Archeologia e Calcolatori
35.1, 2024, 515-520
doi 10.19282/ac.35.1.2024.30

LANDSCAPING AND INTEGRATING DIGITAL ARCHAEOLOGY AND DIGITAL EPIGRAPHY RESOURCES: NEW CHALLENGES AND FUTURE OPPORTUNITIES. INTRODUCTION TO THE SPECIAL SECTION

1. Introduction

The H2IOSC Project ('Humanities and cultural Heritage Italian Open Science Cloud'), funded by the European Union Next Generation EU and the Italian Ministry of University and Research as part of the National Recovery and Resilience Plan (NRRP), started in November 2022 with the aim to establish a federated cluster of the Italian nodes of the CLARIN, DARIAH, E-RIHS and OPERAS[1] research infrastructures in the ESFRI (European Strategy Forum on Research Infrastructures) domain focusing on Social and Cultural Innovation. This endeavour seeks to foster collaboration among researchers from various disciplines in the Humanities, Language technologies and Cultural Heritage sectors, enabling them to engage in data – and compute-intensive research. The Project (https://www.h2iosc.cnr.it/) is structured into eight Work Packages, with archaeologists and epigraphists playing a significant role in the sub-packages pertinent to the E-RIHS infrastructure.

This special section of the journal «Archeologia e Calcolatori» is dedicated to showcasing the research activities conducted by the CNR-ISPC Open Data, Open Knowledge, Open Science research group[2] as part of the H2IOSC Project. The featured research spans across three Work Packages: WP2 (Landscaping & Building Communities), WP4 (RIs Nodes and Resources Interoperability), and WP7 (Community Pilots: Innovative Cross-domain Services and Environments). Our editorial programme envisages publishing three special thematic sections in the journal's 35.1, 35.2 and 36.1 issues. The main objective is to promote and share the research outcomes by providing free access to scientific data, following an approach that aligns with our journal's tradition. At the same time, these publications will enrich the federated cluster of Italian nodes of the leading research infrastructures in the Humanities and Cultural Heritage, by creating awareness that knowledge and interpretation of the past can inform future technological decisions.

[1] http://dariah.cnr.it/; https://www.clarin-it.it/; https://www.e-rihs.it/; https://operas-eu.org/.
[2] https://www.ispc.cnr.it/it_it/2020/05/14/gruppo-open-data/.

2. Improve knowledge on Digital Archaeology resources and classification

As part of the WP2, research activity 2.4 focuses on 'Landscaping the Cultural Heritage and Heritage Science resources and needs panorama'. This activity is dedicated to surveying and detailing the existing resources and requirements within the Cultural Heritage (CH) and Heritage Science (HS) research communities participating in the Italian E-RIHS network. It involves a series of interconnected tasks aimed at identifying and highlighting the available resources and repositories, their specific disciplinary domains, the main services and tools developed, and the best practices, protocols, and standards employed. The objective is to ensure that the production and sharing of datasets adhere to the FAIR principles, promoting their interoperability.

Since the start of the H2IOSC Project, specific attention has been placed on classifying the main cross-cutting themes that characterise the Digital Archaeology research domain and its current integration within Digital CH and HS. In the context of the articles published in the international journal «Archeologia e Calcolatori», this evolution necessitates a fresh examination of the methods employed to classify the wide-ranging areas of application that now span a vast array of technological innovations, posing challenges to conventional classification frameworks. Thanks to the application of Machine Learning (ML) techniques, the implementation of a knowledge map has been instrumental in identifying topic clusters and keywords that exemplify the ten dominant technological approaches[3] to analyse and interpret archaeological data and in highlighting the pivotal contribution of archaeology – particularly computational archaeology – to the multidisciplinary domain of HS[4].

Within the activities illustrated in this special section, we delve deeper into the advancements made in the frame of WP2 and, therefore, into the landscaping activities. The definition of a collaborative framework between the disciplinary communities associated with the four research infrastructures helped us to establish general guidelines for conducting the survey, collecting information and mapping their technological landscape. The main focus of the activity was centred around three primary goals: the development of a web-based questionnaire structure, the selection of candidates for in-person interviews, and the initiation of a semi-automated search for existing resources, tools, software and services.

In addition to the active participation of the E-RIHS community in both the design and the dissemination of the questionnaire, which is set to be distributed

[3] Artificial Intelligence; GIS and Spatial Analysis; Imagery Analysis; Material Culture; Modeling and Simulation; Digital Cultural Heritage; Photogrammetry and 3D Scanning; Remote Sensing; Semantic Technologies; Virtual Reality.

[4] Caravale A., Duran-Silva N., Grimau B., Moscati P., Rondelli B. 2023, *Developing a digital archaeology classification system using Natural Language Processing and Machine Learning techniques*, «Archeologia e Calcolatori», 34.2, 9-32 (https://doi.org/10.19282/ac.34.2.2023.01).

online via the Lime Survey service thanks to the licensing and server hosting provided by CNR (https://www.limesurvey.org/), our research group strategically developed two operational databases: DHeLO and BiDiAr (see Mancuso, D'Eredità, in this special section). They are designed to closely monitor the digital resources of the CH/HS domains, along with their extensive bibliographic references. These tools are intended to be both flexible and versatile, enabling the storage of data collected via the questionnaire distribution and facilitating interaction with other landscaping initiatives promoted by WP2, such as the CLARIN Virtual Language Observatory (VLO). Furthermore, the data collected can serve as input for the E-RIHS DIGILAB platform and contribute to implementing the Open Digital Archaeology and Epigraphy Hubs, which are being developed as an E-RIHS pilot project within H2IOSC's WP7.

As will be extensively discussed, DHeLO's primary objective is to establish an observatory for digital products to provide a deeper insight into the dynamic scenario of CH/HS and better meet its needs in terms of digital tools. The database conceptual scheme enables efficient mapping of digital products, providing, in particular, an in-depth exploration of the progress of Digital Archaeology, including an examination of the ten thematic clusters mentioned above (see note 3). This ongoing data analysis progressively reveals common practices, shared standards, and similar workflows across various research projects and institutions. This holistic approach enhances our understanding of Digital Archaeology's trajectory, fostering improved cooperation and informed decision-making.

P.M.

3. Digital Archaeology: collection and sharing of bibliographic resources

In an effort to create new resources for the Italian node of the E-RIHS infrastructure, a new bibliographic reference tool primarily focusing on Digital Archaeology has been developed to be implemented in the H2IOSC Marketplace. Named BiDiAr, this tool is a product of the CNR Institute's 30 years of activity in the field of archaeological computing. It not only documents the most recent developments but also facilitates the preservation of past events and contexts. Indeed, in recent years, archaeological computing, a relatively young discipline, has seen significant growth in studies and research, due to the widespread integration of digital technology across all the different investigation areas. This has not only broadened the perspectives of analysis, but has also revolutionised the methods of collecting, preserving, and using archaeological data, opening up new horizons of comprehension, interpretation, and dissemination of the past through technology. Analysing the evolution of this discipline through its bibliography provides a historical and conceptual overview, starting from the earliest works focused on a small number of key topics up to the current research

panorama characterised by a diversified and multifunctional use of machines and technologies. The availability of a bibliographical corpus makes it possible not only to benefit from an organised and systematic framework of a constantly expanding discipline, but also to have a grasp of the sectors that today have achieved major relevance while identifying those that are less visited.

The CNR-ISPC Open Data, Open Knowledge, Open Science research group acknowledges the pivotal role of bibliography in understanding a specific disciplinary sector. Over time, the group has conducted different research in this domain, sharing their outcomes online. An important initiative is the Virtual Museum of Archaeological Computing (https://archaeologicalcomputing.lincei.it/), representing the multimedia apex of an international project dedicated to the historical development of computers in archaeology. This project originated from a fruitful collaboration between Accademia Nazionale dei Lincei and Consiglio Nazionale delle Ricerche. The website provides access to several bibliographies, beginning with those associated with prominent scholars in archaeological computing who have been active since the 1950s (J.-C. Gardin, A.C. Spaulding, D.L. Clarke, G.L. Cowgill, J.E. Doran, R. Ginouvès, F.R. Hodson). The portal also contains the Bibliography of Archaeological Computing of the 1990s, a rich repository comprising over 2700 titles spanning the period from 1989 to 2000, which has been the subject of a specific study[5]. The study aimed to describe the international cultural panorama of the 1990s, connecting it with the achievements of the previous decades and with the challenges of the subsequent years, which have deeply influenced the development of Digital Archaeology in the new millennium.

BiDiAr is a new research tool in the field of Digital Archaeology and Cultural Heritage (see Mancuso, D'Eredità, in this special section). It is developed in H2IOSC-WP2 using bibliographic data collected from recent volumes of «Archeologia e Calcolatori», which has been a reference point in this sector since 1990. As described in this special section, the repository archives over 6500 titles using the open-source software Zotero, providing access to a rich body of knowledge and offering an overview of studies that, although focused on recent years, provides a diachronic perspective on this research area, allowing its evolution and trends to be tracked over time. The tool's potential lies in its continuous implementation and updating, enabling ongoing monitoring of the discipline's status. Additionally, it can be linked with other tools developed within the same research group, such as the above-mentioned DHeLO and IADI (A&C Interactive Atlas of Digital Images; https://iadi.archcalc.cnr.it/), aimed at creating a digital ecosystem that offers an integrated approach to data consultation, fostering their sharing and dissemination.

A.C.

[5] Caravale A., Moscati P. 2021, *La bibliografia di informatica archeologica nella cultura digitale degli anni Novanta*, Firenze, All'Insegna del Giglio.

4. Towards a semantic framework for the interoperability of Cultural Heritage resources: the role of Digital Archaeology and Digital Epigraphy

Central to H2IOSC's objective of facilitating data integration and supporting research in the Humanities, Language technologies and Cultural Heritage sectors are the activities that CNR-ISPC is conducting in the frame of the Project's WP4. The CNR-ISPC Milan branch of the Open Data, Open Knowledge, Open Science group leads Activity 4.10, which aims at ensuring interoperability in the domain of Cultural Heritage and at contributing to the elaboration of a shared semantic framework among the four RIs involved in the Project to allow their resources to be onboarded into an integrated, cross-domain environment – the H2IOSC Marketplace. Activity 4.10 is related to the resources which will populate DIGILAB, E-RIHS's digital platform offering virtual access to tools and data in the field of Heritage Science. These activities encompass identification and mapping of metadata schemas, alignment to community standards, ontology design and application.

The contribution by Scarpa and Valente in this special section tackles the need for a survey of Knowledge Organisation Systems in the Heritage sector as a preliminary step for resources' harmonisation and integration, with a focus on semantic artefacts (i.e. knowledge models which are intelligible – and therefore automatedly processable – by machines, as the authors explain). The first result of this survey activity has been the setting up of H-SeTIS – Heritage - Semantic Tools and Interoperability Survey, an open access observatory of semantic artefacts in the Heritage domain, providing descriptions of resources and tools such as ontologies, metadata schemas, thesauri, application profiles and software. H-SeTIS, in turn, has been designed as a tool whose information is mapped to standard schemas and is exposed in interoperable formats through a web-service. The bibliographic catalogue of the semantic survey, whose entries are associated to the records of the H-SeTIS database, is maintained as an open Zotero group.

The use of keywords for describing the disciplinary domain(s) of each resource catalogued in H-SeTIS allows to draw some preliminary qualitative and quantitative reflections on the semantic landscape of the Heritage sector. This shows a longstanding interest and concerted effort towards a shared description of Cultural Heritage-related digital resources in the fields of Archaeology, Museum and Text studies. On the other hand, disciplines concerned with diagnostic analyses of cultural objects suffer from a lack of contents' formalisation and integration with existing semantic tools in the field of Cultural Heritage. To address such issue, CNR-ISPC will design an ontology for modeling the Heritage Science domain that integrates into the H2IOSC semantic platform. Crucial to this final aim is the acknowledgement, among the information catalogued in H-SeTIS, of the development/maintenance status of the described resources, which is an essential aspect to sift through them and provide an up-to-date

panorama of the reliable and usable semantic tools, while acknowledging the closed or abandoned initiatives to document the history of studies in the field.

The integration of cross-domain resources enabled by WP4's activities is going to be applied and tested in H2IOSC through a series of pilot projects, to be developed within WP7. As part of E-RIHS, CNR-ISPC is currently involved in the creation of several digital hubs which span and intersect the multiple disciplinary fields of Heritage Science. Activity 7.4 is working towards the implementation of two case studies related to the fields of Digital Archaeology and Digital Epigraphy that will benefit from the services provided by the H2IOSC RIs in terms of tools and resources for material heritage studies and text studies. Starting from the CNR decennial experience of Diamond Open Access publishing in the Digital Archaeology sector, represented by the journal «Archeologia e Calcolatori», the Open Digital Archaeology Hub will provide archaeologists with accessible resources to experiment with innovative technologies for the extraction of structured information from traditional textual and iconographic corpora, thus stimulating awareness of Open Science in the archaeological field[6]. The Open Digital Epigraphy Hub, on the other hand, aims at building an open platform for the discovering and sharing of resources, tools and methods in the field of Digital Epigraphy, leveraging the experience of the open access 'Digital Archive for the Study of pre-Islamic Arabian Inscriptions' (DASI: https://dasi.cnr.it/)[7].

Indeed, the peculiar nature of the inscriptions – which are both textual and material sources – has put the research on epigraphic corpora at the crossroads of Digital Humanities, Computational Linguistics and Heritage Sciences. 'Epigraph' appears thus as a complex object of study, one that is especially apt to test technological solutions developed within the Project for its interdisciplinary goals. The joint effort that WP7 and WP4 are undertaking to identify a common semantic framework of H2IOSC starts precisely from the case study of 'text-bearing objects', where epigraphs as resources represent the various facets of the textual, art-historical and archaeological inquiries onto which the diagnostic processes of the analytical sciences of Cultural Heritage can be applied. These activities will be described in forthcoming contributions in future issues of this journal.

I.R.

Alessandra Caravale, Paola Moscati, Irene Rossi
Istituto di Scienze del Patrimonio Culturale - CNR
alessandra.caravale@cnr.it, paola.moscati@cnr.it, irene.rossi@cnr.it

[6] Rossi I., Paraciani N. 2021, *IT applications to archaeology and the OA diamond journals' challenge. Enhancing access and reuse of textual and visual resources*, «Archeologia e Calcolatori», 32.1, 325-347 (https://doi.org/10.19282/ac.32.1.2021.18).

[7] De Santis A., Rossi I. (eds.) 2018, *Crossing Experiences in Digital Epigraphy: From Practice to Discipline*, Berlin-Boston, De Gruyter (https://doi.org/10.1515/9783110607208).