

Yiu-Kang Hsu (DBM) will coordinate this task by engaging laboratory partners previously involved in the TWG “Lab Standards”. They have agreed to contribute their service and device information, forming the core of our database. The authorisation and workflow system in DSpace will be configured to enable participating institutions to view, edit, manage, and withdraw their own digital resources. In parallel, we plan to expand the database by harvesting similar data platforms – [E-RHIS](#), [OpenIRIS](#), and [GeoX](#) – to increase the overall coverage of laboratory services.

- Expected Outcomes

With the contribution of real-world use cases, we may improve our metadata profiles and testify their usability to describe laboratory-related activities.

- WP3 DSpace Customisation, Deployment, and Maintenance

- Description and technical implementation

This work package aims to bring the portal to a production-ready state, focusing on its user experience and maintenance plan. Yiu-Kang Hsu will design a user-friendly frontend interface for the homepage, search display, and item pages, while the IT team at the DBM support the delivery of a backup and recovery policy.

- Expected Outcomes

Upon the completion, we create a fully functional search portal for laboratory services with the PID integration, standardised documentation, and dynamic search capabilities. Physical objects/samples from various research within the NFDI4O are able to connect with their relevant analytical activities, forming the Linked Open Data ecosystem.

- Contributions to Commons

The codebase of the entire DSpace instance will be version controlled in a Github repository. The docker images are automatically built from this instance that preserve themes, configurations, and metadata to ensure the full reproducibility of our tool. These resources along with the documentation from WP1 will be thoroughly recorded and published as an open educational

resource via NDI4O Commons as well as MkDocs. construct a complete guidance for the implementation of DSpace for any digital content.

8. Work Plan

Submit a resource-based milestone plan (in the form of a Gantt chart, if applicable) for the planned work phase (6-24 months). The plan should contain a realistic time frame with clearly defined milestones and concrete deliverables. Please take into account the reflections from the TRAIL evaluation and the SC decision on the strategic pillars.

TRAIL Duration: 01.01.2026 to 31.12.2027 (24 Months)					
Deliverable	Strategic Pillars	Tasks	Timeline	Milestone	Required Resources
WP1 Installation Configuration	P1: Standard P2: Tool	Installation	01.01.26 – 01.03.26 (2M)	Custom Resource Submission Forms Completed by 01.09.26	DBM IT
		Metadata Configuration	01.03.26 – 01.06.26 (3M)		Yiu-Kang Hsu
		Handle PID Setup	01.06.26 – 01.09.26 (3M)		Yiu-Kang Hsu GWDG
WP2 Use Cases	P1: Standard P4: Network	Data Collection & Harvesting	01.09.26 – 01.01.27 (4M)	Database ready by 01.01.27	Yiu-Kang Hsu TRAIL partners
WP3 Customisation Deployment Maintenance	P2: Tool P3: Community	Front-end Customisation	01.01.27 – 01.09.27 (9M)	Launch of Portal by 01.12.27	Yiu-Kang Hsu
		Deployment & Maintenance	01.09.27 – 01.12.27 (3M)		DBM IT
		Documentation & Tutorials	01.06.27 – 01.12.27 (6M)		Yiu-Kang Hsu

9. Funding / Own contributions

The resources required for the TRAIL build upon the outcomes of the previous activities performed by Yin-Kang Hsu. He will dedicate 50% of his NFDI4O-funded position to completing the planned tasks with the support from partner institutions, including Katrin Westner (DBM), Stefan Röhrs (Rathgen), Ivan Calandra and Roland Schwab (LEIZA), and Christoph Herm (HfBK).

10. Additions/Additional information

[1] Hsu, Y.-K. (2025). Archaeometry Community Survey. Zenodo.

<https://doi.org/10.5281/zenodo.17218351>.

[2] Hsu, Y.-K., Jan, S., Löffler, I., & Stöllner, T. (2024). NFDI4Objects Temporary Working Group: Metadata Profiles for Analytical Methods in Archaeometry. Zenodo.

<https://doi.org/10.5281/zenodo.10990411>.