



graphia

GRAPHIA

Knowledge Graphs, AI Services and Next Generation Instrumentation for R&D in Social Sciences and Humanities

Work Package WP3

Innovative Solutions and Instruments for the SSH KG

Deliverable 3.1

Catalogue of Next-Gen SSH Instruments and tools

Funding Instrument: Horizon Europe

Call: HORIZON-INFRA-2024-TECH-01

Call Topic: R&D for the next generation of scientific instrumentation, tools, methods, solutions for RI upgrade

Project Start: 2025-01-01

Project Duration: 36 months

Document Identifier: 10.5281/zenodo.19047196



Funded by
the European Union

Funded by the European Union. Grant Agreement number 101188018. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.



Deliverable Information

| | |
|--------------------------|--|
| WP Number: | 3 |
| WP Title: | Innovative Solutions and Instruments for the SSH KG |
| Deliverable Number: | 3.1 |
| Deliverable Full Title: | Catalogue of Next-Gen SSH Instruments and tools |
| Deliverable Short Title: | SSH_I&T_Catalogue |
| Document Identifier: | GRAPHIA-WP3-T3.1-Next-Gen I&T Catalogue |
| Beneficiary in Charge: | CNR |
| Version: | v1.0 |
| Submission Date: | 2026-04-09 |
| Dissemination Level: | PU |
| Nature: | DEM |
| Lead Author(s): | Giovanni Pescarmona, CNR ISPC / E-RIHS; |
| Co-author(s): | Angelo Di Iorio, UNIBO; Tomasz Hoffman, PCSS; Julien Homo, FOXCUB; Matteo Greco, CNR ISPC; Arek Margraf, PSNC; Noushin Najarifagheb, CNR ISPC; Vladislav Nechakhin, TIB; Anthony Pamart, CNRS MAP; Lorenzo Paolini, UNIBO; Silvio Peroni, UNIBO; Marta Soricetti, UNIBO; Frank Uiterwaal, NIOD-KNAW; Xiao Xu, NIOD-KNAW; Jan Wiebelitz, LUH, Valentina Vassallo, CYI; Quentin Vogel, CNRS MAP; Athanasios Koutoupas, CYI |
| Reviewer(s): | Michela Magas, ICF; Marcin Plociennik, PCSS |
| Keywords: | SSH, Tools, Services, R&D, Instruments |
| Status: | Submitted |



Change Log

| Date | Version | Author/Editor | Summary of Changes made |
|------------|---------|----------------------------------|---|
| 2026/01/23 | v0.1 | Giovanni Pescarmona | First release of draft document |
| 2026/03/06 | v0.2 | Giovanni Pescarmona | Major edits and partners contributions |
| 2026/03/24 | v0.3 | Michela Magas, Marcin Plociennik | Revised version from external reviewers |
| 2026/03/27 | v0.3 | Mandy Y Lin | Quality control |
| 2026/04/09 | v1.0 | Giovanni Pescarmona | Included JSON dataset link into the appendix Final release |



Table of Contents

| | |
|--|-----------|
| Deliverable Information | 2 |
| Change Log | 3 |
| List of Figures | 6 |
| Executive Summary | 7 |
| 1. Introduction | 9 |
| 1.1 Purpose and Scope of the Document | 9 |
| 1.2 Task 3.1 - R&D of next-gen SSH Instrumentation and tools | 10 |
| 1.3 Next-gen SSH catalogue | 10 |
| 1.4 Structure of the Document | 11 |
| 2. Methodological Framework | 13 |
| 2.1 From Milestone M2 to Deliverable 3.1 | 13 |
| 2.2 From Milestone M2 to Deliverable 3.1 | 13 |
| 2.3 Metadata Model and Alignment with SSH Open Marketplace | 14 |
| 2.4 Working Dataset, SSHOMP Onboarding, and Final Output | 15 |
| 2.5 Contribution to the SSH Community | 15 |
| 3. Instruments & Tools Implementation of the Graphia Catalogue | 16 |
| 3.1 Implementation Strategy | 16 |
| 3.2 Onboarding of GRAPHIA Instruments and Tools | 17 |
| 3.3 Metadata Alignment and Resource Description | 18 |
| 3.4 Relations, Contextualisation and Discoverability | 19 |
| 3.5 Media, Documentation and Resource Visibility | 20 |
| 3.6 Dataset Export and Future Automation | 20 |
| 4. Catalogue of next-gen SSH instruments and tools | 21 |
| AIOLI - Collaborative platform for 2D/3D spatialized semantic annotations for Heritage Digital Twin applications | 21 |
| CEC | 24 |
| CEX - PDF Parsing and Citation Extraction | 24 |
| CIC - Citation Intent Classification | 25 |
| DACE | 30 |
| EHRI LOD Knowledge Graph and the EHRI Virtual Observatory (EHRI Portal (and APIs) - EHRI Document Blog - EHRI Geospatial Repository) | 32 |
| HYPERMNESIA: A Framework for Reinforced Research Traceability in | |
| Heritage Science Open Data | 37 |
| Massive 3D digitization platform for CH objects and spaces as transversal approach to FIXLAB, MOLAB and DIGILAB | 41 |



| | |
|--|-----------|
| MOVIDA | 44 |
| MuSeSca3D – Multi Sensor Scanner 3D (Integrated MA-XRF / RIS / LIS System) | 46 |
| ORKG | 46 |
| SENNSE | 48 |
| VTL | 50 |
| 5. Conclusion | 52 |
| 9. Additions - Entries in SSHOC | 54 |
| Appendix A. | 55 |



List of Figures

Figure 1 - Example of wood remain NF862 from Notre Dame de Paris 3D annotated dataset within AIOLI viewer linked to open data ([doi:10.57745/4QUWZE](https://doi.org/10.57745/4QUWZE))

Figure 1.1 - Example of wood remain NF862 from Notre Dame de Paris 3D in the Quasi.Modio platform semantically enriched digital model (left) and W7 semantic signature extracted from the KG (right)

Figure 2 - CEX pipeline

Figure 3 - Working Dynamics of CiteFusion

Figure 4 - CEX Web Application Interface

Figure 5 - CIC Web Application Interface

Figure 6 - Example of processing workflow pipeline combined from DACE microservices

Figure 7 - Screenshot of the EHRI Portal (<https://portal.ehri-project.eu/>)

Figure 8 - Screenshot of the EHRI Document Blog (<https://blog.ehri-project.eu/>)

Figure 9 - Screenshot of the EHRI Geospatial Repository (<https://geodata.ehri-project.eu/>)

Figure 10 - Screenshot of the EHRI Knowledge Graph prototype's LODLive interface, featuring web-based navigation of key entities and their relationships: (<https://lod.ehri-project-test.eu/>)

Figure 11 - Anamnesis frontend dashboard

Figure 11.1 - W7 ANAMNESIS concepts (247 entries) referenced in [OpenTheso](#)

Figure 11.2 - Graphical representation of the W7.WHAT section and hierarchical concepts in [OpenTheso](#)

Figure 12 - Anamnesis W7 metamodel for activity description

Figure 13 - Example of metadata scheme for the documentation of photogrammetric capture activity ([from Wiki-based ANAMNESIS user documentation](#))

Figure 13.1 - Json to Graph visualization of W7 metadata of the data acquisition on the wood remain NF862

Figure 14 - Massive 3D digitization instrumental platform ; SQUILLIDEA (top-left), DIGIT-SPIDER (bottom-left), DIGIT-ARCH (middle) and DIGIT-STUDIO (right)

Figure 15 - MOVIDA 2.0 interface

Figure 16 - Screenshot of the ORKG home page (<https://orkg.org/>)

Figure 17 - SENNSE application interface and hardware IoT devices

Figure 17.1 - SENNSE metadata template scheme structured in W7 using ANAMNESIS

Figure 18 - VTL interface



Executive Summary

This deliverable presents **D3.1 – Catalogue of Next-Generation SSH Instruments and Tools**, developed within **Task 3.1 of the GRAPHIA project**. The deliverable documents a set of advanced hardware and software tools contributed by project partners and relevant to the Social Sciences and Humanities (SSH) data lifecycle, from acquisition and processing to enrichment, interpretation, and publication. The instruments and tools described in this catalogue support the generation and management of multimodal data and contribute to the broader GRAPHIA objective of enabling semantically enriched, interoperable, and graph-ready SSH data (i.e. structured and mapped to support ingestion into the GRAPHIA SSH Knowledge Graph).

A key result of the work carried out in Task 3.1 is the consolidation and the adoption of a **structured metadata model for cataloguing instruments and tools**, aligned with the **SSH Open Marketplace** metadata guidelines for tools and services. This choice reflects a strategic decision taken by the consortium during T3.1 coordination phase: rather than developing yet another isolated catalogue, GRAPHIA has chosen to leverage an existing European SSH digital infrastructure that already acts as a discovery portal, contextualised catalogue, and entry point to EOSC for SSH researchers. The SSH Open Marketplace (“SSHOMP”) is explicitly designed to pool and contextualise SSH resources — including tools, services, datasets, publications, workflows and training materials — while supporting contribution, reuse and interoperability through structured metadata and APIs.

In this sense, D3.1 is not limited to a descriptive report. It is conceived as a **demonstrator** of a practical and sustainable catalogue strategy for GRAPHIA. The deliverable combines:

1. a **documentary** layer, describing the methodological framework, the prototype implementation in the marketplace, and the selected instruments;
2. a structured **dataset**, progressively aligned with SSHOMP metadata fields;



This approach allows GRAPHIA to validate the catalogue data model in a realistic environment and to test its usability, visibility, and interoperability without duplicating infrastructure or creating a new siloed service.

The work carried out for D3.1 also generates added value for the wider SSH community by contributing new descriptions, actors, and conceptual enrichments to the SSHOMP ecosystem, including GRAPHIA-related entries and vocabulary extensions relevant to specific domains in the SSH broader ecosystems, such as Heritage Science. The deliverable therefore represents both a project output and a contribution to the wider European SSH service landscape.

In addition, the catalogue strategy adopted in GRAPHIA is designed to remain extensible beyond the scope of this deliverable. The SSH Open Marketplace roadmap foresees the introduction of a **Collections functionality**, enabling curated groups of related resources to be organised and presented as coherent thematic sets. This development will offer an opportunity for GRAPHIA to implement its catalogue of instruments and tools as a dedicated **GRAPHIA collection within SSHOMP**, providing a structured entry point to the project's technological outputs while maintaining their integration within the broader SSH research infrastructure ecosystem. Such an approach ensures that the catalogue can continue to grow, incorporating new instruments, enriched metadata, and additional links to workflows, datasets and use cases developed throughout the project (notably through Tasks 3.2, 3.3 and 4.4).

By leveraging the SSH Open Marketplace as the implementation environment for the catalogue, GRAPHIA also contributes to the broader **European Open Science Cloud (EOSC)** ecosystem. Since SSHOMP acts as an entry point to EOSC for SSH resources, the integration of GRAPHIA instruments and tools within the marketplace enhances their visibility, discoverability, and interoperability within the European research infrastructure landscape.

Finally, D3.1 lays the foundation for subsequent developments within WP3. The catalogue and its metadata model will continue to evolve after March 2026, supporting the refinement of workflows in **Task 3.2**, the implementation of **Task 3.3 use cases**, and the alignment of the described tools with the interoperability and ingestion requirements of **WP2**. In this perspective, D3.1 should be understood as the first operational prototype of a broader, extensible GRAPHIA catalogue of instruments and services.



1. Introduction

1.1 Purpose and Scope of the Document

The purpose of this document is to present the **Catalogue of Next-Generation SSH Instruments and Tools** developed within **Task 3.1 of the GRAPHIA project**. The deliverable provides a structured overview of selected hardware and software solutions that have been, or are in the process of being upgraded, refined, or validated within the project in order to support the acquisition, processing, semantic enrichment, and reuse of data in the Social Sciences and Humanities domain.

The scope of the catalogue includes both **physical instruments** and **digital services**, ranging from environmental sensing devices (SENNSE), data input software (MOVIDA), digitisation ecosystems (MAP-CNRS), and knowledge graph platforms to annotation environments, OCR/OMR tools, and workflow-oriented data processing systems. These solutions are relevant because they enable the production and management of complex multimodal data, often in direct relation to cultural heritage, archival, bibliographic, and scholarly communication use cases.

During the preparation of the deliverable, the consortium clarified that the main objective of D3.1 is not simply to list tools, but to **test and validate a structured workflow for cataloguing SSH instruments and tools** in a way that is reusable, interoperable, and strategically sustainable. For this reason, the work has progressively converged toward alignment with the **SSH Open Marketplace**, which is maintained by DARIAH, CLARIN and CESSDA research infrastructures and offers a well-established framework for the discovery, contextualisation and reuse of SSH resources. SSHOMP does not host the resources themselves, but provides structured metadata, relations between items, and API-based reuse, making it particularly suitable as a target environment for the GRAPHIA catalogue prototype.

This deliverable therefore combines a **catalogue of selected GRAPHIA instruments and tools** with the description of a **cataloguing methodology**, a metadata model, and an implementation strategy based on SSHOMP. It also prepares the ground for



further work, both enhancing currently listed entries and laying the ground for new tools to be added to the catalogue.

1.2 Task 3.1 - R&D of next-gen SSH Instrumentation and tools

Task 3.1 focuses on the **research and development of next-generation SSH instrumentation and tools**, with the objective of improving how SSH data are acquired, processed, made interoperable and interpreted. The task addresses both hardware and software innovation, covering the full data lifecycle and supporting the production of richer, well structured and more reusable digital resources.

Two main innovation lines are addressed within the task. The first concerns the **upgrading of existing hardware and software solutions**, so that they can produce more robust, semantically aligned and graph-compatible data. The second concerns the **development of new digital methods, services and workflows** that support the transformation of heterogeneous and multimodal data into resources that can later be integrated within the GRAPHIA SSH Knowledge Graph.

A central feature of Task 3.1 is its close relation with the other components of WP3 and the wider project architecture. Instruments and tools described in this deliverable are expected to connect with the workflow work developed in **Task 3.2**, with the **use cases of Task 3.3**, and with the semantic and ingestion requirements developed in **WP2**. While the tools described in this deliverable have been developed independently by each partner, several of them already produce structured or semantically enriched outputs (e.g. RDF, W7 metadata, ontology-mapped data) that are relevant to the GRAPHIA data architecture. Assessing their alignment with the WP2 ingestion requirements and the SSH Knowledge Graph data model is planned as a priority for the next project phase. For this reason, D3.1 is not only an inventory of tools, but a first operational step in building a reusable service layer for GRAPHIA.

1.3 Next-gen SSH catalogue

The need for a **catalogue** emerged from the recognition that GRAPHIA involves a significant number of advanced and heterogeneous tools distributed across multiple partners and research infrastructures, often dealing with heterogeneous data



spanning a wide range of different knowledge domains. Without a common framework, these solutions risk remaining fragmented, poorly visible, and difficult to compare or integrate.

During the T3.1 coordination process, the consortium explicitly addressed the potential risk of creating a parallel catalogue. This was judged strategically undesirable and potentially anti-economical, especially given the existence of mature SSH infrastructures already dedicated to resource discovery and contextualisation. In response, the consortium agreed to orient D3.1 toward a solution based on **reuse and interoperability**, rather than duplication.

The **SSH Open Marketplace** was identified as the most suitable reference environment for this purpose. SSHOMP is explicitly described as a **discovery portal**, an **aggregator of curated resources**, a **catalogue contextualising resources**, and an **entry point into EOSC for SSH researchers**. At the same time, it is explicitly *not* a repository and *not* a generic data catalogue, which makes it particularly appropriate for GRAPHIA's needs: GRAPHIA does not need to host tools in a new platform, but to describe, contextualise, and connect them in a shared SSH ecosystem.

The catalogue strategy adopted in D3.1 therefore aims to make GRAPHIA instruments and tools **findable, contextualised, and interoperable**, while also creating the conditions for future reuse within the SSH and research infrastructure communities. This approach is consistent with FAIR principles and with the broader aim of integrating GRAPHIA outputs into existing European ecosystems rather than creating isolated project-specific silos, which would hinder discoverability and practical use, also by industry partners.

1.4 Structure of the Document

The document is structured as follows:

- **Section 2** presents the methodological framework of the catalogue, including the rationale for the chosen strategy, the definition of the data model, and the alignment with SSHOMP.



- **Section 3** describes the prototype implementation logic, including the progressive onboarding of GRAPHIA resources into SSHOMP and the relation between the working dataset and the final exportable output.
- **Section 4** presents the catalogue entries of selected next-generation SSH instruments and tools, including their current state of development and expected upgrades.
- **Section 5** provides concluding remarks and outlines the next steps for the evolution of the catalogue after the delivery of D3.1.
- **Appendix A** contains the analytical dataset underpinning the catalogue and related working material.



2. Methodological Framework

2.1 From Milestone M2 to Deliverable 3.1

Milestone M2 provided the first draft inventory of instruments and tools, contributed by Task 3.1 partners. The milestone was primarily based on a shared spreadsheet and a descriptive report, and its main objective was to identify candidate tools and services to be upgraded or further developed within GRAPHIA.

D3.1 builds on that first phase, but goes beyond it in two important ways. First, it transforms the initial inventory into a more formal **catalogue model**, based on explicit metadata fields and a common structure. Second, it introduces a **demonstrator logic**, in which the catalogue is no longer only documented internally, but is progressively mapped and implemented in a real SSH discovery environment.

This shift reflects the overall T3.1 working logic, where the consortium agreed that D3.1 should test not only the descriptive content of the catalogue, but also its practical viability as a structured and usable resource.

2.2 Demonstrator Nature and Scope of D3.1

According to the Grant Agreement, D3.1 is classified as a DEM – Demonstrator, pilot, prototype, so as a deliverable it must guarantee that it:

- exists as a usable and testable prototype for methods and practices of the community,
- validates a data model, even if not yet fully production-ready,
- demonstrates feasibility and usability,
- documents how the prototype can evolve into a more complete service.

In the case of D3.1, the demonstrator dimension is represented by the combination of:

- a structured metadata model for SSH instruments and tools,
- a tangible dataset, progressively populated by partners,
- a first consolidated set of entries, created or updated in the SSH Open Marketplace,



- the possibility to export the resulting catalogue data via API in machine-readable form.

D3.1 is therefore conceived as a **prototype catalogue implemented through an existing SSH infrastructure**, rather than as a self-standing platform developed from scratch. The demonstrator dimension of Deliverable D3.1 is constituted by the following components:

- A structured metadata model specifically designed for Social Sciences and Humanities (SSH) instruments and tools.
- A tangible dataset, incrementally populated by the contributing partners.
- An initial consolidated collection of entries, established or updated within the SSH Open Marketplace.
- The facility to export the resulting catalogue data in a machine-readable format via an Application Programming Interface (API).

Furthermore, each entry within the catalogue represents an instrument or tool that inherently possesses a **prototype, pilot, or experimental quality**, reflecting its ongoing research and development status within the GRAPHIA project.

2.3 Metadata Model and Alignment with SSH Open Marketplace

The metadata model adopted for D3.1 is based on the **SSHOMP metadata guidelines for Tools and Services**, complemented where necessary by GRAPHIA-specific information needs, such as the documentation of expected upgrades, relevance to the GRAPHIA project objectives, and relations to the SSH Knowledge Graph architecture developed in WP2. This choice was made because SSHOMP already provides a mature, community-oriented and interoperable metadata environment for SSH resources.

The catalogue model includes, among others, the following categories of information:

- identity and description of the tool/service,
- responsible organisation(s),
- type of resource and keywords,
- target audience and mode of use,



- URLs to documentation, helpdesk and access conditions,
- contributor, creator and provider roles,
- state of development and expected upgrade,
- media and thumbnails (where available),
- relations with other resources in the SSH Open Cloud.

This alignment allows GRAPHIA to reuse an existing metadata ecosystem and increases the possibility of future integration, discoverability and reuse across SSH communities.

A full disclosure of the implemented metadata fields is available in the .json dataset that is part of this Deliverable (**Appendix A**).

2.4 Working Dataset, SSHOMP Onboarding, and Final Output

The T3.1 working spreadsheet initially served as the shared environment for collecting and harmonising partner input. A consolidated version of this dataset was originally published as Appendix A of the Milestone 2 document (Draft catalogue of instruments and tools). As the strategy matured, the consortium agreed that this spreadsheet should be understood mainly as an intermediate working dataset, supporting the **transition** toward the actual onboarding of entries into SSHOMP and toward the formalisation of a **machine-readable dataset**, queried through APIs.

Partners were therefore asked to progressively create or update their own tools and services directly in SSHOMP, using the shared spreadsheet as a coordination instrument to maintain **consistency** in level of detail and metadata quality. The longer-term objective is that, once the relevant GRAPHIA entries are in place, the catalogue can be exported as a single JSON dataset via SSHOMP APIs, which can then serve as the machine-readable output associated with the deliverable.

This approach gives D3.1 both a **human-readable documentary dimension** and a machine-readable implementation dimension, in line with the expected nature of a DEM deliverable.

2.5 Contribution to the SSH Community

The work carried out for D3.1 does not only produce a project-specific result. It also contributes to the broader SSH ecosystem by enriching SSHOMP with new entries,



new descriptions, and new conceptual elements. Examples include the introduction of GRAPHIA-related actors such as **GRAPHIA Project** and **E-RIHS ERIC**, as well as new conceptual vocabulary relevant to the SSH and heritage science domain, such as **Heritage Science**.

This aspect is important because it demonstrates that the catalogue strategy adopted by GRAPHIA is not extractive or isolated: it strengthens an existing SSH infrastructure and contributes to the visibility and structuring of research resources beyond the project itself.

3. Instruments & Tools Implementation of the Graphia Catalogue

3.1 Implementation Strategy

The implementation of the GRAPHIA Catalogue of Instruments and Tools follows a strategy based on **integration with an existing SSH discovery infrastructure**, rather than the creation of a new standalone catalogue platform. This approach emerged from the coordination discussions within Task 3.1, where partners recognised that creating a new dedicated catalogue would risk generating additional fragmentation in the SSH ecosystem and duplicating functionality already provided by established infrastructures.

To avoid this outcome, the consortium agreed to implement the catalogue prototype using the **SSH Open Marketplace** (SSHOMP) as the primary discovery and contextualisation environment for GRAPHIA resources. SSHOMP is maintained jointly by the research infrastructures DARIAH, CLARIN and CESSDA and provides a curated catalogue of tools, services, datasets, training materials and workflows relevant to the Social Sciences and Humanities community.

The SSH Open Marketplace functions as a **discovery portal and contextualised catalogue** of SSH resources, offering structured metadata, relations between resources, and API-based access to the catalogue content. Importantly, the platform



does not host the tools themselves but provides descriptions and contextual links that enable users to discover and access external services and infrastructures.

This characteristic makes SSHOMP particularly suitable for the GRAPHIA catalogue: the project does not aim to host tools centrally, but rather to **describe, contextualise, and connect** existing research instruments and services within a **shared European ecosystem**. By adopting SSHOMP as the implementation environment for the catalogue prototype, GRAPHIA ensures that its outputs remain interoperable with existing SSH infrastructures and avoid the creation of a project-specific silo.

Finally, Foxcub, as technical coordinator of the GRAPHIA project, contributed to the T3.1 coordination process by ensuring alignment between the catalogue strategy and the overall **project architecture**, including coherence with the WP2 data acquisition and integration framework.

3.2 Onboarding of GRAPHIA Instruments and Tools

The practical implementation of the catalogue prototype involves the **progressive onboarding** of selected GRAPHIA instruments and tools into the SSH Open Marketplace. Each partner responsible for a tool or service is invited to create or update the corresponding entry in SSHOMP, using the metadata fields defined in the marketplace guidelines for tools and services.

The onboarding process follows a number of common **conventions** agreed during the Task 3.1 coordination meetings. In particular, the roles associated with each entry are assigned as follows:

- **Creator** – the organisation responsible for developing and maintaining the tool or service;
- **Provider** – the infrastructure or organisation responsible for providing access to the service;
- **Contributor** – the project or initiative that contributed to the development or upgrade of the resource.

Within the context of GRAPHIA, this typically results in the following structure:

- the **Creator** corresponds to the partner responsible for developing the tool;



- the **Provider** may correspond to a research infrastructure such as **E-RIHS ERIC** or another hosting organisation;
- the **Contributor** role is mainly (but not exclusively) used to reference the **GRAPHIA project**, reflecting the contribution of the project to the development or upgrade of the resource.

This structure allows GRAPHIA to be explicitly acknowledged within the metadata while ensuring that the ownership and maintenance responsibilities of the tools remain clearly attributed to their respective institutions.

3.3 Metadata Alignment and Resource Description

The entries created within the SSH Open Marketplace follow the metadata structure defined in the **metadata guidelines for Tools and Services**. These guidelines specify a set of mandatory, recommended, and optional metadata fields that ensure consistent description and discoverability of resources.

Among the most relevant metadata elements used in the GRAPHIA catalogue are:

- resource name and description;
- responsible organisations and contributors;
- keywords and disciplinary context;
- access URLs and documentation links;
- intended audience and usage mode;
- support contacts and helpdesk references;
- related resources and contextual links.

The use of this metadata model ensures that the GRAPHIA catalogue entries are both **human-readable and machine-actionable**, enabling their discovery through the SSHOMP interface while also supporting programmatic access through the platform APIs.

In addition, the project has adopted a shared convention to facilitate the identification of GRAPHIA resources within the marketplace: catalogue entries include the keyword **“GRAPHIA Project”**, allowing them to be retrieved collectively through keyword-based search queries.



Crucially, the consortium has prioritized the inclusion of the **recommended metadata fields** from the SSHOMP schema, moving beyond the basic mandatory requirements. This effort ensures that each entry provides a rich and contextualized description, maximizing the visibility and utility of the GRAPHIA instruments for the wider SSH community. By consistently populating these recommended fields, the quality of the catalogue data is enhanced, which is vital for enabling effective searching, filtering, and cross-linking within the SSH Open Marketplace and, by extension, the EOSC ecosystem.

The consistent use of the "GRAPHIA" keyword within all catalogue entries is a deliberate strategic choice that not only enhances immediate discoverability but also anticipates future developments. This consistent tagging will be essential for the subsequent implementation of the catalogue as a dedicated **GRAPHIA Collection** within the SSH Open Marketplace, a feature planned for release in the near future. Once available, this functionality will allow users to seamlessly filter and group all GRAPHIA-contributed instruments and tools, maintaining their coherence as a project output while benefiting from full integration into the broader SSHOMP ecosystem.

3.4 Relations, Contextualisation and Discoverability

One of the advantages of implementing the catalogue through SSHOMP is the possibility to contextualise tools and services within a wider ecosystem of SSH resources. The platform allows the creation of **relations between items**, enabling tools to be connected with datasets, workflows, publications, or other services, notwithstanding their inclusion or development in the GRAPHIA project.

In the case of GRAPHIA, this functionality is particularly relevant because the instruments and tools described in the catalogue may also relate to:

- workflows developed within **Task 3.2**,
- datasets produced through project activities, use cases explored in **Task 3.3**,
- or services already available through participating research infrastructures.

Although the semantic richness of these relations is currently limited — since the marketplace provides a general “related items” mechanism rather than strongly



typed semantic relations — the feature already allows the creation of meaningful contextual links between resources. It should be noted that this limitation means that SSHOMP relations cannot be directly transposed into the GRAPHIA SSH Knowledge Graph without additional semantic mapping. Addressing this gap is expected to be part of the follow-up work in Tasks 3.2 and WP2.

3.5 Media, Documentation and Resource Visibility

To improve the usability and visibility of catalogue entries, partners were encouraged to enrich their SSHOMP descriptions with **media elements**, including screenshots, thumbnails, and interface images. When multiple images are provided, the marketplace automatically generates a carousel that allows users to browse visual representations of the resource.

This feature is particularly useful for tools and services with graphical interfaces, as it allows users to quickly understand the nature and functionality of the resource before accessing external documentation or platforms. Where available, entries also include links to documentation, user manuals, APIs, and helpdesk contacts.

3.6 Dataset Export and Future Automation

An important feature of the SSH Open Marketplace is the availability of **API-based access to the catalogue metadata**. Once the relevant GRAPHIA entries are published and approved, it becomes possible to retrieve their descriptions programmatically through the marketplace API.

This functionality opens the possibility of exporting the catalogue entries as a **machine-readable dataset**, for example in JSON format. Such an export could serve as the final structured dataset associated with Deliverable D3.1, complementing the descriptive documentation provided in this report.

In this perspective, the working spreadsheet used during the early phases of Task 3.1 should be understood primarily as a **coordination and drafting instrument**. As the onboarding process progresses, the authoritative catalogue data will increasingly reside within the SSH Open Marketplace itself.

Future work may explore additional automation mechanisms, including API-based ingestion and synchronisation workflows, enabling more efficient management of the



GRAPHIA catalogue as the number of instruments and services evolves during the project.

4. Catalogue of next-gen SSH instruments and tools

The catalogue presented in this section provides a structured overview of the next-generation SSH instruments and tools developed or upgraded within the GRAPHIA project. The rationale adopted for this catalogue is to ensure that each entry offers a consistent and comprehensive set of information about the respective instrument or tool, covering its purpose, organisation, state of development, and future upgrades. Crucially, this descriptive content is directly aligned with and, where applicable, implemented within the **SSH Open Marketplace (SSHOMP)**, ensuring that the GRAPHIA project outputs are discoverable, contextualised, and integrated into the wider European SSH research infrastructure ecosystem.

Moreover, the inclusion of the MuSeSca3D system by the Cyprus Institute (GRAPHIA partner), demonstrates the catalogue's inclusiveness and extensible nature and anticipates its future enrichment.

Note: instruments and tools are presented in alphabetical order

AIOLI - Collaborative platform for 2D/3D spatialized semantic annotations for Heritage Digital Twin applications

Organisation: CNRS MAP

Short description: [AIOLI](#) is a collaborative platform that aims to bring together the actors involved in documenting, safeguarding, and sharing our cultural heritage



through new practices for a better understanding of heritage objects¹. This service is based on two major technological developments: the democratization of photogrammetry techniques, which allow us to compute a 3D model by correlating images, and the possibility of massively processing and sharing gathered data through the cloud. Powered by remote-computing and containerization performance its innovation relies on a specific development for multidimensional propagation of spatial semantic annotations. Through this propagation process, annotations are automatically reprojected on all the 2D and 3D views of the object (past, present, and future) as described in ([Abergel et Al., 2023](#)). The platform has been used and developed during several research projects from the past decade (see contributors the tool's webpage in SSHOMP). It was the core spatialization engine in the scientific work for the restoration of Notre Dame de Paris (1530 projects, 14000 annotations created)² fostering 180 researchers and heritage scientists contributing to a digital corpus rich of 40 terabytes of KG-compliant and in [opening-data](#). This work and the methodology achieved within the [ERC NDame](#) coordinated by Livio de Luca is currently under deployment in the French infrastructure E-RIHS related project (Equipex+ ESPADON) to ensure its dissemination and availability for a larger community. Moreover, its accessibility in the [ECCCH-ECHOES](#) as vertical application is also under consideration.

¹ Violette Abergel, Adeline Manuel, Anthony Pamart, Isabelle Cao, Livio De Luca. Aioli: A reality-based 3D annotation cloud platform for the collaborative documentation of cultural heritage artefacts. *Digital Applications in Archaeology and Cultural Heritage*, 2023, 30, pp.e00285. [\(10.1016/j.daach.2023.e00285\)](#). [\(hal-04249650\)](#)

² Roxane Roussel, Livio De Luca. An approach to build a complete digital report of the Notre-Dame Cathedral after the fire, using the Aioli platform. 29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage: Humanities and Digital Technologies for Shaping the Future", Jun 2023, Florence, Italy. pp.1359-1365, [\(10.5194/isprs-archives-XLVIII-M-2-2023-1359-2023\)](#). [\(hal-04146083\)](#)



Figure 1 - Example of wood remain NF862 from Notre Dame de Paris 3D annotated dataset within AIOLI viewer linked to open data (doi:10.57745/4QUWZE)

State of the art and expected upgrade: The AIOLI platform stands as a robust solution for generating fast and efficient 2D/3D annotations, serving as a valuable data source for AI-driven approaches within WP4. Its integration into Graphia's instrument catalogue, linked to both SSHOMP and E-RIHS DIGILAB, will significantly broaden access for the cultural heritage (CH) community. While the platform's main server is nearing end-of-life, a new instance is being deployed via ESPADON facilities, with early access to the updated version expected soon for new users and pilot cases. Additionally, an API is planned for release in the coming year as a deliverable of the [EU-funded ECCCH-StratiGraph](#) project (coordinated by CNR), further enhancing its utility. For GRAPHIA's objectives, the focus will be on improving interoperability between AIOLI and the GRAPHIA SSH Knowledge Graph, in particular by exploring how AIOLI's annotation outputs can be mapped to the GRAPHIA data model and ingested via the data acquisition platform developed in WP2. Currently, AIOLI's output is distributed as file dumps, but automating deposits to open repositories—such as GoTriple, where less than 2% of datasets are currently available—could drastically increase data accessibility. Another key development involves enabling interactive access to annotated photogrammetric scenes (distinct from raw photogrammetric datasets like point clouds or 2D annotations) through IIIF-based services, linked to OAI-PMH and CoreTrustSeal-certified repositories. This effort aligns with broader initiatives to strengthen interoperability among French instruments, orchestrated by the SSH infrastructure IR*HumNum (e.g., [ArcheoGrid](#),



the [National 3D Data Repository](#), and [Nakala](#)). The goal is to streamline access to metadata and data derived from reality-based modeling, supported by a harmonized PID framework—using DOIs for annotated datasets and URIs for semantic features—to ensure sustainability and FAIR compliance.



Figure 1.1 - Example of wood remain NF862 from Notre Dame de Paris 3D in the [Quasi.Mode platform](#) semantically enriched digital model (left) and W7 semantic signature extracted from the KG (right)

CEC

Organisation: UNIBO

Short description: CEC (Citation Extractor and Classifier) is a software that performs the automatic annotation of in-text citations in academic papers provided in PDF. It works by applying two steps, carried out by the Citation Extraction (CEX) and the Citation Intent Classification (CIC) components.

CEX - PDF Parsing and Citation Extraction

The Citation Extraction component takes scholarly PDFs as input and produces: (i) a TEI-XML representation encoding bibliographic data (e.g., reference list) and document structure (e.g., sections and paragraphs); (ii) a JSON file that records each in-text reference pointer with its sentence-level citation context and, when available, the enclosing first-level section heading - offering also the possibility to perform semantic alignment of the original PDF section titles with those that conform to the



[Discourse Element Ontology \(DEO\)](#); and (iii) a RDF graph compliant with the OpenCitations Data Model to support semantic interoperability ([Daquino et al., 2020](#)). The alignment of the OpenCitations Data Model with the GRAPHIA data architecture is being assessed as part of WP2 integration activities, with the aim of enabling direct ingestion of CEC's RDF output into the SSH Knowledge Graph. Figure 2 summarizes these steps.

To obtain a structured document representation, CEX builds on a trained version of GROBID, a framework for extracting, parsing and re-structuring raw PDF documents into TEI-encoded XML ([Lopez, 2009](#)). In the context of CEC, GROBID was selected for its modular, trainable architecture, and, mostly, for its ability to detect in-text citation callouts – thereby supporting CEX's primary objective of capturing in-text citations together with their local context.

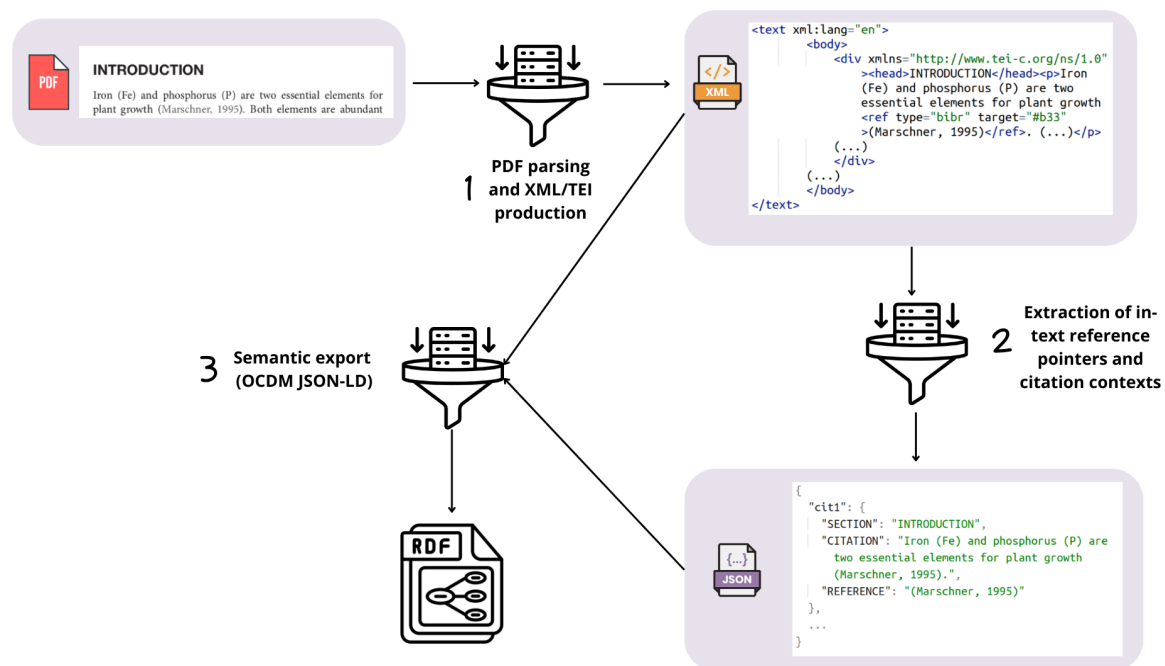


Figure 2 - CEX pipeline

CIC - Citation Intent Classification

The Citation Intent Classification (CIC) component assigns an intent to each extracted citation context. Inputs can be provided either as the CEX JSON output or as lists of tuples encoding citation contexts, optionally including first-level section titles. It supports single-item and batch processing (including compressed archives of multiple



JSON inputs) and returns intent-annotated outputs in the same format, reporting the predicted label, meta-classifier confidence, and base binary-model probabilities.

The core inference engine of the classification service is CiteFusion ([Paolini et al., 2025](#)), a state-of-the-art ensemble model for citation intent classification. Within CIC, the original multi-class classification problem is decomposed into a set of one-vs-all binary classification subtasks, each corresponding to a specific citation intent category. For each binary subtask, two complementary pretrained language models (PLMs) are employed: a domain-specific model, SciBERT ([Beltagy et al., 2019](#)), optimized for scientific discourse, and a general-purpose model, XLNet ([Yang et al., 2020](#)), which captures broader linguistic patterns. Each model is independently fine-tuned to distinguish the target intent from all the remaining classes. For a given citation context, the positive-class probabilities produced by all binary classifiers are then concatenated into a single feature vector and passed to a feed-forward neural network meta-classifier (Figure 3), which reconstructs the original multi-class prediction.

CIC adopts the intent classification schema employed by CiteFusion, grounded in object properties of CiTO and aligned with standard benchmark datasets such as SciCite ([Cohan et al., 2019](#)). In particular, CIC classifies citation contexts as providing background (cito:obtainsBackgroundFrom), reusing methods (cito:usesMethodIn), or reporting or comparing the cited work's results (cito:usesConclusionsFrom). Furthermore, CIC also employs a residual label (cito:citesForInformation) for valid contexts that do not reliably align with the three main citation functions.

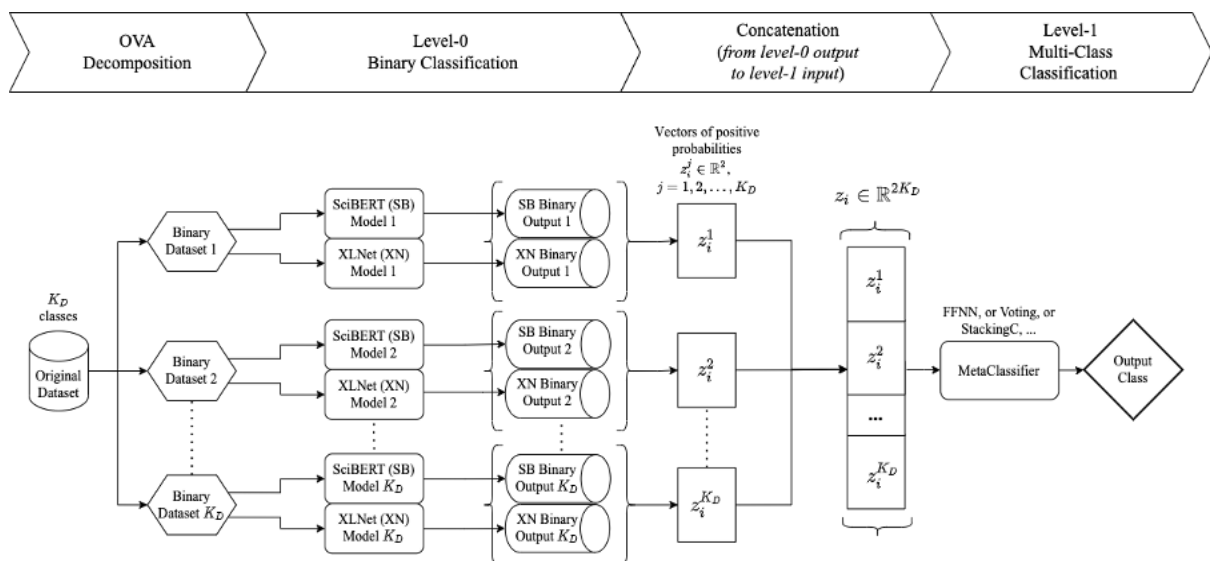


Figure 3 - Working Dynamics of CiteFusion



State of the art and expected upgrade:

The software is complete, publicly available, well-documented, and fully functional. A Docker setup for running both the services is also available on the [GitHub project repository](#). Upon local installation, the extractor and the classifier can be accessed via a web interface (shown in Figures 4 and 5), REST APIs, or the command-line interface.

Future improvements for CEX will focus on enhancing the extraction of section titles, an area that emerged as a bottleneck during the evaluation phase. For CIC, we plan to incorporate SHAP (SHapley Additive exPlanations) ([Lundberg & Lee, 2017](#)) to provide explainability for classified sentences, enabling dynamic and engaging visual representations of the model's output. Additionally, we will allow users to manually adjust classification thresholds. This feature will help researchers distinguish between reliable and unreliable classifications while accommodating the inherent ambiguity of citation contexts.

Moreover, within the scope of the GRAPHIA project, CEC is currently being evaluated for its performance on Social Sciences and Humanities (SSH) articles. Unlike traditional scientific papers that rely on standard end-section bibliographies, SSH publications frequently utilize non-standard formatting, such as endnotes, narrative citations, and hybrid or footnote-embedded references. These complex styles pose a significant challenge for automatic extraction, as current GROBID models cannot always segment them reliably. Consequently, future development will focus on enhancing the tool's extraction capabilities to better handle the idiosyncratic citation practices prevalent in SSH literature.



CITATION EXTRACTOR – CEX

This tool allows you to upload a single PDF file or one archive containing PDFs to process (accepted archives: ZIP, ZST, TAR and TAR.GZ), process it using Grobid and obtain a JSON file containing all the sentences where an in-text citation is present.

By clicking on "Process File", Grobid will perform the processing of the input PDFs and as a result a zip folder, containing the Grobid xml, a JSON file composed by the citations sentences, and a JSONLD file, will be automatically downloaded to your machine. Be patient, it will take few seconds.

N.B. You can process only **ONE PDF** or **ONE ARCHIVE** at the time.

Nessun file selezionato ☐ Perform semantic alignment of sections' headings
☐ Generate JSONld file
Max Workers

You can choose to perform the semantic alignment of the original article's section headings by ticking the box above. When this option is selected, the JSON file, which contains all the sentences with in-text citations, is enriched with a new key called "ALIGNED SECTION." This key stores the aligned section title if it is found.

Some sample PDFs you can use to test the service:

[PDF 1](#) [PDF 2](#) [PDF 3](#) [PDF 4](#)

Figure 4 - CEX Web Application Interface



Classify Your Sentences

This is a tool to classify your sentences according to the categories presented in the landing page.
You can either enter your sentences in the form below, or upload a JSON file containing them.
The result will be displayed in the form of a table, and you will be able to download it as a JSON file.

It is **important to notice** that:

- if you upload a JSON file, you will have to click the "Classify JSON" button,
- if you enter your sentences manually, you will have to click the "Classify" button.

In both scenarios, the result will be displayed in the table below, and you will be able to download it as a JSON file.
Please, do not forget to select your classification mode before classifying your sentences.

How to use

Manually inserted sentences (Manual upload)

Enter your sentences below in the form of a list of tuples.

Each tuple **MUST** contain both section title and citation context.

The format is a LIST, so please do not forget to put squared brackets around it, and the commas to divide the tuples.

In case you do not have the section title, it has to be replaced by an empty string.

Example:

```
[ ('Introduction', "In his 1945  
essay 'As We May Think',...record [Bush, 1945]."),  
 ('', 'The problem has only ...  
nine years [Bornmann and Mutz, 2015].'),  
 (... , ...),  
 ...  
 ]
```

Automatically inserted sentences (JSON upload)

Or you can upload a JSON file:

The structure of the JSON file has to follow some rules.

Each data entry needs an ID as key, and as value it asks for a dictionary with two keys: 'SECTION' and 'CITATION'. The value of 'SECTION' is a string of the section title, and the value of 'CITATION' is a string of the citation context.

For empty elements, the value of 'SECTION' has to be an empty string.

Example:

```
{  
  ID1:{  
    'SECTION': 'Introduction',  
    'CITATION': 'In his 1945 essay  
'As We May Think',...record [Bush, 1945].'  
  },  
  ID2:{  
    'SECTION': '',  
    'CITATION': 'The problem has  
only ... nine years [Bornmann and Mutz, 2015].'  
  },  
  ID3: ...  
}
```

Select your classification mode:

Mixed

With Section Titles

Without Section Titles

Scegli file

Nessun file selezionato

Classify JSON

[('Section title 1', 'Citation context 1'), ('Section title 2', 'Citation context 2'), ...]

Classify

Figure 5 - CIC Web Application Interface



make use of resources available through the IIIF protocol. Utilizing this API requires awareness both on the side of data providers and clients willing to use this data. Within cultural heritage institutions, more and more data sources supporting IIIF Change Discovery API appear, however, sharing their data more widely requires preparing tools that will be able to discover that data. DACE, as an example of the aggregation tool, is a natural candidate for extension with mechanisms supporting the discovery of objects using the IIIF Change Discovery API. This will allow DACE to be enriched with new capabilities based on the latest achievements in the area of sharing high-resolution objects based on the IIIF protocol and, on the other hand, it will encourage data providers to make their services better by implementing support for IIIF Change Discovery, allowing them to increase the visibility of their digital objects at the same time.

Another significant problem for cultural heritage is the quality of digital objects metadata. On one hand, digital objects are often described improperly but on the other hand their metadata lacks contextual information. Such information could be helpful in terms of building the knowledge graph associated with cultural heritage digital objects. Cultural heritage institutions often struggle with the problem of lack of human resources needed to verify and enhance digital objects metadata. This is the area that AI could be used to help with the creation of proper and rich metadata descriptions. DACE processing workflow is a natural place to include the AI module, shaping the digital objects metadata transformation from the original form to the cleaned, enhanced and enriched schema that will be able to feed the GRAPHIA knowledge graph.

The expected upgrade of the DACE system would consist of the following steps:

- implementation of an additional microservice supporting aggregation of digital objects via IIIF Change Discovery API,
- implementation of an additional AI plugin for record processor that will utilize LLMs for cleaning and enriching objects metadata,
- implementation of an additional mechanism feeding the GRAPHIA knowledge graph with information found by the AI plugin, leveraging the data ingestion interfaces developed within WP2.



EHRI LOD Knowledge Graph and the EHRI Virtual Observatory (EHRI Portal (and APIs) - EHRI Document Blog - EHRI Geospatial Repository)

Organisation: NIOD-KNAW

Short description: EHRI's online platform, the EHRI Virtual Observatory, offers a wealth of resources to support researchers in their work. It features a comprehensive directory of institutions holding Holocaust-related collections, online digital editions of primary source material, and tools for visualising and contextualising historical documents.

Three specific services in this observatory are of relevance to the GRAPHIA project.

1. The EHRI Portal (and APIs)

The EHRI Portal offers access to information about institutions holding Holocaust-related archival material and metadata relating to their collections. It is an expanding resource that currently hosts reports detailing the Holocaust history and archival situation in 61 countries, information about more than 2,300 Holocaust-relevant archival institutions, and hundreds of thousands of descriptions of archival materials held by more than 870 institutions across Europe and beyond. All this information can be conveniently explored via dedicated tools for searching and browsing.³

The EHRI Portal provides two JSON-based APIs: a ReST-style search API, and a GraphQL API. Other forms of structured data can be provided by exporting archival descriptions, institution descriptions, and authority files as XML using, respectively, the Encoded Archival Description (EAD 2002), Encoded Archival Guide (EAG 2012) and Encoded Archival Context (EAC 2010) standards. These files are dynamically generated from the portal's database. Additionally, the Portal can provide data via OAI-PMH in Dublin Core, EAD 2002, or EAD-3 formats.

³ Tobias Blanke, Michael Bryant, Michal Frankl, Conny Kristel, Reto Speck, Veerle Vanden Daelen, and René Van Horik. 2017. The European Holocaust Research Infrastructure Portal. *J. Comput. Cult. Herit.* 10, 1, Article 1 (April 2017), 18 pages. <https://doi.org/10.1145/3004457>

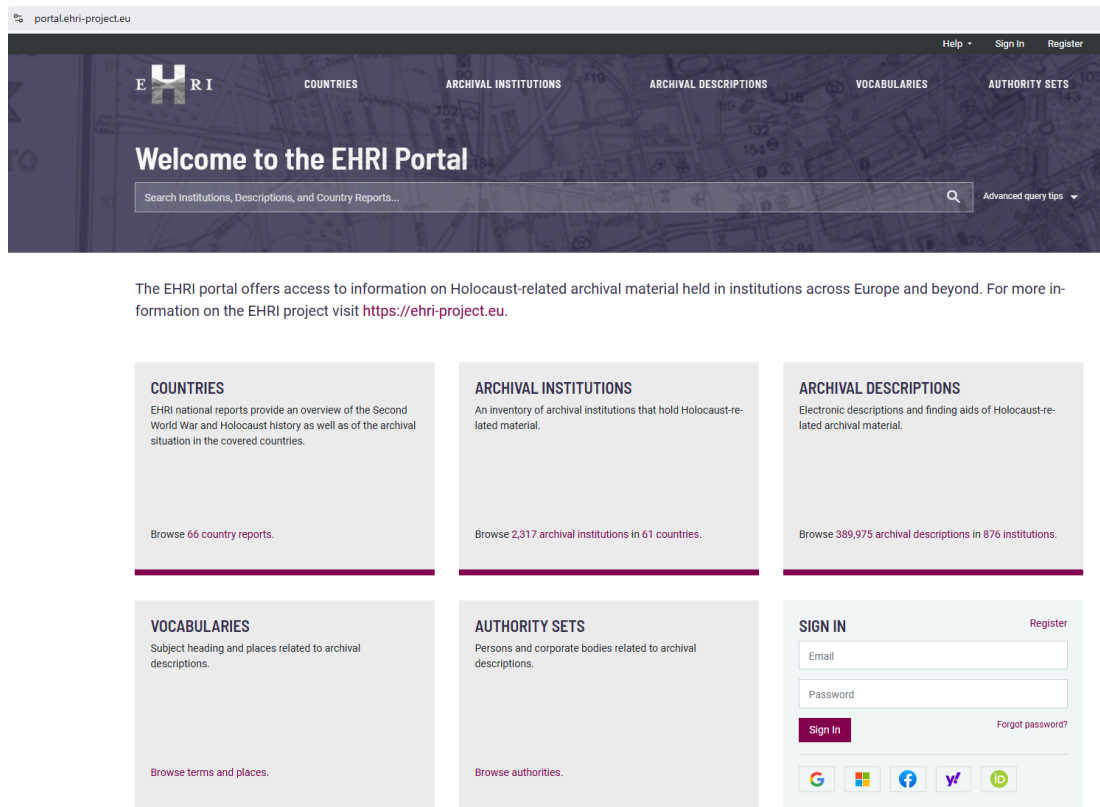


Figure 7 - Screenshot of the EHRI Portal (<https://portal.ehri-project.eu/>)

2. The EHRI Document Blog

The EHRI Document Blog is an open, experimental space that tackles questions related to Holocaust documentation, sources, and digital methodology. While allowing discussions on any issues and questions related to Holocaust documentation, the platform is particularly informed by the provenance and history of the Holocaust-related collections and sources, especially considering their destruction during World War II and the often circuitous path of the surviving materials.⁴

The Document Blog doesn't mandate a specific format or technology—contributors can use any method that can be embedded (or even linked to) into the blog via iframes, thereby allowing for flexibility and openness. A common approach to providing interactive elements within blog posts involves using EHRI's visualisation platform, which uses Omeka and its popular Neatline plug-in for constructing

⁴ Michal Frankl, "Blogging as a Research Method? The EHRI Document Blog" in Holocaust Research and Archives in the Digital Age, eds. Laura Brazzo and Reto Speck, Quest. Issues in Contemporary Jewish History. Journal of the Fondazione CDEC, n. 13, August 2018 DOI: 10.48248/issn.2037-741X/824

interactive presentations that locate items (for example, maps or documents) in space and time alongside textual/narrative elements.

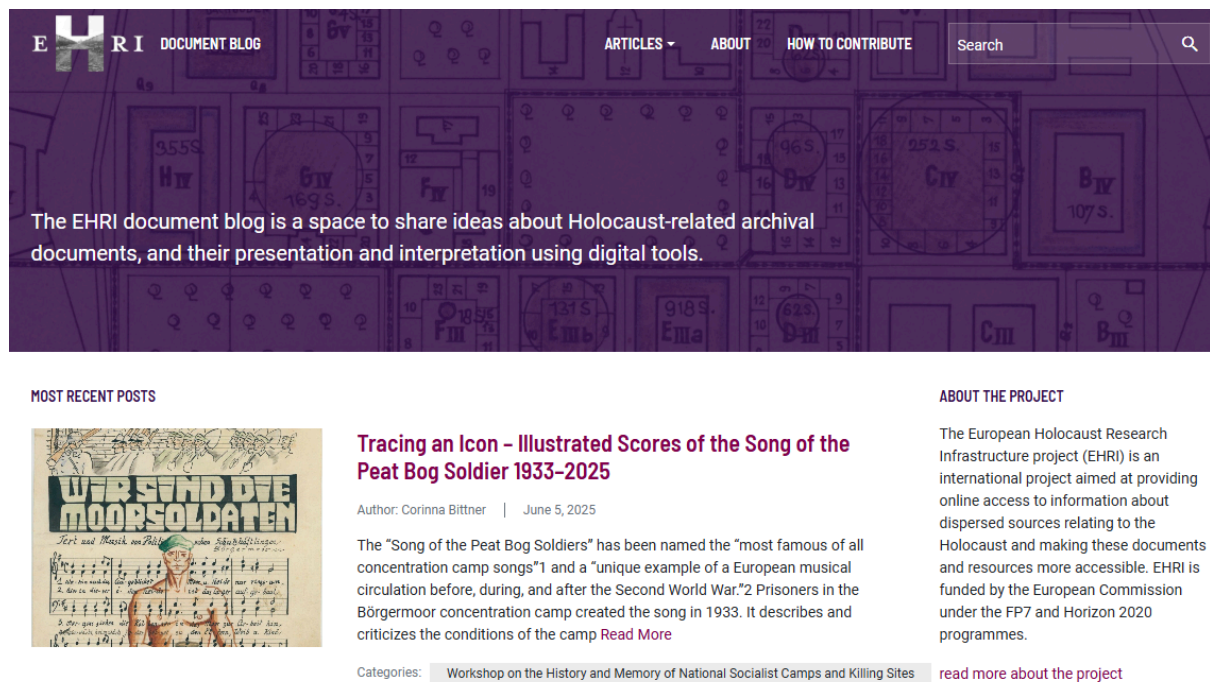


Figure 8 – Screenshot of the EHRI Document Blog (<https://blog.ehri-project.eu/>)

Since 2025, articles on the EHRI Document Blog are assigned DOIs, linked to [external landing pages](#).

3. The EHRI Geospatial Repository

By providing access to data about Holocaust-related places and spaces, the EHRI Geospatial Repository facilitates research driven by spatial and geographic approaches. It builds on the spatial turn in the field of Holocaust Studies: over the past years, the research in Holocaust geographies, as conducted by the Holocaust Geographies Collaborative and others, advanced new perspectives and methods. The EHRI Geospatial Repository provides researchers, projects and organisations with the means to: store and share geodata in a standard compliant, reusable way; search and explore datasets based on their metadata use; download and combine datasets for mapping, data enrichment or spatial analysis. Datasets can be uploaded and saved in a variety of formats (such as shapefiles, CSV, GeoTIFF, etc.). EHRI's staff will check the data, standardise it if needed and internally store it as a GeoPackage. In a next step, the dataset will be pushed into the EHRI Geoserver instance (providing

standardised services such as WMS, WFS, etc.) and linked to the Geospatial Repository which serves as a metadata store.

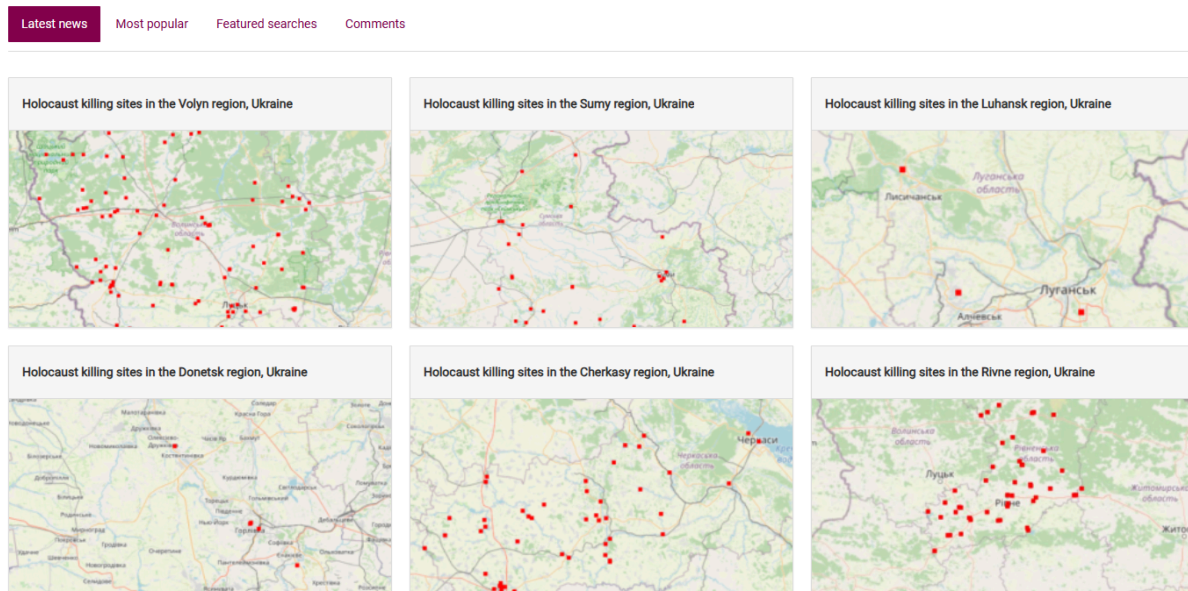


Figure 9 - Screenshot of the EHRI Geospatial Repository (<https://geodata.ehri-project.eu/>)

The EHRI Knowledge Graph (or KG)

By nature, Holocaust data are dispersed, heterogeneous (e.g. multilingual) and complex (sometimes even incomplete). To bring more organisation and semantic coherence to the content that can be found in EHRI's Virtual Observatory, the EHRI Knowledge Graph (or KG) is currently being developed. This Linked Open Data model will closely follow the Records in Contexts (RiC) standard, augmented by properties derived from schema.org and a selection of EHRI-specific elements.⁵

The EHRI KG will provide a SPARQL endpoint and a web-based navigable interface powered by the LODlive tool, allowing users to explore key classes and their relationships in an interactive manner. When deployed in production the KG will be kept continuously in sync with the information on the EHRI Portal from which it is derived.

⁵ García-González, H., Bryant, M. (2023). The Holocaust Archival Material Knowledge Graph. In: Payne, T.R., et al. The Semantic Web – ISWC 2023. ISWC 2023. Lecture Notes in Computer Science, vol 14266. Springer, Cham. https://doi.org/10.1007/978-3-031-47243-5_20



The country reports constitute the main entry point to exploring the data in the EHRI portal. They offer an introduction to the Holocaust research and landscape in a certain country. Therefore, we propose that non-experts users start browsing this KG from one of the proposed countries in order to get an idea of the potential of this new tool. Below you can find a comprehensive list of the countries represented in the EHRI portal.



SPARQL EXAMPLES

In order to show how data is organised and how data can be consulted we provide different examples of SPARQL queries. They cover simple queries, some aggregates to obtain statistical data and some federated queries that query CDEC's and DBpedia's data to expand this KG possibilities.

Federated queries show the potential that semantic technologies can suppose for users. However, even though the federated queries exposed here are carefully selected, take into account that modifying and expanding them can derive into unexpected long computation times.

Figure 10 - Screenshot of the EHRI Knowledge Graph prototype's LODLive interface, featuring web-based navigation of key entities and their relationships: (<https://lod.ehri-project-test.eu/>)

State of the art and expected upgrade: By leveraging AI-based solutions developed in GRAPHIA and connecting the existing EHRI Collection Graph with the SSH KG, EHRI data will be significantly enhanced and more tightly integrated into the overall SSH data ecosystem. This will require assessing the compatibility between the Records in Contexts (RiC) model used by EHRI and the GRAPHIA data architecture, a mapping exercise planned in coordination with WP2. Alongside other “access points” such as the names of the creators and significant places, subject terms characterise the material being described using one or more labels from a controlled vocabulary to narrow the scope of a query being undertaken by cataloguers or end-users. We propose to develop and deploy multilabel classification tools developed specifically for archival cataloguing in both fully-automated and supervised settings, leveraging both ML and LLM approaches for scenarios permitting different degrees of efficiency and accuracy, and where training material is more or less available. Building on established MLC tools like Annif, we will ensure modular integration of both local and cloud-based LLM technologies and conduct qualitative studies of labelling effectiveness in realistic cataloguing settings.



HYPERMNESIA: A Framework for Reinforced Research Traceability in Heritage Science Open Data

Organisation: CNRS MAP

Short description: HYPERMNESIA is a data provenance framework composed of a series of tools dedicated to Heritage Science instrumentation and documentation workflows. The corner stone is the W7 conceptual model ([Ram and Liu, 2009](#) ; [2017](#)) used as a structured yet user-friendly metadata and paradata framework enabling to document and describe digital data lifecycle from acquisition to processing and analysis⁶ in HS interdisciplinary practices. It is composed of several tools integrated in the DIGILAB Catalogue of Service managed by the CNRS-MAP aligned with the E-RIHS IP D5.1 DIGILAB Implementation Plan⁷. It relies on several interoperable layers to foster semantic modeling approaches at the core of HS practices from metadata schemes, to controlled-vocabularies mapping and ontology management. From the 4 instruments available in the [DIGILAB CoS](#), GRAPHIA related tasks within this WP3 will focus on [ANAMNESIS](#) but complementary tools are mentioned below to provide a full understanding of the interoperability framework and issues.

- **Web-based platform (ANAMNESIS):** Is a research-born prototype initially developed within an eponymous project funded by the Fondation des Sciences du Patrimoine. The output is fully operative and open source web-service (back-end, front-end and documentation) enabling to manage W7 structured metadata harmonized with PIDs framework in order to link and enrich from the data source all the digital resources created by HS instrumentation workflows ([Pamart et Al., 2026](#)). The users can customize and share technical schemes according to their own practices and activities. Used as an early enrichment process ANAMNESIS can reinforce FAIR-by-design practices including SSH scholars with no experience in data-science and web-semantic.

⁶ Anthony Pamart, Livio De Luca, Philippe Véron. A metadata enriched system for the documentation of multi-modal digital imaging surveys. *Studies in Digital Heritage*, 2020, 6 (1), pp.1-24. [⟨10.14434/sdh.v6i1.33767⟩](#). [⟨hal-03844472⟩](#)

⁷ De Luca, L., & Guillem, A. (2024). E-RIHS IP D5.1 DIGILAB Implementation Plan. Zenodo. <https://doi.org/10.5281/zenodo.13622848>

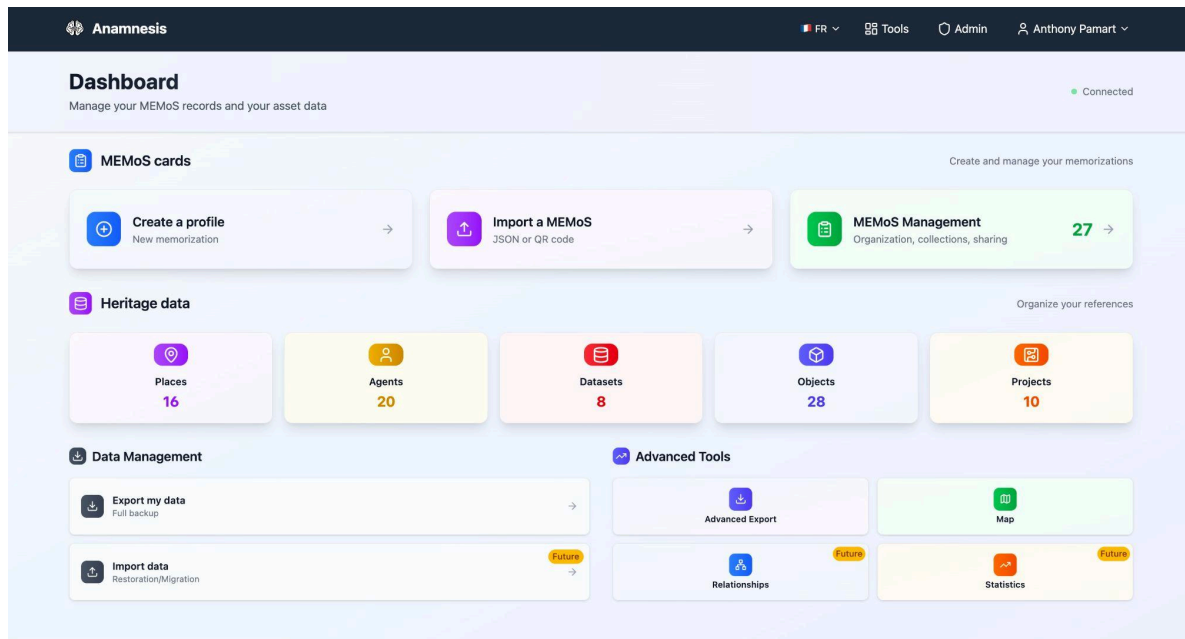


Figure 11 – Anamnesis frontend dashboard

- **Other tools and instruments:** Other layers are necessary to improve data and workflow provenance. One could think about bottom-up enrichment to retrieve and structure metadata from already published resources. To this aim, [METAREVE](#) offers an LLM+NLP approach to extract W7 structured metadata from text and audio files. When possible ANAMNESIS relies on existing PIDs but controlled-vocabularies could be used to stabilize any kind of semantic entities. The tool [Open-theso](#) is a collaborative and multilingual vocabulary management system similar to [Skomos-based](#) instance [Vocabs](#) provided by DARIAH to manage SSHOMP concepts and classes. Managing those layers of semantic modeling is challenging, hence E-RIHS DIGILAB task-force choose [OntolPortal](#) as community-driven tool. The [HSPortal](#) is accessible through the E-RIHS CoS and will be used across all national nodes to share, stabilize and harmonize the management of semantic features.

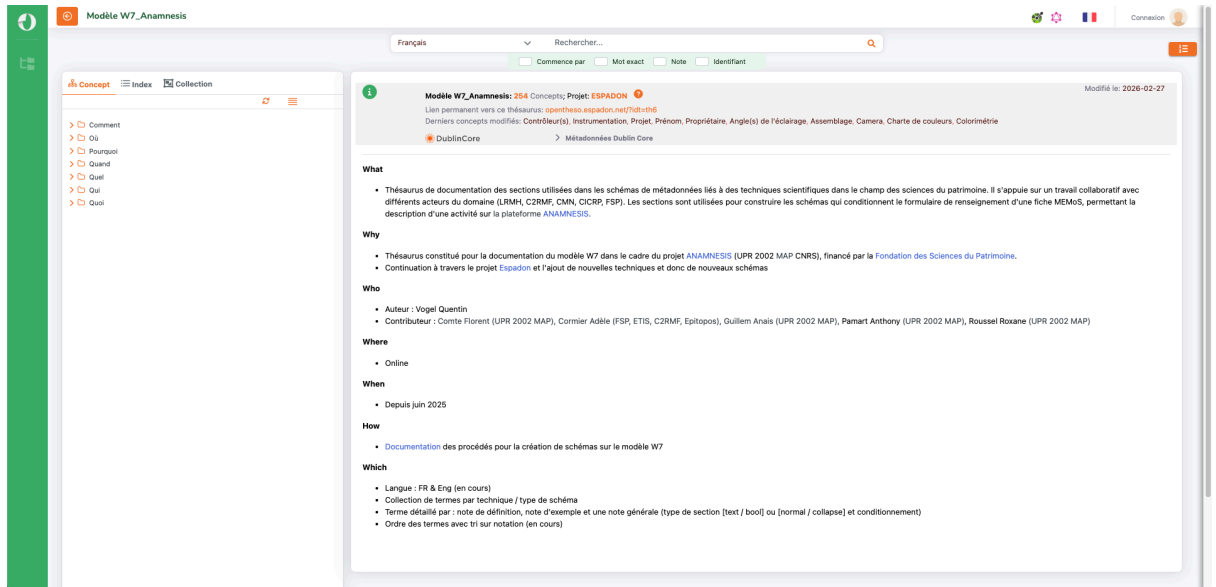


Figure 11.1 - W7 ANAMNESIS concepts (247 entries) referenced in [OpenTheso](#)

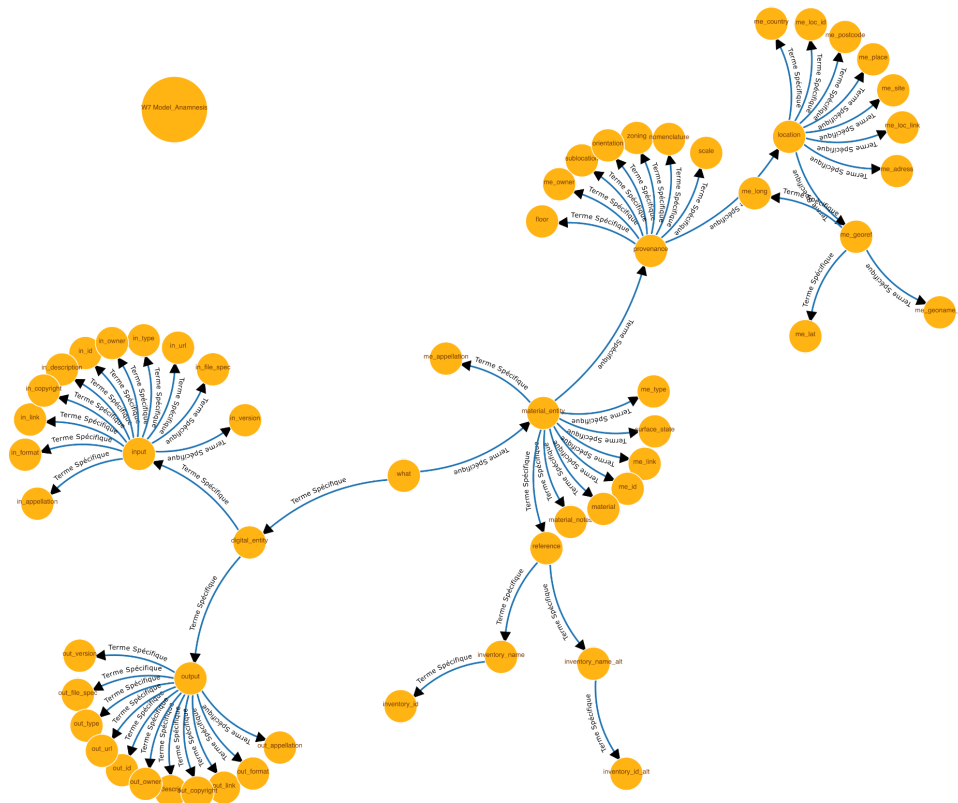


Figure 11.2 - Graphical representation of the W7.WHAT section and hierarchical concepts in [OpenTheso](#)



- ***Semantic modeling (W7)***: the use of this semantic framework could be seen as a contribution in itself. At the CH and HS application level, it has been demonstrated as a versatile and user-friendly approach raising FAIR-by-design awareness and sensitivity toward non-expert SSH practitioners. It enables early conceptual modelling tailored to research data provenance reinforcement. The W7 is an event-oriented semantic structure interconnecting high level concepts such as agent (WHO), time (WHEN), space (WHERE) object (WHAT), instrument (WHICH), method (HOW) and context (WHY) to formalize a wide range of activities. This scheme is versatile enough for already being used in environmental sciences (astronomy and biology) with a potential cross domain application in KG approaches (e.g. LUMEN). CH and HS wise the W7 structured is compliant with CIDOC-CRM and PROV ontologies through mapping approach⁸. Recent works within WP3 have validated the structural compatibility of the W7 model with the GRAPHIA Data Architecture and Data Acquisition Platform, through a mapping exercise with the resource-oriented GoTriple Data Model. This preliminary alignment confirmed sufficient coverage of the W7 metadata categories for GRAPHIA ingestion purposes.

⁸ Anaïs Guillem, Violette Abergel, Roxane Roussel, Florent Comte, Anthony Pamart and Livio De Luca. Bridging the Provenance Knowledge Gap between 3D Digitization and Semantic Interpretation. Heritage Special Issue “Advanced Data Environment in Current Cultural Heritage 3D Digitization Practices, 2025, submitted.

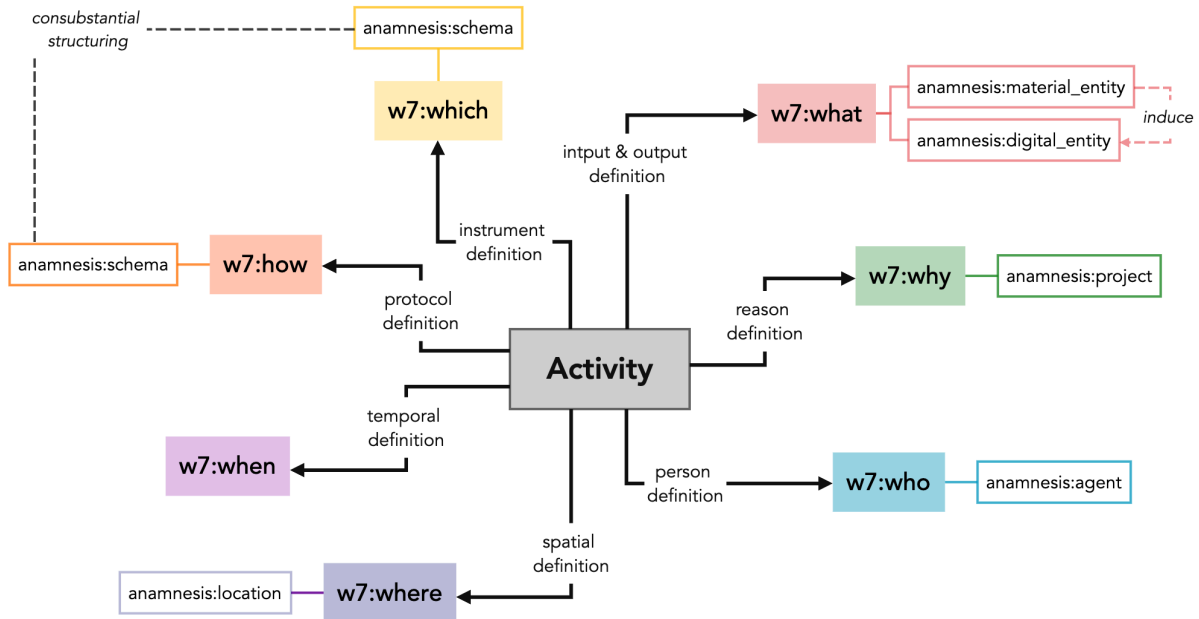


Figure 12 - Anamnesis W7 metamodel for activity description

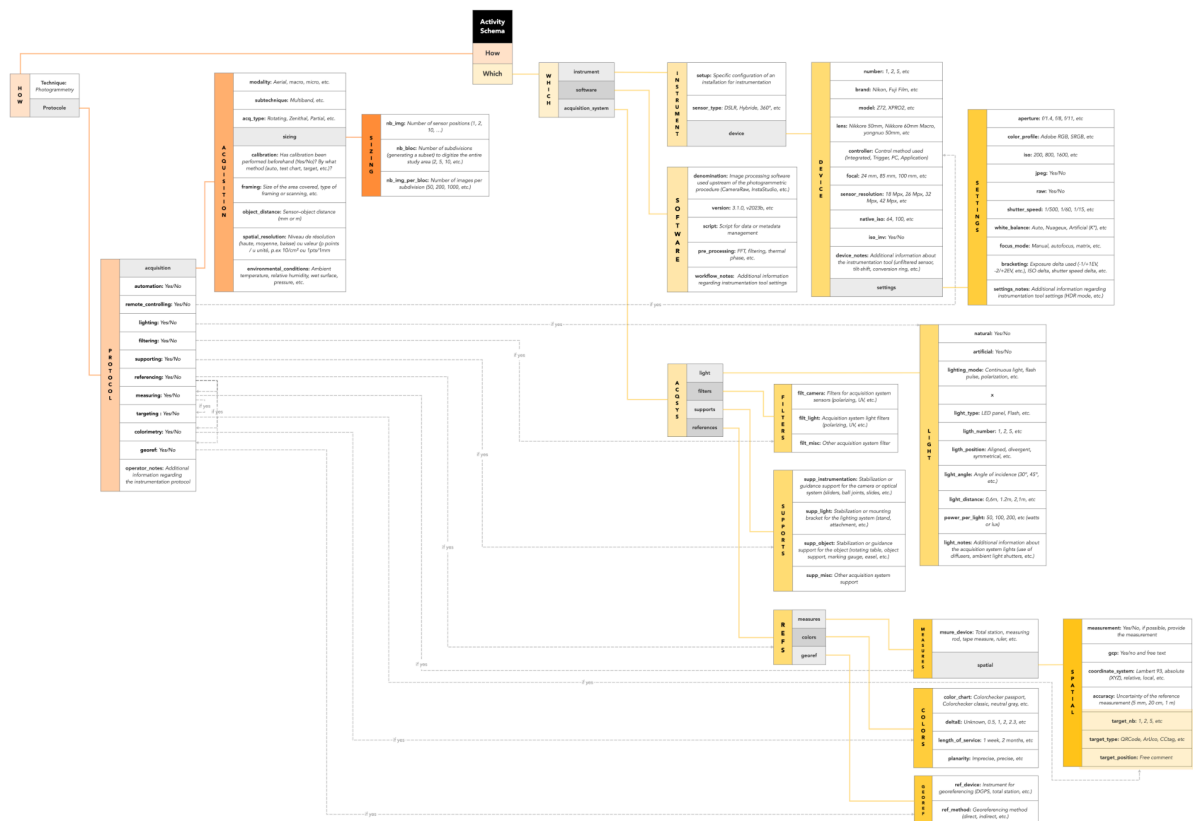


Figure 13 - Example of metadata scheme for the documentation of photogrammetric capture activity
(from Wiki-based ANAMNESIS user documentation)



State of the art and expected upgrade:

ANAMNESIS is accessible in beta version since September 2025. The instrument has been presented in several research driven contexts related to E-RIHS partnership network ([EU funded project DIGITAL-3D4CH](#), training camp related to EU funded project ECCCH-ECHOES, [educational event related to DARIAH](#)) with a common enthusiasm for an intuitive tool bridging the gap between data providers, data curators and data scientists. The beta-testing revealed some bugs and minor corrections along with possible new features drafting a new release during GRAPHIA project. It will be upgraded by transitioning from HAL which is a French PID registry toward extra-national authority files like ORCID and CORDIS to ensure long-term resolvable identifiers for researchers and projects. The semantic stabilization will also be tackled by interoperative reinforcement with all components of HYPERMNESIA (OpenTheso for thesaurii management, Metareve for LLM-based curation, HSPortal as supervision tool of semantic features). ANAMNESIS integrates an advanced export with a mapping function making it already compatible with GRAPHIA interconnection by filedumps (curated JSON). However a proper API will upgrade the instrument from internal (interoperability within HYPERMNESIA) and external sides (populating SSHOMP, connection to GRAPHIA KG, harvested by GoTriple). Additionally a complete local translation of the website into English will be achieved. The documentation package will be upgraded with hands-on tutorial and training materials to ease user adoption and leverage a community of users, potentially providing in a close future KG-ready or KG-compliant open data set. An extension from data provenance to innovative workflow provenance will be the focus of the 3.2 task.



- ***DIGIT-SPIDER instrument prototype (MOLAB/FIXLAB):*** SPIDER is a multi-camera rig automatizing image-based 3D modeling (photogrammetry) dedicated to massive digitization of medium size CH cubical-shaped objects (between 30 cm and 80 cm). It enables fast and efficient generation of 3D digital asset collection. This instrument was used to create 200 3D models.
- ***DIGIT-STUDIO instrument prototype (MOLAB/FIXLAB):*** STUDIO is a similar rig but dedicated to small to extra-small size CH cubical-shaped objects (between 30 cm and 3 cm using its macro extension). This instrument was used to create 950 3D models.
- ***DIGIT-ARCH instrument prototype (MOLAB/FIXLAB):*** ARCH is a multi-camera arch-shaped rig semi-automatizing image-based 3D modeling (photogrammetry) dedicated to massive digitization of multi-scale CH long-shaped objects. This instrument was used to create 130 3D models.
- ***Other potential prototypes to be integrated in the CoS (3):***
 - SQUILLIDEA is a 360° multi-camera rig semi-automatizing image-based 3D modeling (photogrammetry) dedicated to massive digitization of complex CH spaces. It has been developed and optimized to digitize complex and inaccessible architectural space (Notre Dame forest, and Citadel of Marseille underground). It's a prospective and innovative prototype also extensible to multimodal capture (space and sound) and offers immersive perspective through XR extension currently under development.
 - CROSSBOW is a robotized rig enabling high-resolution multiband photogrammetric orthomosaicing for quasi-planar vertical surfaces (canvas, wood-painting, wall-painting etc..). This device is co-realized and co-managed with another E-RIHS France service provider (CICRP). New research-driven initiatives arising in 2026 can accelerate its integration into the CoS.
 - SPHERE is a patented technology available from CNRS since 2018. This system was conceived and developed by the start-up Mercurio Imaging (spin-off from MAP-CNRS) composed of 80 cameras and lights. For a decade, the data captured by this instrument combining multi-view and multi-light modeling was not exploitable because of the paradigm shift. Nowadays, new AI based algorithms unlock new scientific prospects exploring this unique device.

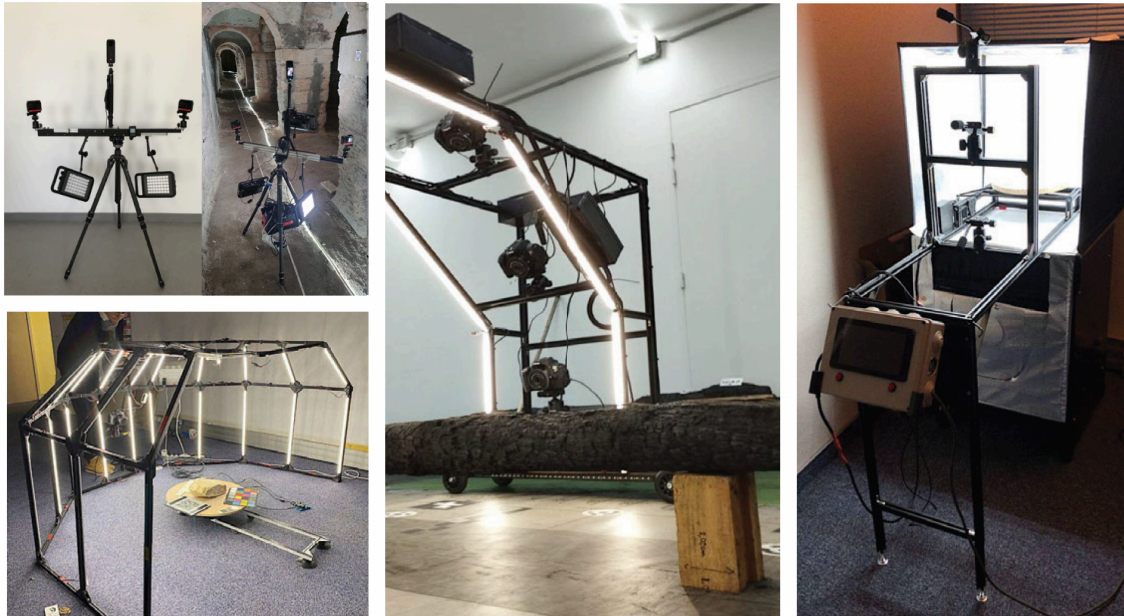


Figure 14 - Massive 3D digitization instrumental platform ; SQUILLIDEA (top-left), DIGIT-SPIDER (bottom-left), DIGIT-ARCH (middle) and DIGIT-STUDIO (right).

- COCORICO united software architecture prototype (DIGILAB):** All the instrumental prototypes and tools above-mentioned are not currently supported by a unified and optimized open-source processing framework. The COCORICO project starting in 2026 for 18 months is funded by the CNRS Innovation OPEN program as part of its pre-maturation phase, focusing on the consolidation of open-source software architectures and interoperability frameworks to prepare for wider adoption and higher Technology Readiness Levels (TRLs). The current prototype, called [MicMacRoom](#) (integrated in SSHOMP) focus on interoperable 2D to 3D image processing based on the two leading open-source solutions for image-based modeling (namely MicMac and AliceVision/Meshroom). The COCORICO project will provide a new set of tools in the form of orchestrated plug-ins managed by the MeshroomResearch graphical environment.

State of the art and expected upgrade: The MAP-CNRS massive digitization platform is an exclusive set of instruments for creating CH 3D digital assets and semantically enriched collection. It has been mostly used in the scientific worksite for the restoration of Notre Dame de Paris providing a collection in construction of 1500 3D models publicly available in ETALAB 2.0 licence accessible in the Dataverse



instance [RechercheDataGouv](https://recherche.data.gouv.fr/). Its integration to E-RIHS MOLAB/FIXLAB CoS jointly to GRAPHIA's catalogue is an opportunity to enrich the current initiatives toward institutional 3D data corpus (e.g. 3D Big Data for the Data Space for Cultural Heritage). GRAPHIA related works will focus on making interoperable this growing data source to be harvested and to feed federated KGs. To this end the HYPERMNESIA framework is the keystone to create KG-ready datasets straight out from innovative instruments while a united APIs framework could streamline the data to metadata anchorage as part of the catalogue of innovative workflows (Task 3.2) completing the current catalogue of instruments. Additionally the DIGIT-STUDIO is expecting a hardware upgrade for technical improvement toward cross-polarization and multi-view stereophotometry to replicate more consistently and accurately the optical appearance of complex materials with high specularly.

MOVIDA

Organisation: CNR (ISPC)

Short description: Movida 2.0 is a new digital application for the collaborative management of data derived from multi-technique, non-invasive investigations of artworks. It is a tool of the Collaborative Virtual Research Environment of the Digilab-IT platform that enables the storage, analysis, visualization, and interoperability of data from diagnostic studies in the field of Cultural Heritage. It will support data lifecycle management for various types of investigations, including point analyses, imaging/mapping techniques, and geophysics analyses. The data produced with the application will then be mapped on the scientific investigation ontology (based on the CIDOC CRM extension CRMSci), becoming part of the Digilab-IT knowledge base. Once the data is published in the semantic data catalog, it will become available for processing, analysis, and visualization through the digital tools that are being implemented or integrated in the platform. The datasets will also be published in the Data Portal to ensure compliance with the FAIR principles and to support Open Science and data reusability.

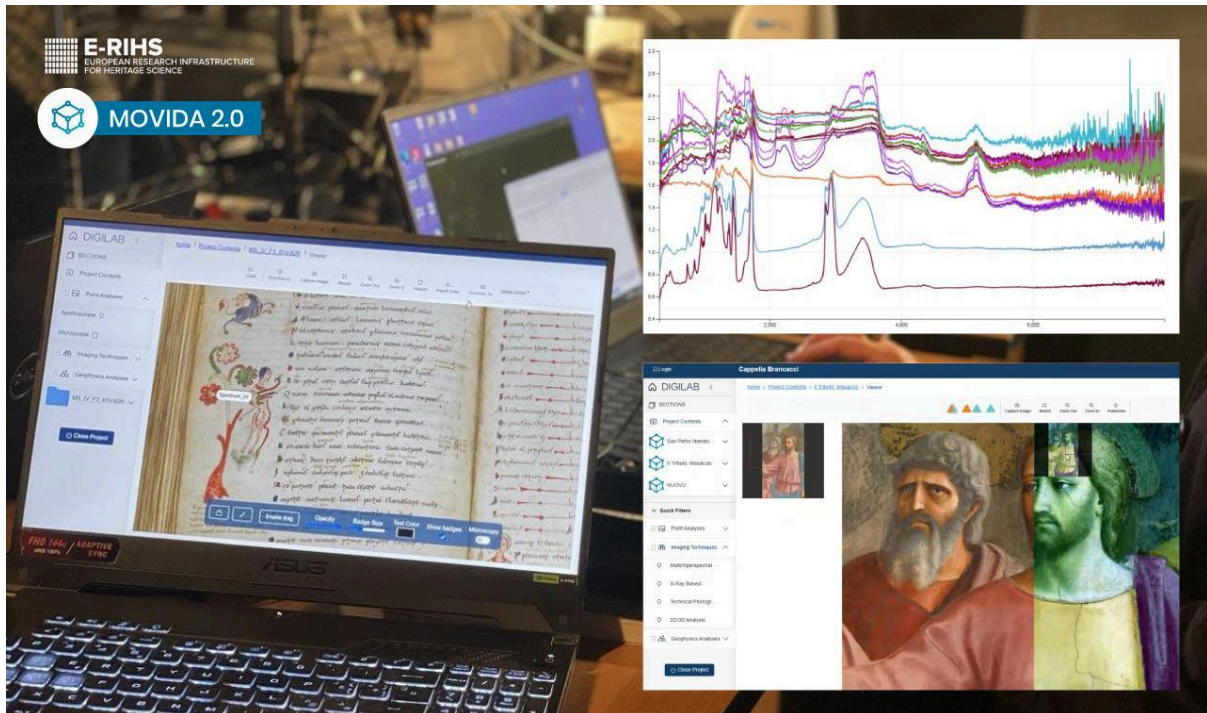


Figure 15 - MOVIDA 2.0 interface

State of the art and expected upgrade: The application is a tool of the DIGILAB platform of the Italian node of the E-RHIS research infrastructure, currently being developed within the H2IOSC project. The core system is approaching TRL 7: it is functional, deployed in an on-premises staging environment closely aligned with the production configuration, and is already being used by researchers to test representative workflows and data. The production release for first users is expected in the coming weeks. Current development activities are focused on the implementation of the ontological mapping pipeline based on CIDOC-CRM (CRMSci extension) for translating the platform's hierarchical data into a Knowledge Graph, with the aim of strengthening interoperability and semantic integration within the Graphia ecosystem.

MuSeSca3D – Multi Sensor Scanner 3D (Integrated MA-XRF / RIS / LIS System)

Organisation: The Cyprus Institute -APAC Labs

Short description: MuSeSca3D is an innovative multi-sensor 3D scanning platform that integrates Macro X-Ray Fluorescence (MA-XRF), Reflectance Imaging



Spectroscopy (RIS), and Luminescence Imaging Spectroscopy (LIS) into a single, compact system. By enabling simultaneous, non-destructive elemental, spectral, and luminescence analysis — including on complex non-flat objects through 3D data integration — MuSeSca3D delivers comprehensive material characterisation with unprecedented efficiency and resolution. Designed for applications in cultural heritage, archaeology, materials science, forensics, and environmental research, the system streamlines analytical workflows while enhancing data fusion, interpretation, and decision-making.

State of the art and expected upgrade: The system currently operates at Technology Readiness Level (TRL) 7 – system prototype demonstration in an operational environment, having been successfully validated under real heritage research conditions.

Although not originally foreseen in the list of tools under T3.1, its inclusion provides clear added value by strengthening the link between T3.1 methodological development and T3.3 implementation in the Cyl case study. It allows developed workflows and protocols to be tested and demonstrated in a real operational context.

The tool is already registered within the E-RIHS Catalogue of Services. However, it is not yet publicly accessible due to an ongoing patent process. The expected upgrade will further enhance system integration and operational robustness, facilitating its transition towards full deployment and service provision.

ORKG

Organisation: TIB

Short description: The Open Research Knowledge Graph (ORKG) is a structured, machine-actionable platform designed to transform scholarly knowledge from unstructured publications into semantically rich, interoperable data. By combining symbolic knowledge representation with neural-network-based information extraction, ORKG encodes research contributions (including methods, results, and experimental setups) into a FAIR format. Its RDF-based graph model, built on domain-specific ontologies, enables the creation of dynamic comparisons and systematic reviews across scientific publications. This infrastructure supports machine-assisted curation through collaborative workflows, where AI extracts initial data from articles and human experts refine it using standardized templates, ensuring consistency and quality. By converting static document-centric knowledge



into a queryable knowledge graph, ORKG accelerates discovery, cross-study analysis, and interdisciplinary knowledge sharing.

State of the art and expected upgrade:

During the second period of 2025 and the beginning of 2026, ORKG has achieved substantial progress in terms of scale, functionality, and technical maturity. The platform has significantly expanded its content coverage, increasing the number of structured research papers from 34,000 to 53,000 (+56%) and growing machine-actionable comparisons from 1,200 to 1,400 (+17%). Additional knowledge artifacts have also expanded, with visualizations rising from 341 to 351 and structured reviews from 45 to 61, reflecting continuous uptake by the research community.

In parallel, new capabilities have been introduced to improve automation and usability. The ORKG Comparison Creator enables semi-automated generation of comparison tables, reducing manual curation effort and accelerating the structuring of scholarly knowledge. Furthermore, the next-generation publishing approach originally developed under ORKG Reborn has been formally launched as TIB Knowledge Loom, providing a production-ready environment for generating “born machine-readable” research outputs directly within computational workflows.

These functional enhancements are supported by substantial technical improvements. The ORKG frontend and backend have been comprehensively reengineered, accompanied by an updated REST API that strengthens interoperability and integration with external systems. In addition, server-side performance has been significantly improved, enabling faster query processing and better scalability.

Within the GRAPHIA project, ORKG's role focuses on providing structured, machine-actionable representations of SSH research contributions that can serve as a federated knowledge source for the GRAPHIA SSH Knowledge Graph. The alignment between ORKG's RDF-based data model and the GRAPHIA data architecture will be further explored in coordination with WP2.

Together, these developments reinforce ORKG's position as a scalable, interoperable infrastructure for machine-actionable scholarly communication.



Figure 16 - Screenshot of the ORKG home page (<https://orkg.org/>)

SENSE

Organisation: CNR

Short description: SENSE is an integrated hardware and software instrument developed by the Italian National Research Council (CNR) for applications in cultural heritage, archaeology, and heritage science. Designed to monitor and analyze microenvironmental conditions, SENSE supports advanced diagnostic processes within both research and practical conservation settings. As part of the SSH Knowledge Graph use case at CNR, it contributes to the real-world deployment of next-generation instruments. Currently at Technology Readiness Level 6 (TRL6), it is expected to advance to TRL7, reflecting enhanced maturity and operational readiness.

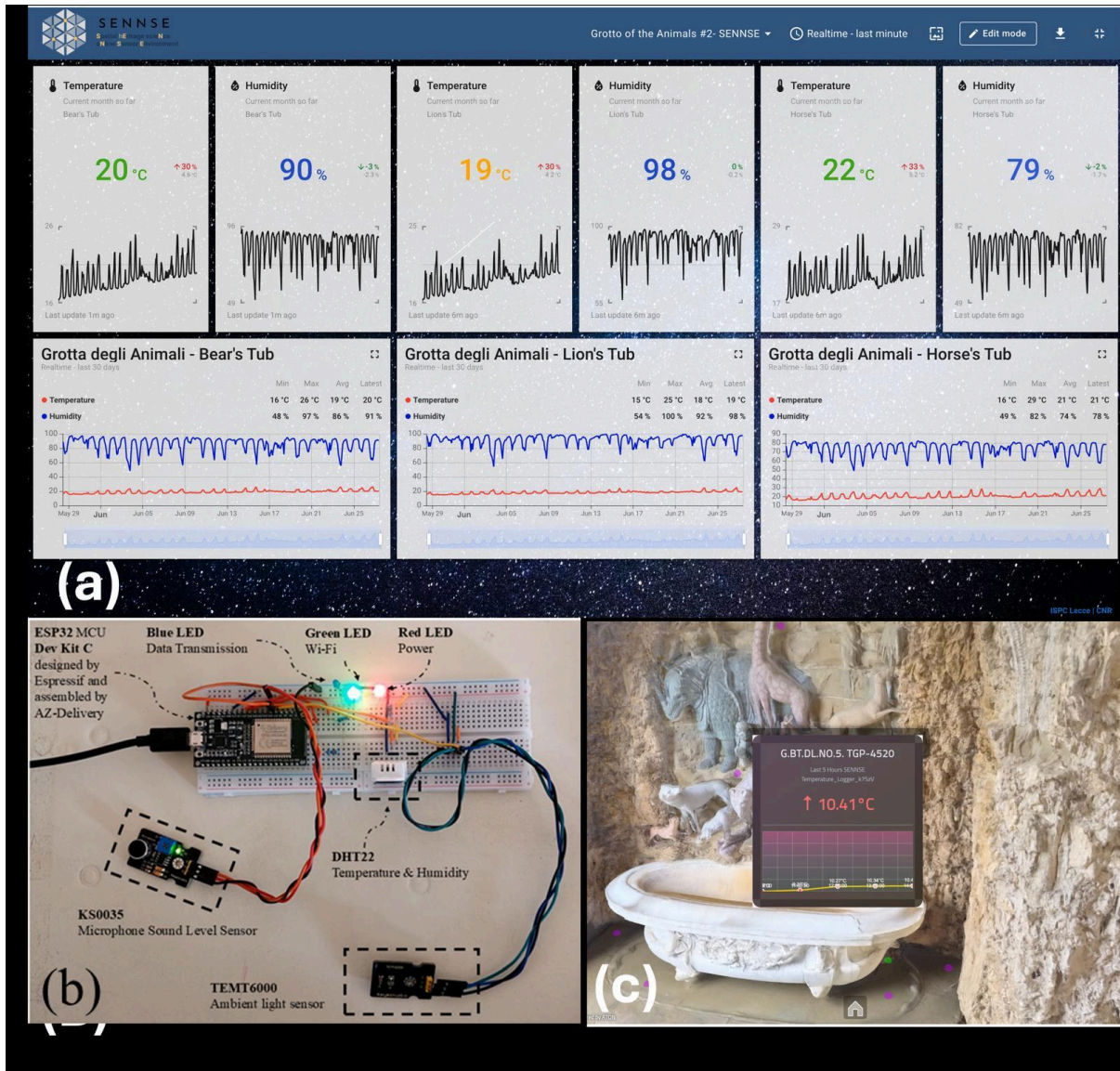


Figure 17 - SENNSE application interface and hardware IoT devices

State of the art and expected upgrade: Currently, SENNSE is used as a tool for a development environment, providing monitoring dashboards, and aiding in data acquisition related to environmental data from cultural heritage sites, such as temperature, humidity, and light. This data is already providing contextual information, helping to interpret cultural artifacts and sites according to their environmental surroundings. An instrumental template scheme has already been integrated within ANAMNESIS to ease early stage metadata enrichment (see Fig x below).



With regard to further developments related to SENNSE, within the GRAPHIA Project, there is a focus on improving data acquisition tools, such as IoT devices, and enhancing the semantic interoperability of SENNSE. Specifically, there is a focus on matching the data generated by SENNSE with the Heritage Science Ontology, so that this data can then be integrated with the Knowledge Graph ecosystem. Future work will assess how this ingestion path can be extended or bridged to feed the GRAPHIA SSH Knowledge Graph developed in WP2.

With this said, there is already a tool, such as the Digilab API, which is currently providing a method to ingest annotated data into the Knowledge Graph. Currently, data generated by SENNSE can be manually prepared and submitted to the Digilab API, while the SENNSE internal API is still under development. Future developments related to SENNSE will allow this process to be automated, where SENNSE can interact with the Digilab API, enabling annotated sensor data to be fed continuously into the Knowledge Graph.

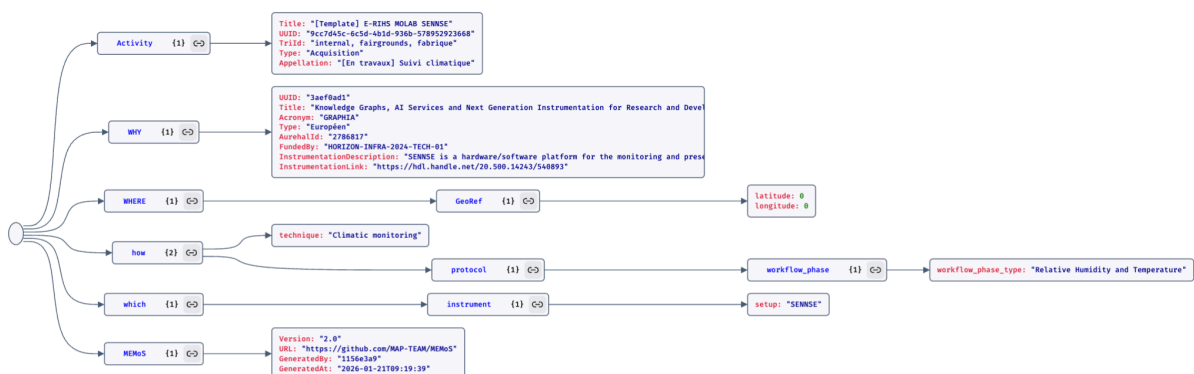


Figure 17.1 – SENNSE metadata template scheme structured in W7 using ANAMNESIS

VTL

Organisation: PCSS

Short description: VTL (Virtual Transcription Laboratory) is an interactive, web-based platform that supports automatic text recognition and collaborative transcription from digitized images. The platform enables users to create personalized projects based on their own image collections or import content through standardized protocols including IIIF (International Image Interoperability Framework) and OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting). VTL provides a comprehensive workflow from image ingestion to text recognition, manual correction, and export in multiple standardized formats including



TEI (Text Encoding Initiative), hOCR, PDF, and EPUB. The platform incorporates machine learning capabilities that improve recognition accuracy over time through user corrections, creating a feedback loop that enhances the OCR model's performance for similar document types and languages.

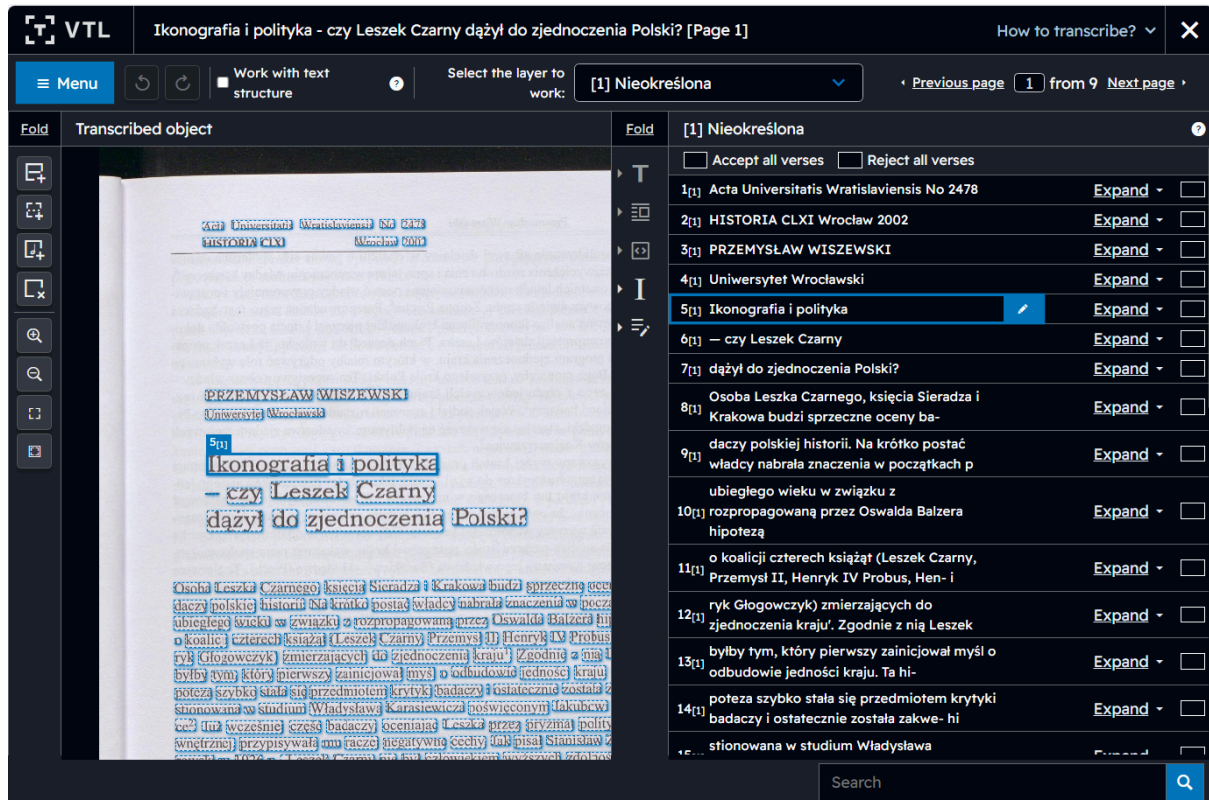


Figure 18 - VTL interface

State of the art and expected upgrade: Currently, VTL operates primarily with natural language text recognition using Tesseract OCR technology, supporting standard contemporary texts with established language dictionaries.

The expected upgrades encompass two major areas of development:

1. **Enhanced OCR capabilities for historical documents:** Implementation of advanced recognition algorithms specifically configured for texts lacking standard dictionaries, including historical documents with archaic language forms and obsolete characters commonly encountered in heritage collections.
2. **Optical Music Recognition (OMR) integration:** Development of specialized recognition modules for printed musical notations, incorporating MusicXML export capabilities. This enhancement will enable the platform to process



historical musical manuscripts and printed scores, expanding its utility for musicological research and cultural heritage digitization projects.

Within GRAPHIA, the primary integration path for VTL outputs lies in the structured text and metadata produced through the transcription process. Future work will explore how TEI-encoded outputs can be transformed into RDF or other graph-compatible formats to support ingestion into the GRAPHIA SSH Knowledge Graph.

5. Conclusion

This deliverable presents the first **operational prototype** of the GRAPHIA Catalogue of Next-Generation SSH Instruments and Tools, developed within Task 3.1. Building on the initial inventory established in Milestone M2, the work carried out for D3.1 has consolidated a structured catalogue model and validated an implementation strategy based on integration with the SSH Open Marketplace (SSHOMP).

A key outcome of this work is the demonstration that the catalogue can be implemented without creating a new standalone platform. Instead, GRAPHIA has adopted an approach based on **reuse and interoperability**, leveraging an existing European SSH discovery infrastructure to describe, contextualise, and connect its instruments and tools. This choice reduces fragmentation within the research infrastructure landscape and ensures that the resources developed in GRAPHIA are visible and reusable within a broader SSH ecosystem.

The catalogue presented in this deliverable should therefore be understood as a **living resource**, rather than a fixed inventory. The list of instruments and tools included in this first version reflects the current contributions of Task 3.1 partners, but it is expected to evolve significantly during the remaining project period. New instruments and services may be added as they are developed or refined within the project, while existing entries can be progressively enriched with additional metadata, documentation, media resources, and links to related workflows, datasets, and use cases.

In addition to expanding the catalogue with new entries, future work will focus on **improving the semantic richness and contextualisation** of existing tools and



services. This may include refining metadata descriptions, strengthening relations between resources, and aligning the catalogue more closely with the technical and semantic requirements of the GRAPHIA SSH Knowledge Graph developed in WP2. Several entries in the catalogue already demonstrate concrete steps toward KG integration: CEC produces RDF output compatible with the OpenCitations Data Model; HYPERMNESIA's W7 model has been mapped to the GoTriple Data Model underlying the GRAPHIA data architecture; and SENNSE data can already be ingested via the Digilab API. These existing bridges provide a primary foundation. As workflows (Task 3.2) and use cases (Task 3.3) mature, further connections between instruments, workflows, and datasets will also be explored.

An important opportunity for the future development of the catalogue is represented by the **planned introduction of a new functionality** in the SSH Open Marketplace: **Collections**. This feature will allow related resources to be grouped and presented as **curated thematic sets** within the marketplace. Once this functionality becomes available (expected late 2026) the GRAPHIA catalogue of instruments and tools could be implemented as a dedicated **GRAPHIA collection** within SSHOMP. This would provide a coherent entry point for users interested in GRAPHIA resources while maintaining the benefits of integration with the broader SSHOMP ecosystem.

Such an approach would allow the GRAPHIA catalogue of next-gen I&T to remain both visible and contextualised, enabling users to discover GRAPHIA tools alongside other SSH resources while still preserving the identity and coherence of the project outputs. In the longer term, the catalogue of instruments and tools described in this deliverable could also evolve toward a broader **GRAPHIA catalogue of services**, linking instruments, workflows, datasets, and use cases developed throughout the project.

In conclusion, D3.1 establishes the foundations for a sustainable and interoperable catalogue strategy for GRAPHIA. By aligning the catalogue with the SSH Open Marketplace and contributing new entries and conceptual elements to the platform, the project strengthens the wider SSH infrastructure ecosystem while ensuring that its own outputs remain discoverable, reusable, and extendable beyond the project lifetime.



Appendix A.

The JSON dataset underpinning this catalogue, which serves as the machine-readable output for Deliverable D3.1, is generated by querying the SSH Open Marketplace (SSHOMP) API. This dataset includes the structured metadata for all GRAPHIA instruments and tools onboarded into the marketplace, adhering to the SSHOMP metadata guidelines. It captures details such as resource identity, responsible organizations, keywords, access URLs, and the relations established with other SSH resources. Due to the dynamic nature of the SSHOMP, the full, most up-to-date JSON export can be retrieved programmatically using the platform's public GET API endpoint, using “graphia” or “graphia project” keywords for search (not case-sensitive) ensuring the data reflects the current state of the catalogue.

The dataset can be found on Zenodo at the following link:
<https://zenodo.org/records/19482278>

Further information on the API endpoint of the SSH Open Marketplace can be retrieved at:

<https://marketplace.sshopencloud.eu/about/api-documentation>

The endpoint of the SSH Open Marketplace API is:

<https://marketplace-api.sshopencloud.eu/api>