A W7 structured metadata framework for reinforcing data provenances in 2D/3D digitization workflows

Anthony Pamart[1], Quentin Vogel[1], Eloi Gattet[1], Daouda Ngom[1], Laurent Bergerot[1]

[1] Modèles et simulations pour l'Architecture et le Patrimoine, UPR 2002 MAP CNRS, Marseille, France.

Keywords: Metadata, Paradata, Data Provenance, real-based modeling, workflows, image-based modeling, range-based modeling, photogrammetry, lasergrammetry, computational photography

Abstract

The documentation and provenance tracking of digitization workflows in Cultural Heritage (CH) face persistent challenges due to fragmented metadata practices and the absence of intuitive frameworks. This paper presents a W7-structured metadata framework, implemented through ANAMNESIS, an open-source platform designed to reinforce data provenance in 2D/3D digitization workflows. By systematically capturing the what, who, where, when, why, how, and which dimensions of digitization activities, the W7 model provides a structured, semantic approach to documenting critical technical and contextual details often lost in current practices. The platform operationalizes this framework with customizable schemas tailored to 2D/3D imaging workflows, ensuring traceability, reproducibility, and interoperability. Through real-case applications, the proposed approach not only preserves data provenance but also enables FAIR-by-design practices, empowering CH practitioners to contribute to a standardized, collaborative knowledge base.

1. Introduction

1.1 General Introduction

Digitization workflows have evolved alongside the technological advances since the early stages of Cultural Heritage (CH) applications. While the main stages—from capture to processing and analysis—remain unchanged, their underlying processes have grown increasingly complex. The methods used to create 2D/3D digital documentation also vary widely, shaped by practitioners' habits, tools, and experiences. This variability, coupled with the constant adaptation of protocols to CH artifacts, often results in a gap in reproducibility. Despite the historical and persistent weaknesses in data provenance and information traceability, experts' choices throughout the modeling process directly influence the quality of the digital output. If we consider real-based modeling as a form of sampling from the physical world, other scientific domains would never reuse a sample without comprehensive knowledge of its content (Plumejeaud-Perreau and et Al., 2019). Yet, this fundamental principle is frequently overlooked in digital sampling—even though the stakes remain equally high. Today, 2D and 3D data are increasingly shared, exchanged, and reused, driven by the rise of open science and FAIR principles. However, aligning data and workflows with these principles remains a tedious and challenging task. While metadata and paradata are well-known cornerstones for data scientists, they are frequently undervalued by data providers. This persistent gap between data providers (e.g. engineers, archaeologists) and users (e.g. curators, archivists) stems largely from the complexity of creating semantically rich documentation. Because of a steep learning curve, field experts often prioritize immediate operational tasks over long-term data usability—effectively sidelining metadata and paradata, despite their critical role in sustaining data lifecycle.

To bridge this gap, we introduce a practitioner-centric framework for metadata enrichment, paired with an open-source web application designed to lower the barrier to semantically rich documentation.

1.2 Research goal

Our approach is grounded in the W7 model, introduced by Ram and Liu (Ram and Liu, 2008, Ram and Liu, 2009, Liu and Ram, 2017) as a semiotic framework to enhance data provenance. The W7 model provides an intuitive and flexible semantic structure to organize metadata, effectively bridging the gap between data providers (e.g. CH practitioners) and data users (e.g. scholars, data scientists). Unlike complex metadata standards, the W7 model is accessible to all, as it relies on answering seven fundamental questions about interconnected concepts: Object (WHAT), Space (WHERE), Time (WHEN), Agent (WHO), Method (HOW), Instrument (WHICH) and Context (WHY) as shown in Fig. 1. While the W7 model's simplicity is its strength, it also presents a limitation: it is neither a formal metadata schema nor a high-level ontological framework. However, this very characteristic enables it to facilitate dialogue between low-level and high-level semantic enrichment. Since 2022, the W7 model has been applied in Cultural Heritage (CH) and Heritage Science (HS) practices, demonstrating its versatility and adaptability across various scientific domains (see section 2.1). Cultural Heritage Science (CHS)practices are not inherently FAIR-compliant, and adopting open-science principles requires a paradigm shift. Currently, documenting research protocols and digital activities is often perceived as an additional burden, leading to its avoidance or neglect. Without user-friendly tools that anticipate minimal semantic structuring, documentation is frequently sidelined. The FAIR principles, particularly the reusability pillar, introduce significant constraints for existing and future data, emphasizing the

need for enhanced provenance and traceability. Scientific communities, including CHS and SSH, have access to numerous semantic frameworks (ontologies, thesauri, controlled vocabularies, and metadata schemas). However, these frameworks are often disconnected from practitioners' needs and experiences. Mastering semantic integration requires expertise comparable to that needed for instrumental, computational, or analytical techniques. For reproducibility and repeatability, data lineage must be supported by workflow management, tracking, and provenance. The primary objective of this research is to strike a balance between the realities of scattered practices and the theoretical constraints of semantic structures, enabling fine-grained documentation of digitization activities.

Our hypothesis is that the paradigm shift toward FAIR-by-design approaches is twofold. First, it involves empowering practitioners by providing an interfacing tool that enhances accessibility and awareness of metadata management, built upon a versatile framework tailored to their needs. Second, it aims to use this tool as a semantic bridge, centralizing user-structured documentation to facilitate curation and enable alignment with broader semantic frameworks (e.g. metadata schemes, reference models, ontologies), thereby ensuring interoperability and long-term reusability.

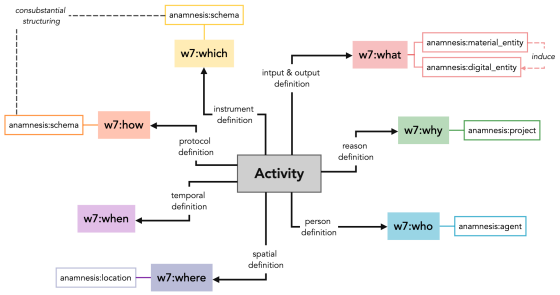

Figure 1. Diagram of W7 model describing CHS digitization activities.

2. State of the art

2.1 Related works

The MAP CNRS inherits decades of research work in ICT application to CHS documentation framework. Our research unit has been exploring the semantic enrichment of 3D modeling since 2000's (Blaise and Dudek, 2002) that one could consider today as the early stage of semantically enriched 3D digital model. Since then, the construction of Heritage Digital Twin (HDT) from a technical and methodological point of view has evolved by different means, with them the workflows to fill the knowledge gap between a static resource (text, 2D or 3D data) and correlated multidimensional data corpus (real based and parametric based 4D model enriched with semantic layers (Roussel et al., 2025)). The foundation of such semantically enriched models was achieved by the anticipation of two aspects ; the underlying semantic layers and reinforced data provenance of any modeling activity (Guillem et al., 2025). Our laboratory explored the semantic enrichment of 2D/3D models by two approaches, i) by spatial annotations (cf. Aioli platform) ii) by metadata and paradata framework. This work focuses on the second approach.

In addition to this innovative approach to convey geometric and semantic modeling methods, a reflexive approach regarding research protocol management, has been conducted. Memoria is an IS developed since 2007 specifically to contribute to a formal and sharable elicitation of research workflows (Dudek and Blaise, 2023). Activities are distributed in five goal-oriented groups dedicated to: data collection and acquisition [78 activities], data filtering and treatement [44 activities], data analysis [51 activities], added value procedural activities [56 activities], and finalisation [68 activities]. These 310 activities are associated with descriptors allowing a multi-level specification of the nature of a given activity (Dudek et al., 2022). The principles and experiences of Memoria remain valid and inspiring despite a granularity gap to facilitate users's adoption. Lately another complementary prototype has been developed; it took the form of an open-source web-application with an intuitive UI/UX aiming to induce an early metadata enrichment practice.

2.2 Data and workflow provenance in CHS

Data provenance in CHS domains is a recurring yet unresolved issue. One of the first requirements to maintain traceability is metadata, for which the community could choose between several standardized schemes (Greenberg, 2005), such as LIDO (Pitzalis et al., 2011), CARARE, or more recently EDM (Isaac and et Al., 2025), until today without a consensual adoption (Skinner, 2014). This first necessary step is followed by more complex framework (Doerr and et Al., 2014) leading to conceptual modeling to be fitted on top of real-based modeling (Carboni and et Al., 2016, Lauro and Lombardo, 2025). The theoretical reference model for CH application, to be named, the CIDOC-CRM has evolved with the practices over the time to encompass digital activities (CRM-Dig), conservation and restoration intervention (CRM-Rp) and move toward HDT Ontology (Theodoridou and et Al., 2025). Nowadays, it requires the accumulation of several semantic layers to overlay semantically enriched documentation of digitization-based studies (Kuroczyński et al., 2024). The trend is toward a standardization of CH practices, which are, in essence — singularity of CH object requires context tailored methods — and by experience unstandardized. The main difficulty is conciliating this multifaceted issue (i.e. FAIR approaches, data provenance, workflow traceability and scientific reproductibility (Massari et al., 2025)) and coping it with user-oriented methods. Indeed the orchestration of severals ontological level remains expert computer scientist skills not to be confused with Cultural and Heritage scientist (Belhajjame and et Al., 2015). Nowadays current practices implies a tedious data curation to reach FAIR compliance (Thomas and Brochu, 2022). Metadata are often consider as late enrichment method (e.g. at archival step only (Steiner and Koch, 2015)) while it should be a dynamic approach from the source to the output (Samuel and König-Ries, 2022) and beyond dissemination or storytelling (Barzaghi et al., 2025).

While Electronic Lab Notebooks (ELNs) are widely adopted in fields like microbiology, organic chemistry, and quantum physics for traceability and reproducibility (Mitchell and et Al., 2022), their use remains marginal

in Social Sciences and Humanities (SSH) and Cultural Heritage Science (CHS). Existing platforms, such as the cloud-based eLabFTW, are tailored to experimental sciences and fail to meet the unique documentation needs of CHS—requiring significant adaptation for direct application. Despite their potential for documenting research provenance and workflow lineage (Schröder and et Al., 2022), ELNs have yet to gain traction in domains where documentation practices differ fundamentally from those in applied sciences. An attempt has been made to transpose ELN for CH resource management (Schroer and Mudge, 2017), providing a tool dedicated to image-based documentation. While there are many solutions for CH data management, visualization and sharing, there is a lack of solutions to centralize metadata (Spettu et al., 2024). In this context, W7 approaches have already shown results to fill this gap in other domain (Arora et al., 2019) or in web-semantic (Orlandi et al., 2010). The W7 is an operational projection of the categories of Bunge's high-level ontology toward documentary use, making it compliant with other semantic layers.

3. ANAMNESIS Web-Platform

ANAMNESIS is a web application designed to manage metadata and paradata using W7-structured forms. Users are guided to describe their digital activities and resources by answering seven simple questions, which lead to the completion of a metadata form called *MEMoS* (see Fig. 2). The platform aims to document digitization activities, complementing descriptions of the physical object and its associated data with additional paradata.

The WHAT section describes the object of study (material entity) and its digital version (digital entity) as shown in Fig. 3. The WHERE and WHEN sections describe the spatial and temporal context of the activity. The WHO section details the agents and institutions involved. The WHY section provides contextual information, such as projects and funding sources. These first 5Ws are stabilized using Persistent Identifiers (PIDs) connected via APIs with other services (Geonames, ORCID, CORDIS, etc.) as illustrated in Fig.4. The last 2Ws, HOW and WHICH, then specify the techniques performed and the instruments used in digitization activities. These sections can be created and modified directly in the platform's graphical interface or imported using documented JSON schemas to allow the implementation or updating of technical schemes. The current templates already enable the documentation of workflows for the most commonly used imaging techniques in Cultural Heritage, such as photogrammetry, lasergrammetry, and computational imaging (RTI, Technical Photography), while documentation for other techniques (Infrared Thermography, Climate Monitoring, Sampling, etc.) is in the prototyping phase. Currently, most of the available schemas (±50%) are devoted to the instrumentation and data acquisition phase.

By organizing and sharing their metadata through the platform, users not only adhere to best practices on a cumulative and collaborative scale but also contribute to a centralized ecosystem—pooling activities, locations, stakeholders, and protocols. The platform further supports this process by enabling the management and sharing of custom templates and tools (e.g. prefilled forms for batch processing), fostering effortless documentation and collective enrichment. This creates a new paradigm for scientific exploration and serendipity. To support this, ANAMNESIS includes a shared collection system, enabling easy centralization, grouping, and dissemination of practices, entities, or reference resources—alongside any form of scientific communication, whether formal or informal. This article explains the construction of these schemes to encourage contributions from experienced users. User adoption is supported by graphical and wiki-based documentation, already available online (see link below). The documentation of the W7-ANAMNESIS schemes, including multilingual support, is managed by controlled vocabulary tools OpenTheso, which provide comprehensive definitions and stabilized URIs for interoperability purposes.

From a technical standpoint, the ANAMNESIS web service is open-source (GNU AGPLv3 license). The backend uses a SQLite database, which, despite its limitations, communicates efficiently with the frontend through API-based requests. The UI is developed using the Vue.js framework and supports multiple languages. The UX adopts a minimalist design (see Fig. 2) to ease metadata input, replacing cumbersome HTML forms with an efficient, user-centric interface. The deployment of the web application, currently hosted on the ESPADON distributed infrastructure, is based on containerization (Docker), anticipating multiple instances and scalability management. The platform can also be installed locally for use as a stand-alone application. All the resources are available following this link : https://hnurl.fr/ccw17d.

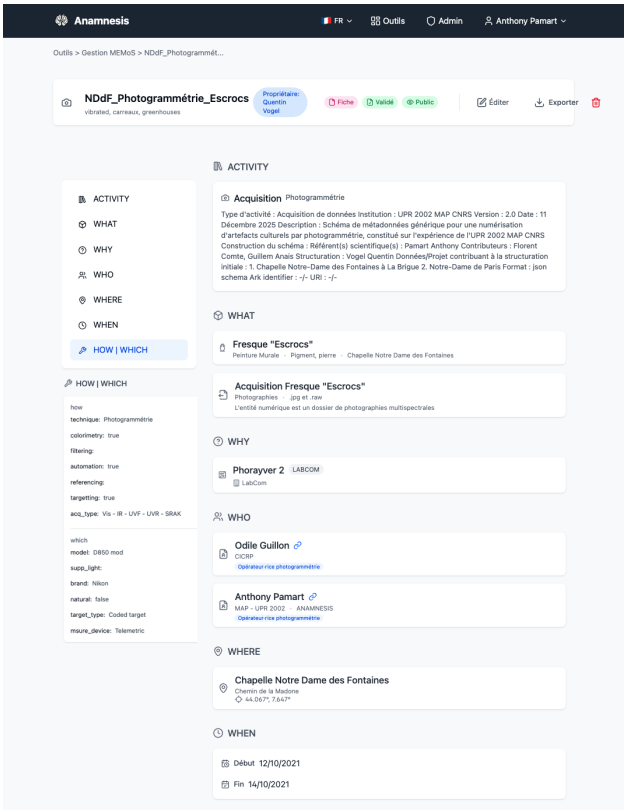

Figure 2. Screenshot of ANAMNESIS UI showing a filled MEMoS form.

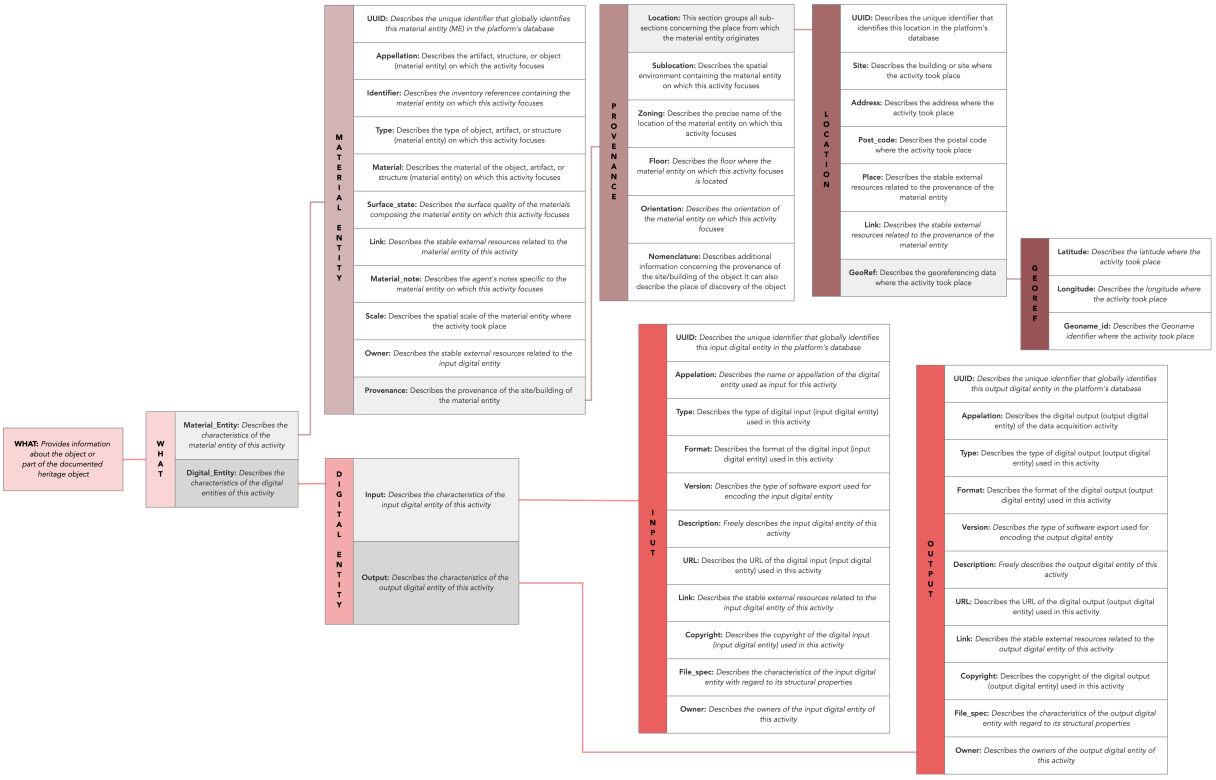

Figure 3. Detailed view of the W7 WHAT section, including all entry fields and their expected values (with definitions).

4. Application to CH Digitization workflow management and tracking

Using ANAMNESIS as documentation framework help practionners to structure their activities with flexible yet homogenous schemes. The already shemes implemented within the platform enable to reinforce data provenance of digitization workflow, from raw data toward the dissemination of enriched digital models. Past and current digitization practices are composed of instrumental, computationnal and analytical phases toward deliverable in multiple forms (e.g. scientific publication, dataset, exhibition, storytelling etc). ANAMNESIS enable to document this data stream with dedicating schemes for each category of activity in data acquisition, processing and analysis to complete the input/output operative chain. In this results and discussions section, the main activities and schemes are presented in subsection 4.1 and first outcomes are given in subection 4.2 from a pilot case studies and on-going collaboration with practionners of other imaging and measurment techniques.

4.1 Structuring and Linking Digitization Activities

Data Acquisition Schemes for Instrumental Phases:
A digitization practice begins with a data acquisition step, guided by established protocols and best practices for each technique, as well as shaped by user experience and expertise in defining the correct protocol and parameterization. During this instrumental phase, multiple choices are made, but many of these details are often omitted from technical metadata (e.g. photogrammetric baseline vs. EXIF). The technical scheme for real-based modeling enables documentation of critical aspects such as lighting, filtering, referencing, or pre-calibration setups. This information is typically undocumented and lost in current practices, despite its well-known impact on data quality, uncertainty, and reliability. W7 schemes enable precise documentation of how data have been collected and which instruments were used, and link this information to other provenance details (e.g. operators and affiliations, research context, time-span). Such paradata, though critical for long-term archiving, open-access deposits, and data papers, are often lost during the data lifecycle.

Data processing schemes for computationnal phases:
The activity category related to processing enables documentation of computational phases, taking as input the output of a data acquisition *MEMoS* sharing a material entity and created new digital entities. Linked to the prior step, these schemes are dedicated to specifying the outputs of technical processes and methods. The examples provided in the platform include photogrammetric and lasergrammetric processing schemes. The image-based 3D reconstruction workflow has been implemented from a report generated by a processing software. It structures the main parameters for Structure from Motion (image matching and pose estimation), spatial referencing, dense matching (depth maps, optimization, etc.), and specifies options for each generated output (point cloud, mesh, texture, DEM, orthophoto). The lasergrammetric processing scheme includes registration approaches and sampling methods. The main outcome of these processing schemes is to keep track of quality control and uncertainty aspects that could impact data (re)use. This step should be seen as an intermediate phase, leading to the next stage of the workflow.

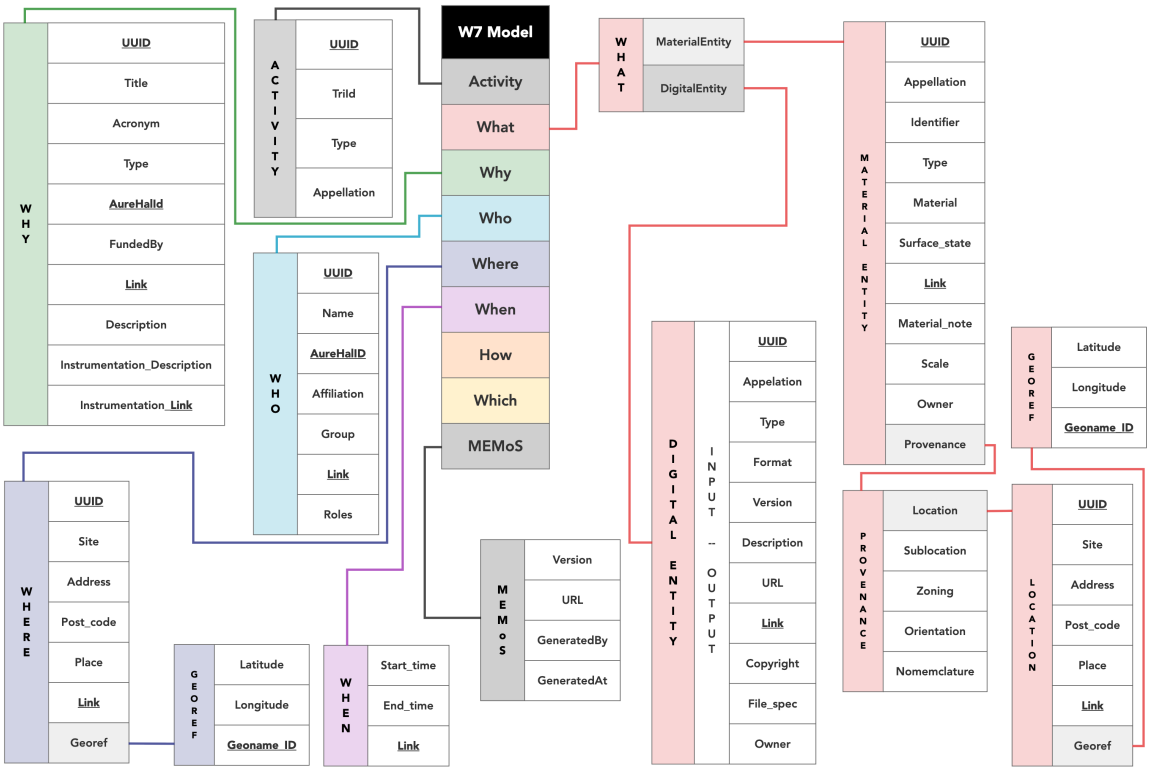

Figure 4. Overview of the core W5 sections, highlighting fields linked to **Persistent Identifiers (PIDs)**.

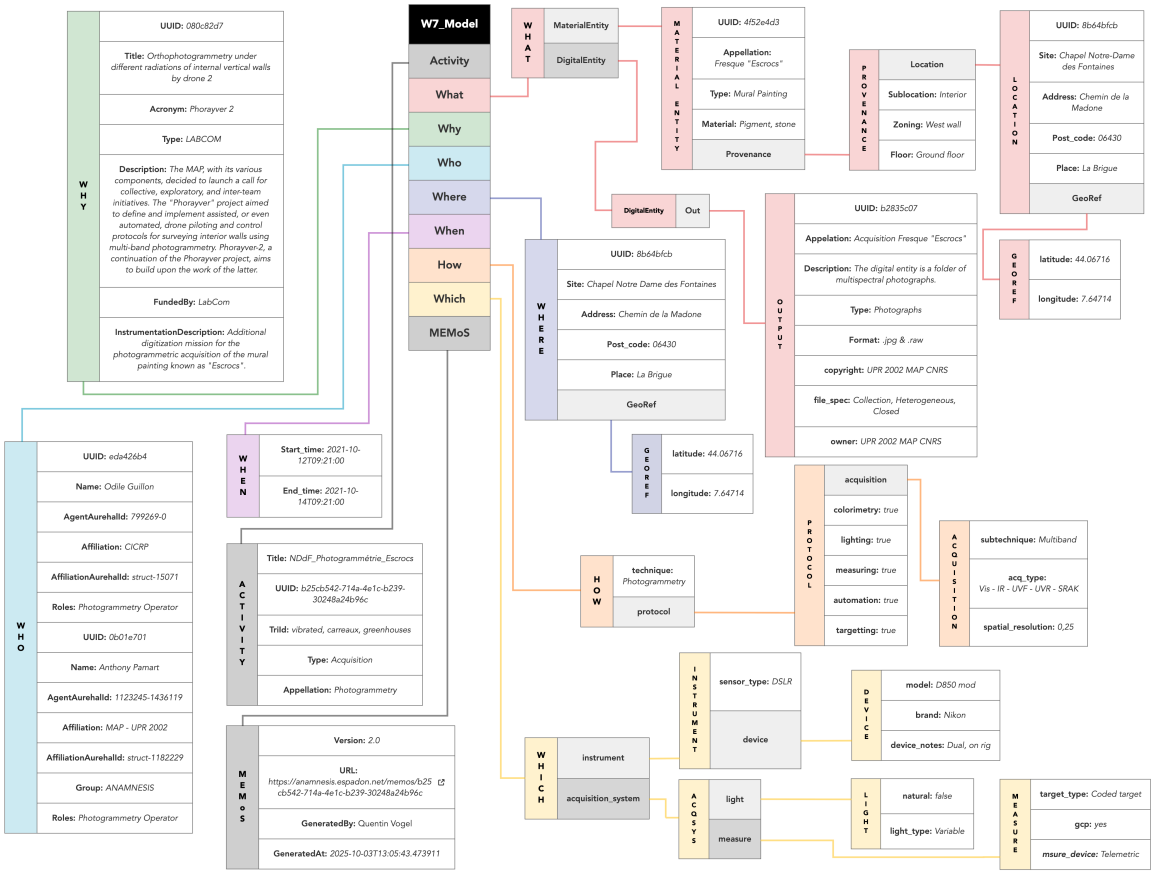

Figure 5. Example of a completed MEMoS using the photogrammetric acquisition scheme.

Data analysis schemes for analytical phases:
The data generated by the digitization pipeline can serve as the source for an analytical phase. Activities in this category include data fusion, enrichment, classification, segmentation, or multi-layer visualization. Unlike previous steps in the workflow, analysis activities are expertise-driven methods and are thus more challenging to categorize and structure into metadata schemes. An attempt has been made to formalize two approaches based on internal experience. Even tough the proposed scheme were curated manually, an exploratory approach involves extracting W7 schemes from scientific publications using fine-tuned Large Language Models (LLMs). The first example is the 2D to 3D spatial annotation process, based on experiments using the AIOLI collaborative platform (Roussel et al., 2025). This schema aims to transcribe the practice of 2D/3D enrichment by synthesizing semantic structure and features (e.g. number and type of layers, number and typology of annotations, descriptors, and use of controlled vocabulary). The second example is a data fusion approach derived from an experiment to build an enhancement feature from multisource data (Pamart et al., 2023). This schema can be reused for any data co-registration approach with advanced feature descriptions.

Other schemes for delivering phases:
Finally, a miscellaneous category has been implemented to include complementary steps or activities that do not fit conventional workflows. This category is more prospective but anticipates further use and advanced practices. The envisioned activities in this category are twofold. The first involves advanced data curation and interoperable mapping, aiming to project data into open and FAIR practices. A schema is currently under construction for depositing a massive data collection into an open repository (e.g. data warehouse or data lake), requiring strictly curated metadata. A similar approach is foreseen to foster data compliance for the connection to federated knowledge graphs. These schemes can be seen as structures synthesizing the data stream, compiling and organizing all activities composing a workflow and all semantic layers added in a data collection. The second use of this special category is the metadata consolidation of innovative tools and instruments, such as smart devices and IoT sensors, hybridizing instrumental, computational, and analytical processes simultaneously(De Luca et al., 2026).

4.2 Case studies and examples

In order to demonstrate the practical use of ANAMNESIS, this subsection will present a examplified case-study and an on-going work to implement new schemes with expert validation.

Using Ananmesis to document multimodal digitization practices:
A small chapel in south of east of France has been documented among several digitization campaings during the past decade creating a multisource digital model aiming to support further investigation for conservation and restoration purposes. For the instrumentation, the several surveys data performed by range-based and image-based techniques were documented using W7 schemes for photogrammetry and lasergrammetry (see Fig.5). It provides an historic overview of the digitization campaign revealing the operations of many experts applying imaging techniques on different part of the chapel belonging to different research contexts. Such provenance information with enriched paradata are usually difficult to retrieve if not lost in current digitization practices. Each source was processed using conventional IBM or RBM and documented with schemes tailored for data processing providing all traceability information from a practical point of view to retrieve co-registration approach and spatial overlapping between the source. Finally, these outputs were compiled into a single digital model enriched with a fusion index documented with an ad-hoc analysis scheme. This data fusion activity summarized the enhancement features in synthetic form providing fast and efficient access of analytical procedure. Once implemented in the Misc category, a additional archiving schema could be use retrospectively to expose the structure of the data set available in a open-data repository. This complete documentation framework compose a collection of MEMoS accessible within ANAMNESIS using a UUID centralizing all the ressources available on this case-study.

Using ANAMNESIS as pivotal interface to structure W7 scheme for CH activities:
The beta testing phase among partners positionned ANAMNESIS not only as an efficient user-oriented metadata management system but also pivotal tool to induce metadata awareness among various CHS practionners. The initial schemes concerning common shared practices like photogrammetry and lasergrammetry presented to beta-testers picked their intererest and curiosity to implement new schemes for their own techniques. Indeed, the current limitation is that a new activity or technique requires a schema for being documented within ANAMNESIS. Two approaches are currently explored to bridge this gap. The first one, is a dialogue with expert to help them formalizing their schemes for which the W7 intuitivity has been proven really helpful. Next step will focus on Compentency Questions to derive from ad-hoc dialogue to procedural and stabilized routine for new scheme implementation. In the administrator panel, the platform's graphical interface the user to construct and organize themselves the new schemes of activity. Interestingly we discover that W7 is flexible enough to adapt at differents levels ; not only at activities and techniques scales but also for projects, institutions, partnerships. The second one is a generalization approach targeting generic schemes even tough reaching such high-level of abstraction is challenging.

In order to ease the management of metadata schemes, we rely on Opentheso thesaurus manager to source the explanation of terminologies with a precise and expert qualification of the latter from the outset (see link given above in section 3). Indeed, although the initial schemes reflect specific in-house practices, efforts are now underway to engage partner institutions and colleagues—who explore similar subjects using diverse techniques—in order to broaden the framework's applicability and relevance. For example, the Sampling and Climate monitoring schemes are currently being validated by the expert, as part of a joint effort with a dedicated working

group (Cormier et al., 2024) of the Espadon equipEX. Starting from undocumented practice and gray literature specific to this expertise field, the user-friendlyness of ANAMNESIS and W7 framework simplified the structuration of a metadata schema from scratch. From this experience a user documentation package has been initiated (see link given above in section 3) aiming to autonomize practionners schemes creation and implementation.

4.3 Future works

From a technical standpoint, ANAMNESIS—though fully operational as a web application—remains a first-release research prototype. Minor bugs and optimizations require attention, and planned features (e.g. a metadata inspector, import tool to auto-fill existing information like EXIF data) are already prioritized in the development roadmap. From a methodological perspective, while the W7 model offers versatility, it remains underutilized CHS. Future work includes standardizing W7 as a stabilized schema using existing frameworks (e.g. schema.org, Ontoportal) to address interoperability challenges that will meet the previous point with on-going API development. Although recent studies (Guillem et al., 2025) demonstrate W7's compliance with ontological layers (e.g. CIDOC-CRM, PROV-O) for data provenance, mapping with other metadata schemes (GoTriple, EDM, Dublin Core, DataCite) is a priority. Currently, ANAMNESIS+W7 relies on bottom-up, expert-driven creation and validation. To complement this, a top-down approach—using AI to extract metadata, schemas, and workflows from reports or articles—will be explored. Collaborations with the GRAPHIA project will leverage fine-tuned models, RAG (Retrieval-Augmented Generation), and chatbots to document techniques lacking dedicated schemas (e.g. point-based, spectral, or hyperspectral imaging in HS) or other activities in SSH/Digital Humanities. ANAMNESIS is set to be integrated into the E-RIHS Catalogue of Services as an innovative DIGILAB tool.

5. Conclusion

In this paper, we demonstrated how the W7 semantic framework can effectively document and preserve digitization workflows, initially applied to 2D/3D imaging but extensible to broader CHS practices. We introduced ANAMNESIS, an open-source, user-centered platform already operational for CH practitioners, offering non-rigid but structured schemas to describe common 2D/3D digitization activities. The platform stabilizes workflow documentation by centralizing and linking multiple PID layers (UUIDs, URIs, ARKs, and DOIs), ensuring traceability and interoperability. To streamline metadata and paradata documentation, ANAMNESIS provides customizable templates for data acquisition, processing, and analysis stages. This work contributes to promoting the W7 framework, with real-case CH studies illustrating its practical benefits toward reinforced data provevance workflows. We encourage experts to enrich existing schemas or create new ones based on their domain-specific needs. Ultimately, this approach aims to structure and reveal shared activities, fostering FAIR-by-design in Cultural Heritage Sciences and beyond, with potential for cross-domain applications.

Acknowledgements

The ANAMNESIS platform benefited from State aid managed by the Agence Nationale de la Recherche (French National Research Agency) and supported by the Paris Seine Graduate School Humanities, Creation, Heritage, Investissement d'Avenir ANR-17-EURE-0021 – Fondation for cultural heritage sciences. This project has been supported by the EquipEx+ ESPADON bearing the reference ANR-21-ESRE-0050 for the deployment and the hosting of the platform. Additional funding was provided by the CNRS through the MITI interdisciplinary programs (TEATIME) for source code improvement and optimization. This work is supported in the context of the EU-funded INFRATECH GRAPHIA project funded by the European Union Grant agreement ID: 101188018, as part of WP3 task 3.2 for the cataloguing of Innovative Workflows for Data Management and Connection to the SSH-KG.

Roles and contributions

Contributors: Anthony Pamart (AP), Quentin Vogel (QV), Eloi Gattet (EG), Daouda Ngom (DN), Laurent Bergerot (LB). CRediT (https://credit.niso.org/): Conceptualization : AP; Data curation : AP, QV; Funding acquisition : AP; Investigation : AP; Methodology : AP, QV; Project administration : AP; Resources : AP, QV, EG, DN, LB; Software : EG, DN, LB; Supervision : AP, QV; Validation : AP, QV, EG, DN; Visualization : AP, QV; Writing – original draft : AP, QV; Writing – review & editing : AP, QV.

References

Arora, S., Balhara, D. R., Sapra, D. P., 2019. Data Provenance Techniques and Semantics Using W7 Model-A Review. 10(3), 66–70. https://www.ijarcs.info/index.php/Ijarcs/article/view/6418.

Barzaghi, S., Colitti, S., Moretti, A., Renda, G., 2025. From Metadata to Storytelling: A Framework For 3D Cultural Heritage Visualization on RDF Data.

Belhajjame, K., et Al., 2015. Using a Suite of Ontologies for Preserving Workflow-Centric Research Objects. *Journal of Web Semantics*, 32, 16–42.

Blaise, J.-Y., Dudek, I., 2002. 3D models as visual interfaces in the documentation of the architectural heritage: experiment on the defensive system of Cracow. IASTED (ed.), *Visualisation, Imaging and Image Processing Proceedings of 2nd IASTED International Conference*, ACTA Press, Spain, 746–751.

Carboni, N., et Al., 2016. Data Provenance in Photogrammetry Through Documentation Protocols. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, III-5, 57–64.

Cormier, A., Roqui, D., Bourgès, A., Grozavu, N., Surma, F., Guillon, O., Labouré, M., 2024. Correlating climate change data with the behaviour of cultural heritage materials. *Colloque Climat et Impact 2024*, 12.

De Luca, L., Comte, F., Pamart, A., 2026. Dür.air: reconciling acquisition and interpretation in cultural heritage field documentation. *In Press - 3D ARCH 2026 Ancona*.

Doerr, M., et Al., 2014. A Framework for Maintaining Provenance Information of Cultural Heritage 3D-models.

Dudek, I., Blaise, J.-Y., 2023. Research workflows, paradata, and information visualisation: feedback on an exploratory integration of issues and practices - MEMORIA IS. *Peer Community In Archaeology*. https://shs.hal.science/halshs-04091200.

Dudek, I., Blaise, J.-Y., Rabefandroana, M., Bénistant, P., Renaudin, N., 2022. MEMORIA - nomenclature of activities. Technical report. https://shs.hal.science/halshs-03744846.

Greenberg, J., 2005. Understanding Metadata and Metadata Schemes. *Cataloging & Classification Quarterly*, 40(3-4), 17–36.

Guillem, A., Abergel, V., Roussel, R., Comte, F., Pamart, A., De Luca, L., 2025. A Bridging the Provenance Knowledge Gap between 3D Digitization and Semantic Interpretation.

Isaac, A., et Al., 2025. Making the europeana data model a better fit for documentation of 3D objects. M. Ioannides, D. Baker, A. Agapiou, P. Siegkas (eds), *3D Research Challenges in Cultural Heritage V: Paradata, Metadata and Data in Digitisation*, Springer Nature Switzerland, Cham, 63–74.

Kuroczyński, P., Bajena, I. P., Cazzaro, I., 2024. The Scientific Reference Model—A Methodological Approach in the Hypothetical 3D Reconstruction of Art and Architecture. 7(10), 5446–5461. https://www.mdpi.com/2571-9408/7/10/257.

Lauro, V., Lombardo, V., 2025. Towards a Structured and Scalable Metadata Framework for Photogrammetric Documentation: Ontological Traceability and Practical Implementation. *Proceedings of the 30th International Conference on 3D Web Technology*, ACM, 1–7.

Liu, J., Ram, S., 2017. Improving the Domain Independence of Data Provenance Ontologies: A Demonstration Using Conceptual Graphs and the W7 Model. *Journal of Database Management*, 28(1), 43–62.

Massari, A., Peroni, S., Tomasi, F., Heibi, I., 2025. Representing Provenance and Track Changes of Cultural Heritage Metadata in RDF: A Survey of Existing Approaches. *Digital Scholarship in the Humanities*, fqaf076.

Mitchell, S. N., et Al., 2022. FAIR data pipeline: provenance-driven data management for traceable scientific workflows. *Philosophical Transactions of the Royal Society A*, 380(2233), 20210300.

Orlandi, F., Passant, A., Champin, P.-A., 2010. Semantic Representation of Provenance in Wikipedia. *CEUR*, CEUR, 670, Shanghai, China.

Pamart, A., Abergel, V., de Luca, L., Veron, P., 2023. Toward a Data Fusion Index for the Assessment and Enhancement of 3D Multimodal Reconstruction of Built Cultural Heritage. *Remote Sensing*, 15(9).

Pitzalis, D., Niccolucci, F., Cord, M., 2011. Using LIDO to handle 3D cultural heritage documentation data provenance. *9th International Workshop on Content-Based Multimedia Indexing (CBMI 2011)*, 37–42.

Plumejeaud-Perreau, Q., et Al., 2019. Towards Better Traceability of Field Sampling Data. *Computers & Geosciences*, 129, 82–91.

Ram, S., Liu, J., 2008. A Semiotics Framework for Analyzing Data Provenance Research. *Journal of Computing Science and Engineering*, 2(3), 221–248.

Ram, S., Liu, J., 2009. A new perspective on semantics of data provenance. *Proceedings of the First International Conference on Semantic Web in Provenance Management - Volume 526*, SWPM'09, CEUR-WS.org, Aachen, DEU, 35–40.

Roussel, R., De Luca, L., Guillem, A., Comte, F., 2025. A cathedral of spatialised annotations portraying the multidisciplinary study of Notre Dame de Paris. S. Campana, D. Ferdani, H. Graf, G. Guidi, Z. Hegarty, S. Pescarin, F. Remondino (eds), *Digital Heritage*, The Eurographics Association.

Samuel, S., König-Ries, B., 2022. End-to-End Provenance Representation for the Understandability and Reproducibility of Scientific Experiments Using a Semantic Approach. *Journal of Biomedical Semantics*, 13, 1.

Schroer, C., Mudge, M., 2017. A Context Metadata Collection and Management Tool for Computational Photography Projects. 2017(1), 99–104. http://www.ingentaconnect.com/content/10.2352/issn.2168-3204.2017.1.0.99.

Schröder, M., et Al., 2022. Structure-Based Knowledge Acquisition from Electronic Lab Notebooks for Research Data Provenance Documentation. 13(1), 4. https://doi.org/10.1186/s13326-021-00257-x.

Skinner, J., 2014. Metadata in Archival and Cultural Heritage Settings: A Review of the Literature. *Journal of Library Metadata*, 14(1), 52–68.

Spettu, F., Achille, C., Fassi, F., 2024. State-of-the-Art Web Platforms for the Management and Sharing of Data: Applications, Uses, and Potentialities. *Heritage*, 7(11), 6008–6035.

Steiner, E., Koch, C., 2015. A Digital Archive of Cultural Heritage Objects: Standardized Metadata and Annotation Categories. 20, 255–260.

Theodoridou, M., et Al., 2025. Towards the Definition of the Heritage Digital Twin Ontology for the European Collaborative Cloud for Cultural Heritage. S. Campana, D. Ferdani, H. Graf, G. Guidi, Z. Hegarty, S. Pescarin, F. Remondino (eds), *Digital Heritage*, The Eurographics Association.

Thomas, S. A., Brochu, F., 2022. Curation at the Point of Measurement and Traceability of Measurement Workflows. *Measurement: Sensors*, 23, 100399.