

## BEYOND MONITORING. REIMAGINING DHELO AS A LINKED OPEN DATA INFRASTRUCTURE FOR CULTURAL HERITAGE RESEARCH

### 1. INTRODUCTION

This article continues and expands on a line of research initiated within the H2IOSC project. As part of Work Package 2 (WP2 - Landscaping and Building Communities) the team at the CNR Institute of Heritage Science (ISPC) – representing the Italian node of the E-RIHS infrastructure – conducted a wide-ranging survey of the digital research landscape. This investigation focused on resources, practices, and communities operating across the domains of Cultural Heritage (CH), Heritage Science (HS), and Digital Archaeology (DA) (see CARVALE, MOSCATI, ROSSI, in this special section).

In earlier works the motivations behind the creation of a disciplinary observatory were presented (CARVALE *et al.* 2024; LUZIETTI *et al.* 2024; MANCUSO, D'EREDITÀ 2024, in press), along with the development of two digital resources designed to support this objective: the DHeLO database (Digital Heritage Landscaping Platform – <https://dhelo.cnr.it/>), intended for the cataloguing of digital products, tools, and research projects, and the BiDiAr bibliographic collection (<https://bidiar.cnr.it>). While this initial phase of the project helped to identify several key challenges affecting the visibility, discoverability, and reuse of digital data in the CH, HS, and DA domains, the present article focuses on the outcomes of a second phase of work, centred exclusively on DHeLO. During the last year, the platform has been rethought and redeveloped in light of a broader reflection on its role and potential, opening new possibilities for its evolution beyond its initial scope.

### 2. INSIGHTS FROM THE LANDSCAPING ACTIVITIES

The evolution of DHeLO has been shaped by a series of landscaping activities conducted within WP2, including targeted questionnaires and interviews addressed to members of the DH, DA and HS communities. These activities, whose results have been recently unfolded (MANCUSO *et al.* in press), have provided valuable insights into current expectations and needs surrounding the production and management of digital data, particularly in relation to the FAIR principles and the broader ambitions of Open Science.

The need for landscaping activities in order to take the reference community into account during the design phases of an infrastructure is now well established in the scientific literature (ROSS *et al.* 2024), implying that infrastructures should not be viewed merely as technical structures linking

information and technologies (BENARDOU *et al.* 2017), but as systems that become progressively meaningful through their use and users, by connecting, networking, managing, and linking otherwise independent elements (PAWLICKA-DEGER 2021). In this context, understanding research communities goes beyond compiling inventories of existing resources or cataloguing available technologies, but requires a sustained and methodologically grounded effort to observe, analyse, and interpret the habits, expectations, and working conditions of the individuals and institutions involved in the production and use of digital content (MARTTILA, BOTERO 2021).

This analysis must also consider the normative and operational frameworks that shape their practices, including how value is assigned to digital resources, what constitutes trustworthiness or interoperability, and which barriers hinder data reuse and sharing. Within the H2IOSC WP2 framework, these conceptual reflections favoured a mixed-method approach, in which quantitative tools, such as structured surveys, are integrated with qualitative methods, including interviews and open-ended consultations, to document both explicit patterns and less visible dynamics that shape digital practices by the actors of the research community (LUZIETTI *et al.* 2024).

Among the various outputs of the landscaping activities, the interviews proved particularly valuable in informing the restructuring of DHeLO. Designed to complement the findings of the initial questionnaire, a series of semi-structured interviews was conducted with a selected group of scholars from the SH6 disciplinary sector, with the goal to gain deeper insight into how members of the academic community engage with digital data and to explore their expectations, needs, and concerns, regarding data management and sharing. Although limited in number, the sample of 25 interviewees was selected to reflect a broad range of academic backgrounds and professional experiences. Particular attention was paid to ensuring a balanced representation across career stages, following the same logic adopted for the questionnaire. The group consisted of 7 students, 8 early-career researchers, 5 mid-career scholars, and 5 senior academics.

Not surprisingly, one of the most critical issues identified when discussing community needs in the context of planning a data infrastructure is the semantic ambiguity surrounding the very concept of ‘data’, which is interpreted in multiple and sometimes conflicting ways (HUGGET 2020, 59). During the landscaping phase, and especially during the interviews, it became apparent that the term is used to denote a wide variety of materials, ranging from basic lists of artefacts and raw excavation records to remote sensing outputs and unprocessed 3D scans. This diversity of interpretations reflects the heterogeneity of research practices within the field and points to a more fundamental question: how can digital infrastructures effectively accommodate such a wide range of data types, each with its own contextual nuances and methodological implications?

This conceptual ambiguity also has practical repercussions, particularly when it comes to data sharing and open access practices (MARWICK, BIRCH 2018, 127-129). For many of the interviewed, the publication of written outputs in open access is considered sufficient to fulfil the requirement of data sharing. In this view, the structured presentation of research findings in a scholarly article is seen as inherently linked to the dissemination of the underlying data. However, this perspective coexists with another, more data-centric approach that advocates for the explicit publication of structured, reusable datasets as independent research products.

The coexistence of these perspectives suggests a tension between legitimate traditional forms of scholarly communication and emerging paradigms of data openness and reusability. A nuanced scenario that also reflects a cultural tension between written knowledge and digital content. Still, written publications, both print and digital, seems to remain the primary vehicles for scholarly recognition and dissemination, often perceived by the interviewed as more reliable and enduring than digital datasets (SCHMIDT *et al.* 2016). Even when digital data are made accessible, they are frequently prepared and perceived as interactive complements to traditional research outputs, rather than as standalone objects of knowledge with their own life cycle and potential for reuse. This perception not only shapes data management practices but also influences the role attributed to digital infrastructures, which are often regarded more as functional supports than as fully integrated ecosystems within the scholarly routine.

### 3. DHeLO. DIGITAL HERITAGE LANDSCAPING PLATFORM

These and other reflections that emerged from the landscaping activities have made it increasingly clear that DHeLO could play a more active role within the broader research infrastructure landscape. As discussed in earlier contributions, the platform was originally conceived within the H2IOSC project as a digital observatory for the CH and HS domains, aimed at collecting and indexing metadata on research products, tools, and projects in order to better understand the research community through its outputs (MANCUSO, D'EREDITÀ 2024, in press). One idea that gradually emerged during the development phase of the platform was the potential value of a system capable of aggregating and indexing existing datasets, interactive resources, and research projects. It became evident that, rather than functioning solely as a virtual observatory operating behind the scenes, the platform could evolve into a tool designed to serve the community through an active use. The envisioned service would not require contributors to conform to a specific data model, but would instead focus on collecting and exposing digital contents already available on the web and making them accessible through a unified, searchable interface; thus, responding to a gap that seems to be present in the current infrastructure ecosystem.

While several initiatives exist that allow or require the deposit of data, few seems to be designed to include or describe external web-based resources such as stand-alone WebGIS platforms or autonomous databases. As a result, a substantial portion of valuable digital content remains difficult to find and contextualise across domains. Seen from this perspective, the initial data collection efforts carried out for DHeLO provided a valuable foundation upon which to build. The metadata gathered in the first phase of the project offered a structured point of departure for designing a more open, connected, and semantically rich infrastructure.

### 3.1 From relational database to Linked Open Data: a change of platform

The transition from a relational database to a more flexible, community-oriented infrastructure represented a key turning point in the evolution of DHeLO (MANCUSO, D'EREDITÀ 2024). Instead of opting for a complete redesign, the platform underwent a gradual transformation aimed at aligning it with the principles of Linked Open Data (LOD). This approach facilitated the mapping and contextualisation of diverse resources, often accessible online with minimal or generic metadata or described only within scholarly literature. By restructuring the platform around LOD-compatible logics, it became possible to capture and represent these resources more efficiently, while maintaining a lightweight and adaptable aggregation model.

This conceptual shift also led to significant changes in the technical infrastructure underpinning DHeLO. Initially, the platform was built on

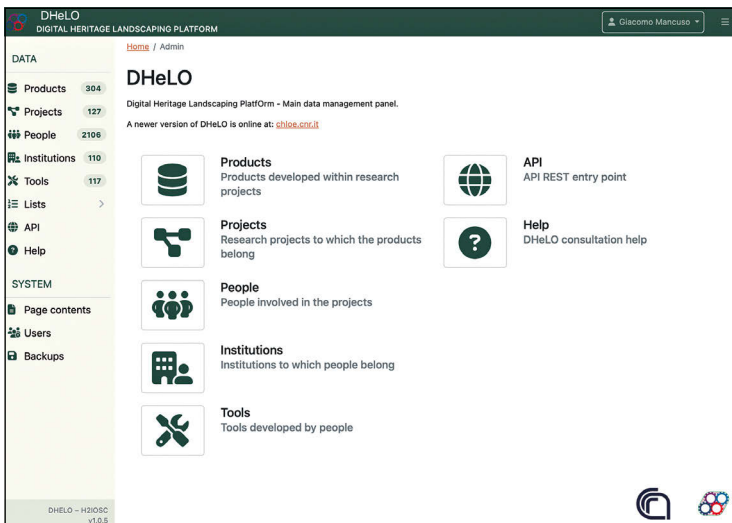


Fig. 1 – Front-end of the previous version of DHeLO, currently offline.

a MySQL relational database combined with a custom HTML interface, a configuration that has proven useful in the project’s early stages, especially for rapid prototyping, internal usage, and data visualisation (Fig. 1). However, as the platform expanded, several limitations became apparent, particularly in terms of scalability and long-term maintenance, as the system required ongoing technical support from IT specialists.

To address these challenges, the project team opted for a different architecture. After evaluating various alternatives, the choice fell on Omeka S, an open-source web publishing platform (<https://omeka.org/s/>). Designed to accommodate multi-site environments and support semantic interoperability, Omeka S functions as a content management system (CMS), facilitating structured data management, while also allowing for visual customisation through templates and CSS (Fig. 2). At its core, Omeka S provides powerful yet user-friendly backend for managing structured metadata, as well as a

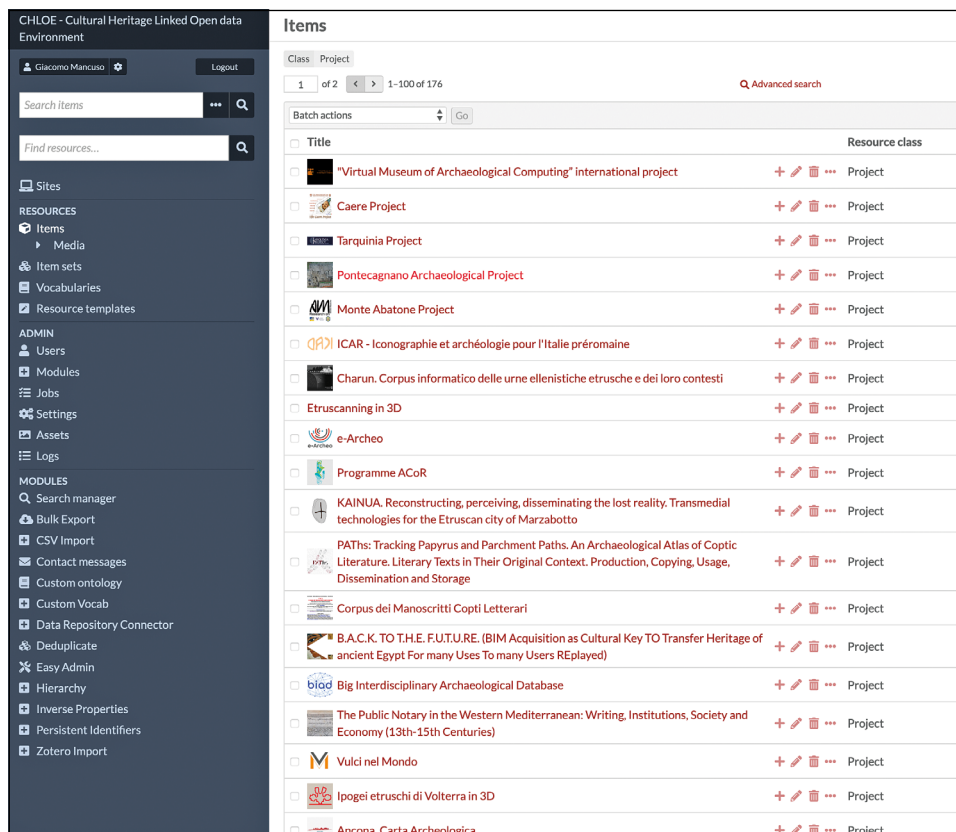


Fig. 2 – Back-end of the new version of DHeLO with Omeka S (<https://dhelo.cnr.it>).

flexible publishing layer for creating public-facing sites that present digital collections in curated and accessible ways. Notably, the platform can be managed effectively by users with no specific technical background, making it especially suitable for collaborative and multidisciplinary environments. While commonly adopted in the digital humanities and museum sectors, its use in DA project is still limited, though steadily increasing (RUEFF 2024, with a selection of case studies).

A key advantage that Omeka S provided to DHeLO lies in its modular architecture, which significantly expands the platform's capabilities, through an ecosystem of plugins and extensions that allow to integrate advanced functionalities without compromising the system's core structure. Among these, the Collecting module played a particularly important role, as it allows users to submit metadata for content that can then be ingested into the system. This capacity to involve the community directly in content curation supports a participatory infrastructure model, where users are not just passive consumers but active contributors to the platform's growth and sustainability (JIROTKA *et al.* 2013); an approach that aligns well with the idea of a community-oriented infrastructure, where data contributions are validated and curated with minimal oversight by an editorial board.

Another critical factor in choosing Omeka S was its conceptual and institutional alignment with Zotero, a well-known reference management tool ([www.zotero.org/](http://www.zotero.org/)); being both platforms developed within the same research centre. This synergy proved particularly valuable, given the conceptual relationship between DHeLO and BiDiAr, a bibliographic database initially created with Zotero and now being integrated into Omeka S. The two infrastructures will interact within the CHLOE framework (Cultural Heritage Linked Open-Data Environment - <https://chloe.cnr.it>): this integration is conceived as a Omeka S shared ecosystem capable of supporting both platforms over time, while also reflecting the central role that the scholarly community continues to attribute to bibliographic resources (Fig. 3).

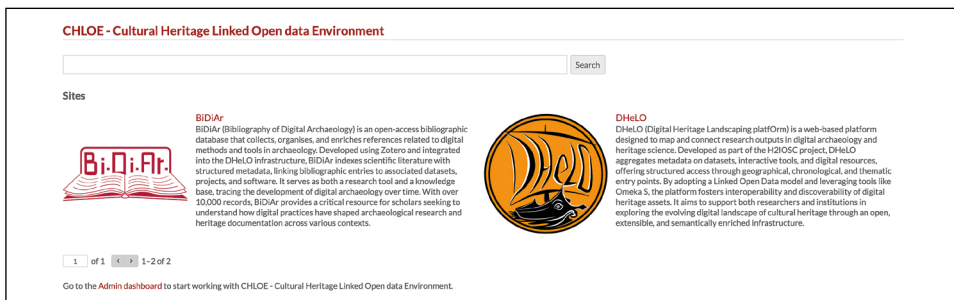


Fig. 3 – Homepage of the CHLOE network site with Omeka S (<https://chloe.cnr.it>).

Furthermore, Omeka S offers native support for controlled vocabularies and ontology-based metadata management. The platform is inherently compatible with RDF-compliant ontologies and can incorporate external vocabularies defined in SKOS or OWL. This capability enables alignment with well-established semantic standards, including Dublin Core, Friend of a Friend, and schema.org. The adoption of Omeka S also benefitted from the contribution of the H-Setis platform, developed within the project to collect and share semantic resources in the CH and HS domains (SCARPA, VALENTE 2024), that has been instrumental in establishing a conceptual and technical framework for DHeLO that is compliant with community standards.

### 3.2 From relational database to Linked Open Data: rethinking the data model

The adoption of a LOD infrastructure required a partial reorganisation of the DHeLO data model (Fig. 4). In its first configuration, the system was structured around six interconnected tables: persons, institutions, tools, research products, product types, and projects (MANCUSO, D'EREDITÀ 2024, 525-528). These tables served to document the various entities and their relationships within the broader context of cultural heritage research. A revision of these existing entities revealed significant overlap with classes defined in two widely used ontological frameworks: FOAF (Friend of a Friend) and the DC Terms (Dublin Core Metadata Terms). These vocabularies, although relatively lightweight, offer a solid and widely compatible foundation for describing individuals, organisations, research outputs, and related entities. Consequently, the core tables were mapped to the following semantic classes: *foaf:Person* for individual researchers, *foaf:Organization* for research institutions, *dctype:Software* and *dctype:Dataset* for digital tools and data products, respectively, and *foaf:Project* for research initiatives.

An important improvement introduced during this restructuring was the definition of *dctype:InteractiveResource* as a distinct class. This new category was designed to represent digital tools that act as interfaces for accessing data,

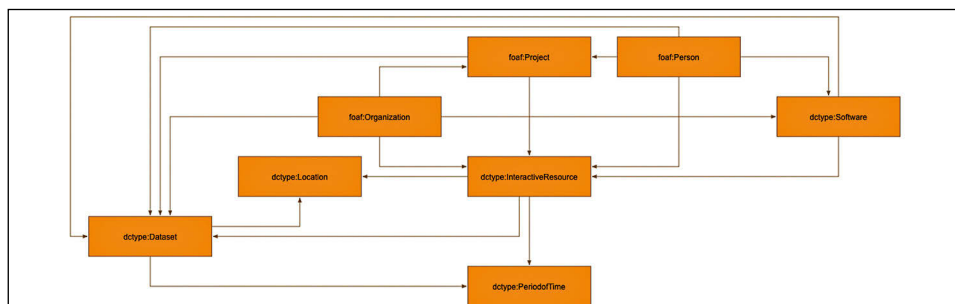


Fig. 4 – Conceptual scheme of the relations among classes in the new DHeLO data model.

such as visualisation platforms, WebGIS applications, or multimedia exploration tools. In the previous version of the platform, these interactive resources had been included in the tools table, a choice that somewhat diminished their intellectual specificity. Furthermore, this classification made it challenging to clearly define their relationships with other entities. By designating *InteractiveResource* as a separate class, it became possible to better represent

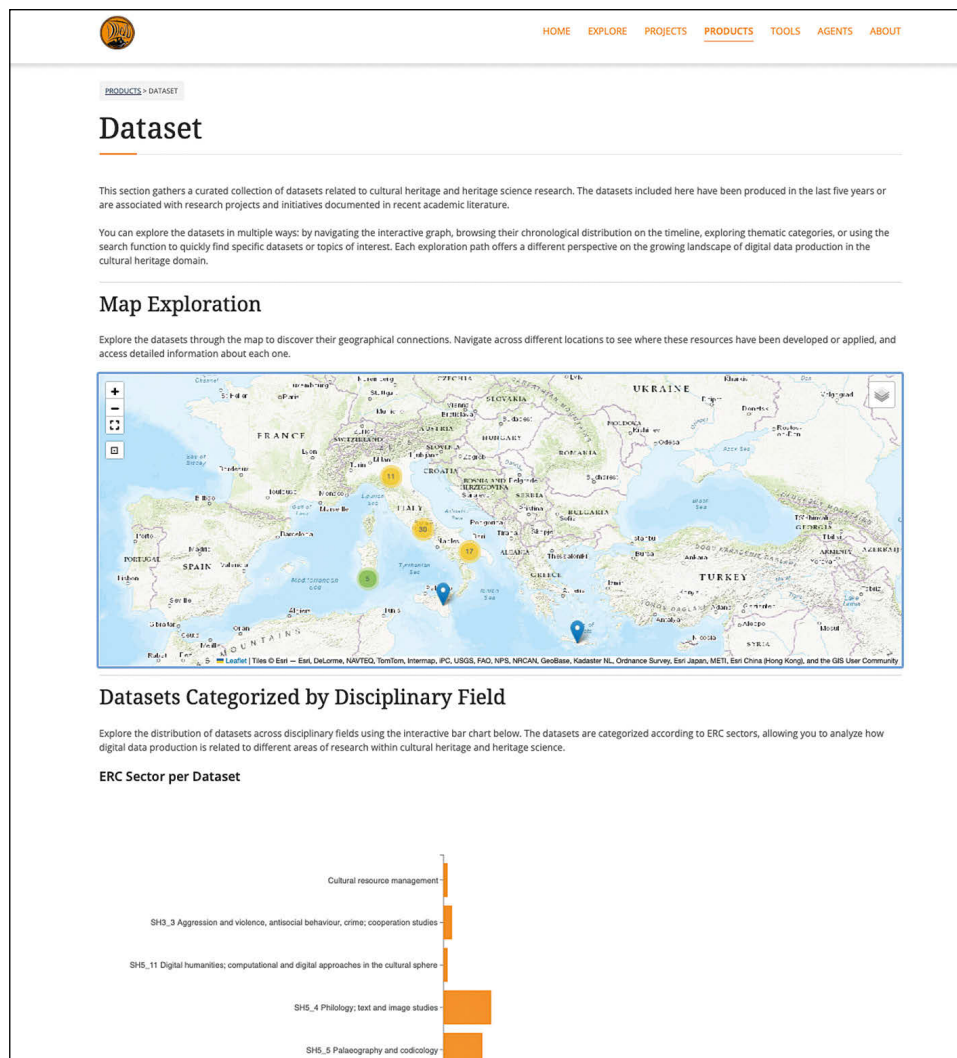


Fig. 5 – Starting page for the datasets exploration. In evidence the interactive map and graph for geographical and thematic exploration.

their dual function as both access tools and scholarly outputs (Fig. 5). This adjustment also allowed these resources to benefit from the same thematic, geographic, and chronological indexing applied to datasets, greatly improving their visibility and interoperability within the platform.

Regarding the thematic classification, the system maintained continuity with the previous version, adhering to the taxonomy developed by the journal «*Archeologia e Calcolatori*», providing a well-established framework for categorising research outputs within the field of archaeological computing. As for spatial and temporal metadata, the platform had already adopted a gazetteer-based approach in its initial configuration, and this method was preserved. However, the need to support advanced visualisation features, such as interactive maps and timelines, prompted a refinement in how spatial and temporal references were modelled. To facilitate more dynamic and flexible data presentation, spatial and temporal references were restructured as independent items. Specifically, two new classes were introduced: *dcterms:PeriodOfTime* for chronological data and *dcterms:Location* for geographic references.

On the temporal side, this strategy allowed for consistent integration with the PeriodO gazetteer, maintaining continuity with previous practices. On the spatial side, it enabled the combination of data from multiple gazetteers, addressing the limitation that could arise from relying solely on a single resource, such as Pleiades, which might lack representation for exclusively modern locations. Establishing specific items for places and periods not only enhanced the precision of metadata but also facilitated semantic linking, filtering, and visual exploration of resources across spatial and temporal dimensions. Additionally, the thematic, geographic, and chronological indexing functions as a key entry point for users which is particularly valuable given the inherent challenges of designing and anticipating navigation paths in LOD-based platforms (CANDELA 2025; Fu *et al.* 2024).

#### 4. CURRENT STATUS AND FUTURE PROSPECTIVES

Although the data collection process is still ongoing, the volume and diversity of material currently indexed in the DHeLO platform provide a solid foundation for preliminary analysis. The methodology adopted for harvesting and selecting data continues to follow the criteria outlined in MANCUSO, D'EREDITÀ (2024), which guided the initial mapping of research products and projects, strongly focused on the Italian research panorama. As of April 2025, DHeLO includes metadata for 336 datasets and 103 interactive resources, along with 67 software tools directly associated with their production. These are linked to 176 research projects, reflecting the contributions of 1534 individuals and over 300 research institutions. The digital outputs are geographically mapped to 354 sites and temporally anchored across 104 historical periods.

During the mapping process, metadata were also collected for several datasets described in the academic literature but not openly accessible online. These include research outputs that lack a stable web presence and therefore cannot be directly linked through the platform. Out of a total of 336 datasets currently recorded in the system, 162 fall into this category. While these datasets are not displayed in the public interface, their metadata are retained within the backend for internal monitoring and analysis. This approach reflects DHeLO's dual role: as a discovery tool connecting existing resources, and as an observatory tracking digital production across the CH/HS landscape. In order to appear publicly, each record must be associated with an external access point; where this is absent, the dataset remains private but contributes to the statistical backbone of the platform.

In conclusion, the development of DHeLO, in its current LOD-based configuration, has reaffirmed a key principle in the design of research infrastructures: the centrality of the community. Far from being a passive user base, the community of scholars, developers, and professionals working across CH, HS, and DA has been conceived as a source of insight, validation, and even reorientation throughout the reshaping of this tool. The work carried out through WP2 has not only offered the methodological tools to observe this landscape but has also contributed a framework that helped shape the platform itself, grounding its evolution in the needs, values, and expectations expressed by its future users. This dynamic relationship between infrastructure and community is particularly evident in both the layered architecture of DHeLO and its broader potential. By integrating metadata about research products, projects, people, and bibliographic references, the platform can lay the foundation for a domain-specific knowledge graph in which connections are not imposed but emerge organically from the fabric of research activity (MOSCATI 2021). The forthcoming integration with BiDiAr, in response to a clear need from a community that still communicates predominantly through literature, will further strengthen this relational potential, allowing users to navigate between datasets, publications, contributors, and institutional affiliations. This integration is taking shape both within DHeLO and BiDiAr themselves, and as part of the broader framework of the Open Digital Archaeology Hub (CARVALE *et. al.*, in this special section).

Beyond its role as a monitoring tool, DHeLO is increasingly envisioned as a research resource. By exposing curated metadata in open and interoperable formats, and by linking entities through established ontologies and controlled vocabularies, the platform enables new forms of exploration and reuse that align with the initial stages of FAIR compliance. In this sense, DHeLO is not merely documenting the state of digital practices but is actively contributing to their discoverability, reuse, and conceptual integration. Looking ahead, DHeLO's trajectory is marked by both continuity and transformation. On the one hand, the platform will continue to serve as a space for collecting,

organising, and exposing knowledge about digital production within the CH/HS research community. On the other, it is poised to support increasingly exploratory, research-oriented uses, contributing not only to infrastructure, but to epistemology: helping to visualise disciplinary boundaries, track methodological trends, and identify gaps or asymmetries in digital engagement. In this evolving role, DHeLO exemplifies a model of infrastructure that is not simply technological, but intellectual; an interface not just for accessing resources, but for reflecting on the shape and trajectory of research itself.

GIACOMO MANCUSO

Istituto di Scienze del Patrimonio Culturale - CNR  
giacomo.mancuso@cnr.it

### *Acknowledgements*

This research was supported by H2IOSC Project - Humanities and cultural Heritage Italian Open Science Cloud funded by the European Union NextGenerationEU - National Recovery and Resilience Plan (NRRP) - Mission 4 'Education and Research' Component 2 'From research to business' Investment 3.1 'Fund for the realization of an integrated system of research and innovation infrastructures' Action 3.1.1 'Creation of new research infrastructures strengthening of existing ones and their networking for Scientific Excellence under Horizon Europe' - Project code IR0000029 – CUP B63C22000730005. Implementing Entity CNR.

### REFERENCES

- BENARDOU A., CHAMPION E., DALLAS C., HUGHES L. (eds.) 2017, *Cultural Heritage Infrastructures in Digital Humanities*, London, Routledge (<https://doi.org/10.4324/9781315575278>).
- CANDELA G. 2025, *Browsing Linked Open Data in Cultural Heritage: A shareable visual configuration approach*, «Journal on Computing and Cultural Heritage», 18, 1, 1-15 (<https://doi.org/10.1145/3707647>).
- CARAVALE A., MOSCATI P., ROSSI I. 2024, *Advancements of the H2IOSC Project: Enhancement of digital resources in the Cultural Heritage field targeting archaeology and epigraphy. Introduction to the Special section*, «Archeologia e Calcolatori», 35.2, 463-470 (<https://doi.org/10.19282/ac.35.2.2024.48>).
- FU B., LAMBRIX P., PESQUITA C., VEGA-GORGOJO G. 2024, *LOD4Culture: Easy exploration of cultural heritage linked open data*, «Semantic Web», 15, 5, 1563-1592 (<https://doi.org/10.3233/sw-233358>).
- HUGGETT J. 2020, *Is big digital data different? Towards a new archaeological paradigm*, «Journal of Field Archaeology», 45, sup1, S8-S17 (<https://doi.org/10.1080/00934690.2020.1713281>).
- JIROTKA M., LEE C.P., OLSON G.M. 2013, *Supporting scientific collaboration: Methods, tools and concepts*, «Computer Supported Cooperative Work (CSCW)», 22, 4, 667-715 (<https://doi.org/10.1007/s10606-012-9184-0>).
- LUZIETTI R.B., SPADI A., GIAMPIETRO N., MANCUSO G., CARAVALE A., D'EREDITÀ A., CARADONNA M., MOSCATI P., QUOCHI V., MONACHINI M., DEGL'INNOCENTI E. 2024, *Digital Humanities and Heritage Science: Moving from landscaping to a dynamic research observatory in an Open Science Cloud*, in A. DI SILVESTRO, D. SPAMPINATO (eds.), *Me.Te. Digitali. Mediterraneo in rete tra testi e contesti, Proceedings del XIII Convegno Annuale AIUCD (Catania 2024)*, «Quaderni di Umanistica Digitale», 599-565 (<https://dx.doi.org/10.6092/unibo/amsacta/7927>).

- MANCUSO G., CARVALE A., D'EREDITÀ A., MOSCATI P. in press, *Merging Knowledge and Tools in Heritage Science and Digital Archaeology. Practices from H2IOSC WP2*, presented to DIGITAL HERITAGE 2025. *The Premier Global Forum where Culture Meets Cutting-edge Technology* (Siena 2025).
- MANCUSO G., D'EREDITÀ A. 2024, *DHeLO and BiDiAr: New digital resources within the H2IOSC Project*, «Archeologia e Calcolatori», 35.1, 521-542 (<https://doi.org/10.19282/ac.35.1.2024.31>).
- MANCUSO G., D'EREDITÀ A. in press, *Two FOSS resources for Cultural Heritage and Digital Archaeology in the H2IOSC project*, presented to ARCHEO.FOSS XVIII 2024. *International Conference on Open Software, Hardware, Processes, Data and Formats in Archaeological Research* (Chieti 2024).
- MARTILA S., BOTERO A. 2021, *Infrastructuring for collective heritage knowledge production*, in M. RAUTERBERG (ed.), *Culture and Computing. Design Thinking and Cultural Computing. HCII 2021. Lecture Notes in Computer Science*, Springer International Publishing, 95-112 ([https://doi.org/10.1007/978-3-030-77431-8\\_6](https://doi.org/10.1007/978-3-030-77431-8_6)).
- MARWICK B., BIRCH S.E.P. 2018, *A standard for the scholarly citation of archaeological data as an incentive to data sharing*, «Advances in Archaeological Practice», 6, 2, 125-143 (<https://doi.org/10.1017/aap.2018.3>).
- MOSCATI P. 2021, *Digital archaeology: From interdisciplinarity to the 'fusion' of Core competences*, «magazén. International Journal for Digital and Public Humanities», 2, 2, 253-274 (<https://doi.org/10.30687/mag/2724-3923/2021/04/004>).
- PAWLICKA-DEGER U. 2021, *Infrastructuring digital humanities: On relational infrastructure and global reconfiguration of the field*, «Digital Scholarship in the Humanities», 37, 2, 534-550 (<https://doi.org/10.1093/lcf/qab086>).
- ROSS J., SHEAIL P., TERRAS M., SCHOFIELD C. 2024, *Infrastructure Futures for Digital Cultural Heritage*, University of Edinburgh (<https://doi.org/10.5281/zenodo.13710265>).
- RUEFF B. 2024, *Dealing with post-excavation data: The Omeka S TiMMA web-database*, «Peer Community Journal», 4, 1-14 (<https://doi.org/10.24072/pcjournal.366>).
- SCHMIDT B., GEMEINHOLZER B., TRELOAR A. 2016, *Open Data in Global Environmental Research: The Belmont Forum's Open Data Survey*, «PLOS ONE», 11, 1, e0146695 (<https://doi.org/10.1371/journal.pone.0146695>).
- SCARPA E., VALENTE R. 2024, *A resource hub for interoperability and data integration in Heritage research: The H-SeTIS database*, «Archeologia e Calcolatori», 35.1, 543-562 (<https://doi.org/10.19282/ac.35.1.2024.32>).

## ABSTRACT

This article presents the transformation of DHeLO (Digital Heritage Landscaping Platform) from a relational database into a Linked Open Data (LOD) infrastructure, designed to enhance the discoverability, interoperability, and reuse of digital resources within Cultural Heritage (CH), Heritage Science (HS), and Digital Archaeology (DA) research. Initially conceived within the H2IOSC project as a virtual observatory for cataloguing digital products, tools, and research projects, DHeLO has been restructured as a service-oriented system in response to emerging community needs and a broader reflection on its role. The transition was guided by insights gained from landscaping activities conducted within Work Package 2, including questionnaires and interviews with members of the DH, DA, and HS communities. The new configuration of DHeLO integrates metadata on research outputs, projects, people, and bibliographic references, laying the foundation for a knowledge graph that mirrors the research landscape. A key step in this evolution was the adoption of Omeka S as the core platform, chosen for its modularity, semantic interoperability, and synergy with Zotero, which supports the integration of the BiDiAr bibliographic database. In this form DHeLO aspires to become an active research resource, moving beyond mere monitoring to foster exploratory analysis and linked knowledge networks.