ARIADNE: A Data Infrastructure for the Archaeological Research Community

by Alessia Bardi, Massimiliano Assante and Francesco Mangiacrapa (CNR-ISTI)

The ARIADNE infrastructure provides tools and services for researchers to address archaeological grand challenges that require discovery and analysis of information scattered across different thematic and geographically distributed sources.

ARIADNE (Advanced Research Infrastructure for Archaeological Dataset Networking in Europe) is an infrastructure for the archaeological research community funded by the European Commission via the projects ARIADNE (from 2013 to 2017) and ARIADNEplus (from 2019 to 2022). [L1] The infrastructure offers a suite of services and tools to support research in archaeology in different phases of the research life cycle: tools for data visualisation (e.g. 3D modelling, geoportal); virtual research environments (VREs) for data analysis, data sharing and human-centred collaboration; and a knowledge base of archaeological resources, accessible as Linked Open Data and searchable via a web portal.

At the end of the ARIADNEplus project (December 2022), the 41 partners [L2] successfully extended the scope of the first ARIADNE project, delivering a digital infrastructure based on the D4Science e-infrastructure [3] developed and maintained by CNR-ISTI, with: (i) a knowledge base of 3.5 million archaeological resources including archaeological reports, findings, inscriptions, archaeological sites and monuments from archives and repositories in Europe and beyond (e.g. Argentina, Japan); and (ii) a set of Virtual research Environments (VREs) where archaeological researchers can find different tools for data exploration and analysis.

The knowledge base was built via an aggregation system, capable of collecting detailed descriptions of archaeological resources (metadata records) in different formats, transforming the records according to a common data model based on the CIDOC-CRM ontology, enriching them with dating information via PeriodO [L3] and subject terms via Getty AAT [L4]. For each dataset, an automated workflow has been devised via the D-NET Software Toolkit, a software for the realisation of aggregative metadata infrastructures developed by CNR-ISTI [1], and configured to apply a mapping defined with the 3M Editor developed by FORTH-ICS [2]. As a result of the mapping, each input metadata record is transformed into an RDF (Resource Description Framework) record suitable for ingestion into the ARIADNE knowledge base, an instance of GraphDB [L5]. For the enrichment with dating information, providers were invited to curate an authority on PeriodO, while for the enrichment with Getty AAT, they defined a mapping between their local terms and Getty AAT terms with the Vocabulary Matching Tool developed by University of South-Wales [L6]. SPARQL queries were then used to exploit such information to enrich the records in the knowledge base. The final step of the automated workflow is the publishing on the ARIADNE portal [L7], developed by the Swedish National Data Service, where the resources can be searched and filtered by different criteria (e.g. by location, by historical period, by subject, by contributor).

The VREs provided by ARIADNE are designed to meet the needs of its target community and scenarios. These environments are dedicated to exploring and developing various solutions for specific research questions and serving specific communities of practitioners in the archaeology domain. For instance, the Analytics Lab VRE [L8] offers archaeologists and scholars a virtual laboratory and set of tools for analysing and manipulating data to answer their research questions. The VRE allows users to import their own data files and those from the ARIADNE knowledge base, rearrange and analyse them, and generate statistical reports based on typology, geographical distribution and other factors. The results can also be visualised in the form of complex graphs and thematic maps. The D4GNA VRE [L9] (Dataset for the National Archaeology Geoportal) instead collects data intended to converge into the Italian National Archaeology Geoportal (GNA) under the auspices of the Italian Ministry of Culture. As Figure 1 shows, it allows any user to freely access [L10] and consult documentation related to archaeological excavations in Italy, and authorised users to collaboratively manage these contents. Users can
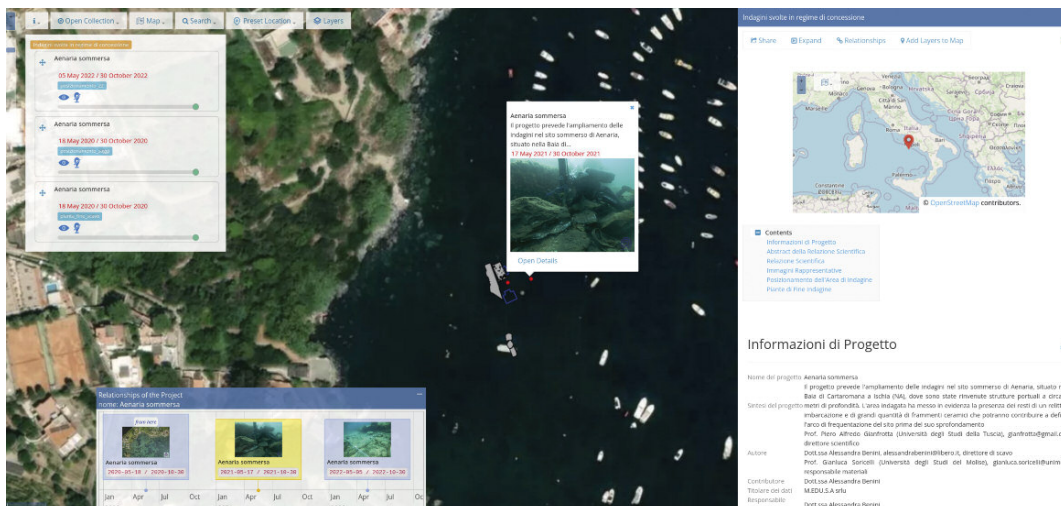
Figure 1: The geoportal viewer of D4GNA VRE.

discover the products either by interacting with the web-map or by searching for products by keywords, author, date and other metadata, and preview the products in a variety of formats, including maps, images and documents. This VRE is equipped with a dedicated service (geoportal) acting as the underlying storage and management layer for the products and their metadata, responsible for ensuring the persistence and accessibility of the data, and for enforcing the access control policies defined by the system administrators. Its data model is rich and flexible, designed to support the discovery and sharing of the products, and to allow users to easily link and connect products based on their metadata and contents.

In conclusion, the ARIADNE infrastructure provides a comprehensive solution for archaeological research, including a range of services and tools to help researchers throughout the research life cycle. From creating data management plans using templates, to visualising data using 3D modelling and geoportal tools, this offer is complemented with VREs designed to be user-friendly and accessible to users of all levels, whether they are experts or novice researchers. This allows researchers to focus on their work, rather than struggling with complex technology. Researchers can store their data securely and share it with their collaborators with ease, knowing that their data is protected by the highest levels of security and privacy. This helps to foster collaboration and innovation in the research community, as researchers can work together to tackle complex problems and make new discoveries. The knowledge base of archaeological resources is accessible as Linked Open Data and can be easily searched through the web portal, ensuring researchers have access to the resources they need to succeed.

Links:
[L1] https://ariadne-infrastructure.eu/
[L2] https://ariadne-infrastructure.eu/partners/
[L3] https://perio.do/en/
[L4] https://kwz.me/hwW
[L5] https://graphdb.ontotext.com/
[L6] https://vmt.ariadne.d4science.org/vmt/vmt-app.html
[L7] https://ariadne-portal.d4science.org/
[L8] https://ariadne-infrastructure.eu/the-ariadneplus-lab-vre/
[L9] https://gna.d4science.org/
[L10] https://gna.d4science.org/d4gna-data-viewer

References:
[1] P. Manghi, et al., "The D-NET Software Toolkit: A framework for the realization, maintenance, and operation of aggregative infrastructures", Program Electron Lib, vol. 48, no. 4, pp. 322–354, 2014. https://doi.org/10.1108/PROG-08-2013-0045
[2] Y. Marketakis, et al., "X3ML mapping framework for information integration in cultural heritage and beyond", Int. J. Digit. Libr., vol. 18, pp. 301–319, 2017. https://doi.org/10.1007/s00799-016-0179-1
[3] M. Assante, et al., "Enacting open science by D4Science", Future Generation Computer Systems, vol. 101, pp. 555–563, 2019. https://doi.org/10.1016/j.future.2019.05.063

Please contact:
Alessia Bardi, CNR-ISTI, Italy
alessia.bardi@isti.cnr.it

Boosting Open Science in the IPERION HS Research Infrastructure with OpenAIRE

by Alessia Bardi (CNR-ISTI) and Laura Benassi (CNR-INO)

IPERION HS is a research infrastructure that supports researchers in the field of heritage science, an interdisciplinary domain studying cultural and natural tangible heritage. This article describes how IPERION HS makes the research outputs open and accessible to the community and monitors its impact thanks to the services offered by the OpenAIRE infrastructure.

IPERION HS [L1] is a European infrastructure project working in the interdisciplinary domain of heritage science to understand, study and protect natural and cultural heritage. It contributes to establishing a pan-European research infrastructure on heritage science (E-RIHS). Led by the Italian National Research Council (National Institute of Optics), the consortium is composed of 24 national nodes across Europe and the US and collaborates with ICCROM, the intergovernmental organisation dedicated to the preservation of cultural heritage worldwide.

IPERION HS offers training and access to a wide range of high-level scientific instruments, methodologies, data and tools for advancing knowledge and innovation in heritage science. The research is actually focused on a better interoperability to data, sample materials, methods and instruments. Together with the users, the IPERION HS community is collecting an incredible amount of data and samples and working on the development of new protocols and instruments.

Making heritage science data open, findable and accessible is one of the main objectives of the IPERION HS infrastructure, whose Open Science strategy does not only include open access to publications, but also FAIR (findable, accessible, interoperable and reusable) and open data and methodologies. Tracking and monitoring the success of the strategy calls for technical tools capable of gathering together the research outputs of the IPERION researchers, which are published and deposited across different venues and repositories, and analysing the gathered data according to different perspectives, such as the types of outputs, their access rights, the policies of the journals they are published in, the availability of persistent identifiers and rich descriptive metadata, and the affiliations of the authors.

In addition, IPERION HS aims to support and ease the discovery of heritage science resources. Due to the high multidisciplinary aspect of the field, heritage science research resources are scattered in different repositories, archives, journals across the world and across many disciplines. Heritage science researchers, therefore, do not have a single web portal for searching for heritage science resources. On the contrary, they have to use many different thematic portals or the portals of large domain-agnostic aggregators. In the first case, re-