Please cite this as: Correia, M.J. and Santos Silva, A. 2023 DB-HERITAGE Building Materials Data Aggregation in ARIADNE - challenges and opportunities, Internet Archaeology 64. https://doi.org/10.11141/ia.64.5

DB-HERITAGE Building Materials Data Aggregation in ARIADNE - challenges and opportunities

Maria J. Correia and António Santos Silva

DB-HERITAGE provides open and free sharing of wide-ranging technical data on hundreds of samples of building materials from diverse periods, extracted from Portuguese built heritage all over the world. It has been developed to improve know-how for historic building materials and as a basis for developing best practices for built heritage conservation. It incorporates both physical and digital repositories for building material samples and related data, providing tools for the systematic recording of data concerning the history, properties, and performance of materials used in Portuguese built heritage.

DB-HERITAGE targets different communities, researchers, and stakeholders involved in the preservation of archaeological and architectural heritage. It provides a systematisation of building assets, within their related environmental, social, and cultural contexts, as well as displaying technical and scientific information on structural elements, built components and their constituent materials. Although the rationale of the wide context sustaining the DB-HERITAGE concept is clear, coordinating communities and reduced time-to-value represent extra requirements for data quality, improved tools, and an efficient management plan.

The participation in ARIADNEplus challenged DB-HERITAGE to get the most out of standardised procedures and the FAIR principles, strengthening its data management plan and practices. Data processing has been improved by deploying a common ontology and further developing standards, shared semantics, and identifiers. Updated protocols for data sharing and detailed information on provenance have also been developed to enhance data reuse.

This article presents an overview of the aggregation process of DB-HERITAGE data into ARIADNEplus. It includes a summary of DB-HERITAGE's strengths and of the challenges faced within the scope of the aggregation process, with examples of some of DB-HERITAGE's major outputs. Additionally, it considers the benefits and opportunities provided by participation in ARIADNEplus.

1. Introduction

In recent decades, several studies have been carried out on the composition and characteristics of building materials, as well as on their deterioration and ageing mechanisms (e.g. archaeological site of Tróia conservation condition assessment - Santos Silva *et al*. 2006). These studies, aimed particularly at archaeological and architectural built heritage, hold the utmost importance for improving research and for preserving built heritage.

The characterisation of materials and assessment of their condition are fundamental for planning conservation and restoration interventions for heritage assets, including selecting suitable and compatible materials and techniques. Therefore, data on built heritage and related building materials are essential in order to improve knowledge about ancient materials' history, properties, and techniques; for performance modelling benchmarking of materials; for innovation on new materials; and to support the decision-making process for the conservation and restoration of the built heritage.

Likewise, the analysis of data about the conservation condition of built heritage, including the history of interventions and characteristics and performance of building materials, is essential for developing sustainable and resilient policies, and for innovation and research progress aimed at prolonging the lifespan and use of built heritage. Besides safeguarding building materials-related research, there are also demands for further innovation and development regarding built heritage protection.

Figure 1: DB-HERITAGE project logo and page header image (see http://db-heritage.lnec.pt/index_en.html)

Besides the historical value and its legacy for present and future generations, the compilation and sharing of information about built heritage have significant cultural, artistic, and educational value. This information, especially if embedded in conservation strategies, contributes to a greater public awareness of built heritage and to so-called cultural heritage tourism. This also favours local employment, particularly in areas that use traditional arts and materials, and it contributes to the settlement of populations in less developed inland parts of Portugal.

Currently, however, a large amount of information on materials and ancient techniques, essential for supporting informed decision-making on asset management, is scattered and therefore at risk of being lost. Moreover, much of the information available on the study of building materials is managed by different individuals and institutions in Portugal and is far from being systematised. Despite the various information systems for cultural heritage currently available in Portugal, there are noticeable problems in getting the most out of FAIR data (Correia *et al*. 2019). Mandatory data properties and descriptions are limited, and there is usually restricted access to data. Despite being indexed, they are not easily interoperable, findable, and accessible.

This has justified the creation of DB-HERITAGE, which aims to improve knowledge of the conservation and restoration of architectural and archaeological built heritage (Figure 1). DB-HERITAGE was developed within a project funded by the Portuguese Foundation for Science and Technology (FCT) entitled 'Database of construction materials with historical and heritage interest' (Project ID PTDC/EPH-PAT/4684/2014). This was a joint project between the Universities of Aveiro, Évora and NOVA from Lisbon and included the collaboration of several Portuguese entities besides the National Laboratory for Civil Engineering (LNEC), the organisation responsible for project coordination.

DB-HERITAGE aims to support not only engineering and technology, natural sciences, and social and historical fields of research, but also to contribute to the improvement of the construction industry (e.g. building material producers and bodies working in the design, management or execution of the conservation, restoration, and maintenance of heritage).

The whole concept of DB-HERITAGE is aimed at preserving building material samples and related data, keeping their provenance, and providing all the information required by different communities and end-user groups. DB-HERITAGE displays a hierarchical description of built heritage and its linked data, including details of a structure's historical and technical evolution in both space and time. Among its strengths, DB-HERITAGE includes provision for open-source and free access for multiple users provided by common policies, but with specific control of different types of accounts and clearances. It is also flexible, scalable, and easily adaptable to user needs.

DB-HERITAGE includes open-source tools for recording and obtaining free access to the data on the history, properties, and performance of building materials used in Portuguese built heritage all over the world. It also displays and provides an easy management system for different physical repositories of samples of building materials used in different periods from prehistory to the present.

While the concept behind the DB-HERITAGE project and system is simple, its implementation is intricate given the different requirements established by the targeted stakeholders and end-user groups. It has faced major challenges in encouraging stakeholders and researchers to curate, enrich and preserve research and technical data and material samples, as well as to approve a common data management plan and to adopt common standards, practices, FAIR principles, and policies. It has also faced challenges related to acknowledging the importance of data sharing by the community and implementing a more effective value chain. Given the current geopolitical, social, and economic contexts, funding the medium and long-term stewardship of digital and material resources is also a primary challenge.

During an upgrade process, it has been recognised that there is a need to advance standards for interoperable data from which shared information can easily be extracted and used. Data management improvement, as well as data aggregation into the ARIADNEplus infrastructure, required the use of metadata standards, vocabularies, ontologies, and other common principles and practices. Thus, as described here, the aggregation process brought significant benefits and further opportunities to improve DB-HERITAGE.

2. Joint Developments in ARIADNEplus

Participating in ARIADNEplus helped DB-HERITAGE deal with most of its major challenges. Data management strategy and related best practices, including the FAIRness of data, have been addressed in ARIADNEplus (e.g. Richards 2023). Standardised data management procedures, including the use of relevant metadata standards, vocabularies, ontologies, and other common principles and practices, have been provided, adopted, and further developed by ARIADNEplus and its participating partners for different major areas of archaeological research. DB-HERITAGE has benefited from the common workflow, which has created opportunities for further improving the DB-HERITAGE data management plan.

ARIADNEplus led to increasing awareness of the benefits of data-sharing and enhancing research and management policies and practices for built heritage assets. Experts from different domains could gather their requirements for achieving a more effective integrated process throughout the value chain (e.g. processing, semantics, integration, and analytics), thus providing FAIR access to archaeological datasets.

The large size of the research community committed to consolidating and extending the ARIADNE infrastructure resulted in major advancements and a wider sharing of information, knowledge, and technologies across different disciplines and communities.

LNEC experts have been mainly involved in the work on standing structures and analytical data, collaborating to establish the required Application Profiles (AP). Data and practices available at DB-HERITAGE were used as a case study for discussing a simplified AP for data on standing structures, which was afterwards used for mapping DB-HERITAGE data to the ARIADNEplus ontology. Analytical and structured data on construction materials' history and properties from DB-HERITAGE also contributed to establish the requirements of the AP for the study of inorganic materials.

Following the ARIADNE guidelines, the DB-HERITAGE vocabulary terms have been matched to the Getty AAT using the ARIADNEplus Vocabulary Matching Tool.

Additionally, Portuguese terms and scope notes, which will improve multilingual vocabularies and queries in Portuguese, have also been proposed.

3. Aggregation Process and DB-HERITAGE Practice

The DB-HERITAGE streamlined data model includes groups of data structured in four major areas (Figure 2). The built heritage asset and its components link all related data clustered on the main areas: environment characteristics, material properties and history of interventions (Santos Silva *et al*. 2017a). The data also include information on studies and analytical data on the properties and performance of building materials.

Building materials data are specifically systematised by type and function, addressing the primary technical and historical subject matters on materials and construction practices. Thus, the materials are linked to their provenance, which is hierarchically displayed and detailed, including information on the evolution of built assets and of their components in both space and time, and according to their functions and use contexts.

Figure 2: Simplified schema of DB-HERITAGE model with illustrating figures from the Roman archaeological site of Tróia

Before incorporating data in the ARIADNEplus Knowledge base, a simplified application profile was used to create the mappings for first-level data of DB-HERITAGE on standing structures. DB-HERITAGE data attributes addressing the central questions (what, when and where) were mapped to suitable classes and properties of the ARIADNEplus ontology (AO-Cat).

The approach comprised mapping to a common ontology, including using or joining existing attributes of DB-HERITAGE (e.g. landing page, id, subject, title, spatial and temporal coverages, and place name), creating new attributes (e.g. publisher and contributor), searching the archive history (e.g. issued and modified dates, and creator), and adopting agreed semantics (e.g. data type). An example of first-level data from the Roman archaeological site of Tróia, as mapped to AO-Cat, is illustrated in Figure 3. Other examples of DB-HERITAGE datasets aggregated and displayed in the ARIADNEplus infrastructure portal can be found at https://portal.ariadne-infrastructure.eu/resource/d5b7c17df9910eeb6a718dbb6892bd5555f523a2e2e147fd969ea6f6802bda04.

This basic approach made the mapping of the DB-HERITAGE data attributes to suitable AO-Cat classes and properties straightforward. Subsequently, the ARIADNEplus team applied the mapping service X3ML to the DB-HERITAGE first-level data, producing the collection for aggregation. Further procedures included an interactive process of sample data testing, feedback, and data improvement, with the required amendments, also using the Activity Dash online tool and the virtual research environments in the ARIADNEplus Cloud.

To improve the interoperability of DB-HERITAGE data, the original subjects have been matched to the domain-specific vocabulary Getty's AAT, and a description and indexing of the historical periods have been provided using the public domain Gazetteer PeriodO. Improvements to DB-HERITAGE data provenance have also been made to further enable data aggregation and improve reusability.

The process of aggregating the data collection on standing structures available at DB-HERITAGE has been straightforward and successful. Nevertheless, there are much more granular and detailed datasets, whose specificity is especially relevant for managing and preserving the built heritage, that still need to be mapped.

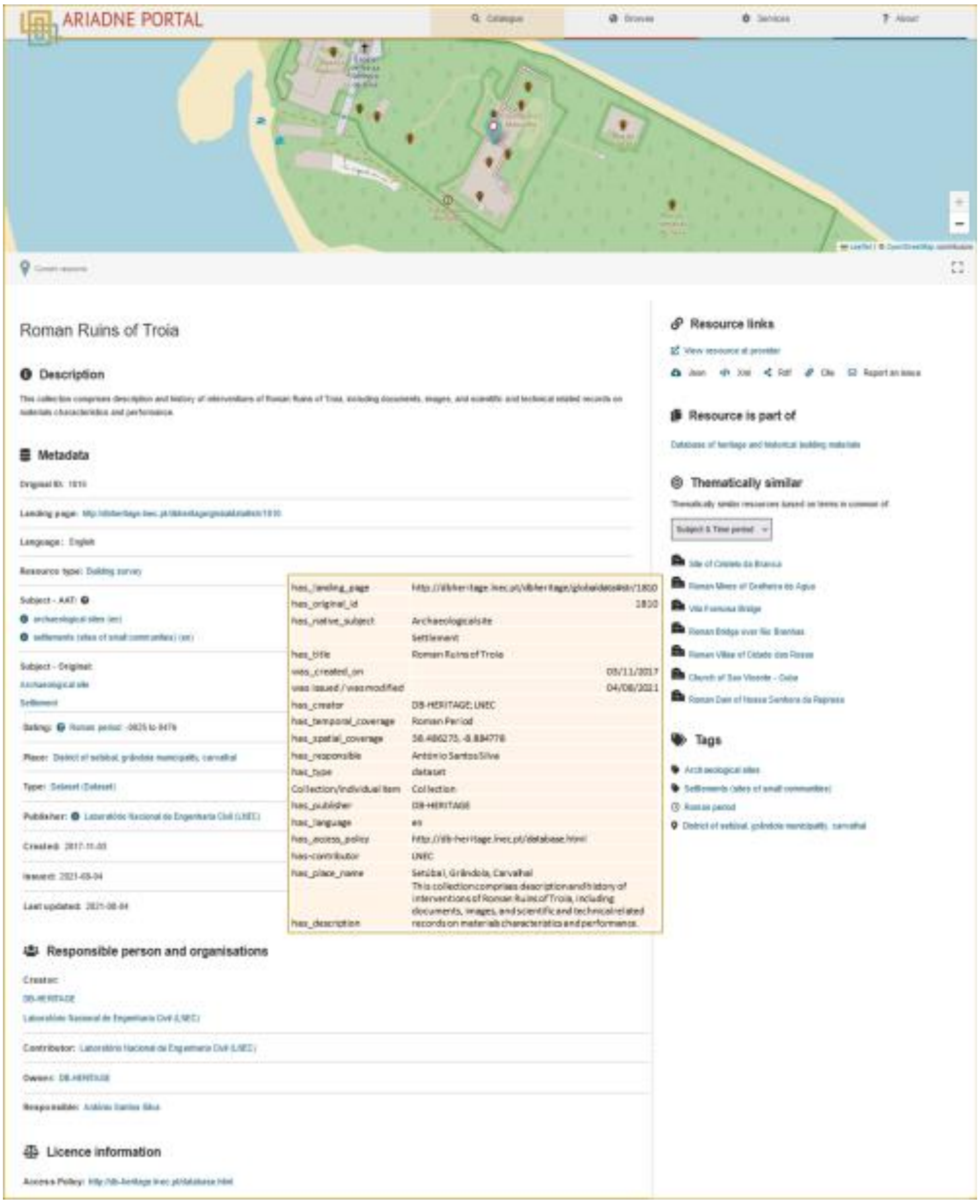

Figure 3: Roman archaeological site of Tróia. Example of first-level data of the collection mapped to AO-Cat and respective dataset accessed in ARIADNEplus portal (https://portal.ariadne-infrastructure.eu/resource/20b5910039c4fff947e506be8483279359772O7dc5376c9ffb80898cd88c7af8)

4. Other Non-aggregated Data of DB-HERITAGE

The full description of standing structures requires a multi- and cross-disciplinary approach, involving the cooperation of different experts and stakeholders. Standing structures are sensitive to their surrounding contexts and related impacts, such as those associated with human actions, environmental exposure, and load-bearing conditions (Fig, 4). These contexts and actions change with time and place, even within the asset and its constituent parts (e.g. microenvironment). Safety, serviceability, and durability requirements should also be considered, especially complying with changes in function and value. Intangible social and cultural (Santos Silva *et al*. 2017b), and material aspects are essential to address the requirements of multiple communities and end-users of the information.

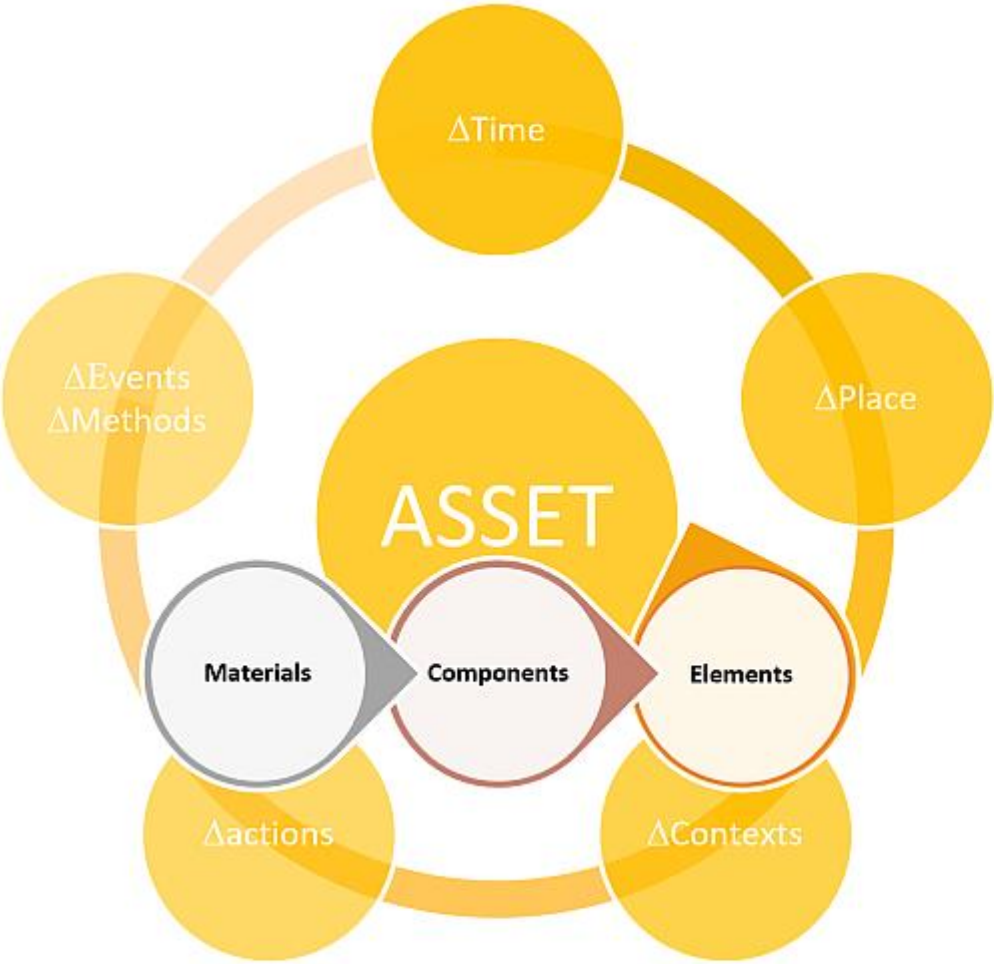

Figure 4: Schematics of the information required for the full description of an asset according to the DB-HERITAGE model

Information on materials (e.g. stone, mortar, concrete, paints, ceramics, metals, wood) and assessment of their state of conservation is also essential for the improvement of current knowledge and for research and innovation, as well as for providing recommendations on materials and techniques to be used for conservation and restoration interventions in built heritage assets.

The detailed description of built heritage assets, including the relationship between their components and their evolution, integrating all the surrounding environment, is very complex. Typically, the information brings together the impact of different events experienced by an asset over time at different scales. Likewise, the specificity of the technical data for each of the various methodologies of observation and analysis adds further complexity to data harvesting.

For example, the Tróia archaeological site comprises various structures in distinct places with different elements and constituents produced for diverse purposes. Through time the surrounding contexts and actions on the asset have changed. Additionally, it has also been affected by different human-made events. The respective data systematisation and integration, as illustrated and outlined in Figures 5 and 6, is therefore essential for a comprehensive interpretation of the asset performance.

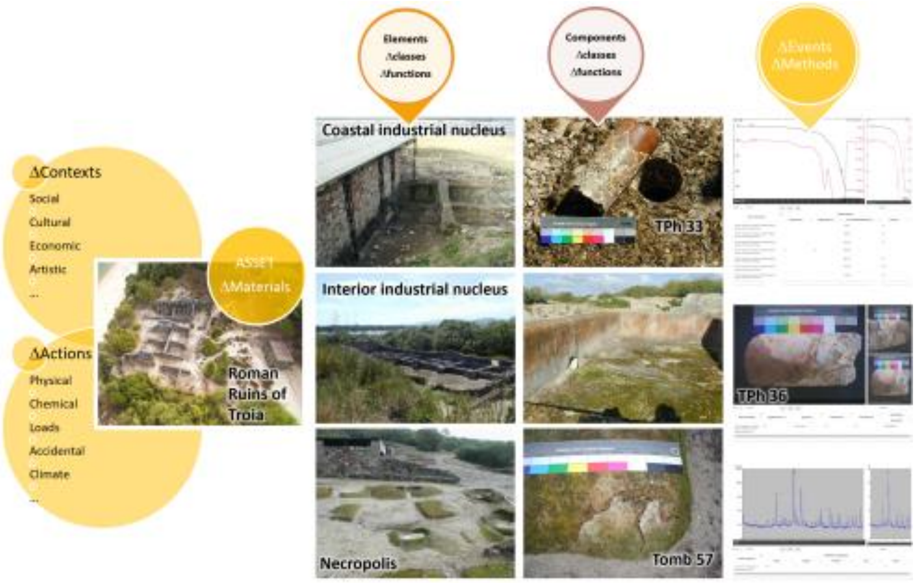

Figure 5: Roman archaeological site of Tróia. Outline of data systematisation illustrated with examples taken from an inspection event
(see http://dbheritage.lnec.pt/dbheritage/historicaldata#/event/55/str/1810)

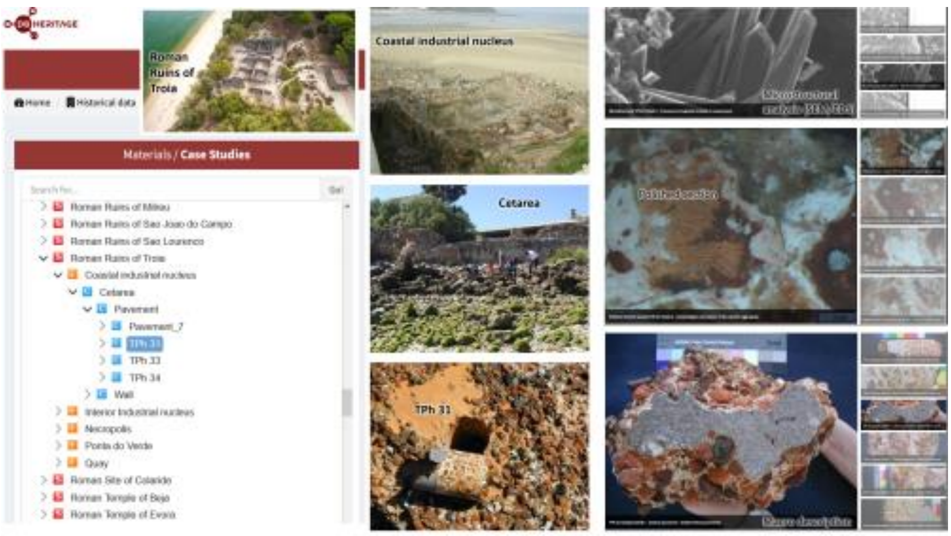

Figure 6: Roman archaeological site of Tróia. Example of analytical data of specimens obtained from a sample (TPh 31) of the pavement of a cetarea (salting tank)
(see http://dbheritage.lnec.pt/dbheritage/historicaldata#/event/55/str/2441)

The primary reasons why these data were not integrated relate to the complexity of the information mesh necessary for data quality enhancement, including details on provenance, changes through space and time, as well as the effects of the multiple contexts (e.g. climate, social, and cultural). DB-HERITAGE data on events (e.g. conservation, inspection, monitoring) could also have been aggregated, eventually with some semantic adaptations. However, the complete assessment of the complex association of data and its evolution in time and space, also depending on the context, has been prioritised.

Specific concepts of the CIDOC CRM and its extensions (e.g, CIDOC CRMba and CRMhs) are required to address all the requirements for the item-level description of built structures and their components, including building materials.
The CRMba provides valuable concepts for standing structures composed of various entities providing topological relations between them. The entities and properties for describing scientific activity (CRMsci) related to heritage science (CRMhs), as further discussed, and developed within ARIADNEplus, can also be used to attain item-level aggregation of building materials data. Other technologies and standard workflows, such as HBIM, should also be considered to provide added value to the information.

5. Benefits and Opportunities Provided by ARIADNEplus

ARIADNEplus has not only produced an integrated data infrastructure, but it has also established a framework for innovation in the methodology and practices of data-based research in archaeology. The association of common ideas, methods, and tools, and the provision of free services, primarily focusing on user needs, integrates a shared strategy concerning better use of resources and cost-savings. Reduced activities for more effective data management, through the value chain, also represent a potential increase in productivity.

The adopted data harmonisation strategy for providing FAIR access to resources may also facilitate multidisciplinary enrichments, broader sharing of information, improved knowledge, and technology transfer across different fields of activity. The use of information beyond research is also expected, for example by contributing to knowledge-based decision-making in asset management and evidence-based policy making.

Through the integration of data in the research infrastructure ARIADNEplus, DB-HERITAGE has contributed to an increase In FAIR data, improving data storage, processing, and sharing capacity beyond the Portuguese context. Via ARIADNEplus, DB-HERITAGE is now aligned with European digital strategy and has reinforced the recommendations for preserving and sharing archaeological and architectural built heritage data.

Extra opportunities for consolidating comprehensive, active, and informed data-sharing communities have provided a strong foundation for encouraging citizens and researchers to curate, preserve and share data on built heritage. The networking and partnerships developed within ARIADNEplus and FAIR data are also envisaged as enhanced opportunities for long-term support and further upgrades of DB-HERITAGE. Savings should also be produced by improving the DB-HERITAGE system research tools and data reuse.

6. Predicted Future Actions and Impacts

DB-HERITAGE aims to improve its contribution to an increase in the storage capacity, processing and sharing of interoperable data, clearly aligning its aims with European strategy and with worldwide recommendations for the preservation and sharing of built architectural and archaeological heritage data. The anticipated impacts range across diverse sectors, from society to industry, bringing changes in the human way of living and in the use of resources. Envisaging contributing to major changes, preserving the past while innovating, DB-HERITAGE plans to be actively involved in different actions in the near future as it is today.

DB-HERITAGE will search for new opportunities for expanding the implemented system, improving its related resources, extending its application domains to new user communities and stakeholder groups, and facilitating the sharing and reuse of data on archaeological and architectural heritage. Progress is also expected on improving communication and promoting the active and inclusive participation of society in both gathering and sharing data. Future actions will also focus on answering the current challenges of analysis in order to preserve the historical, archaeological, and architectural heritage while enhancing its impacts on present and future generations.

Acknowledgements

The authors acknowledge ARIADNEplus (H2020 Programme, contract no. H2020-INFRAIA-2018-1-823914), the National Laboratory of Civil Engineering (LNEC) and Fundação para a Ciência e Tecnologia (FCT - project PTDC/EPH-PAT/4684/2014) for their support.

Bibliography

Correia, M.J., Santos Silva, A., Neto, F. and Martins, M. 2019 'Portuguese cultural heritage data access tools and policies – anticipated impacts of ARIADNEplus' in J. Richards and F. Niccolucci (eds) *The ARIADNE Impact*, Budapest: Archaeolingua Foundation. 107-112. https://zenodo.org/record/4319058

Richards, J.D. 2023 'Joined up Thinking: Aggregating archaeological datasets at an international scale', *Internet Archaeology* **64**. https://doi.org/10.11141/ia.64.3

Santos Silva, A., Correia, M.J., Rato, L. and Matos, N. 2017a 'Building materials with historical and heritage interest – a database in progress' in *PROHITECH'17. 3rd International Conference on Protection of Historical Constructions, Lisbon, Portugal, 12 – 15 July, 2017*, Lisbon.

Santos Silva, A., Menezes, M. and Correia, M.J. 2017b 'The archive of historical materials in the digital era as cultural value' in M. Menezes, D. Costa and J.D. Rodrigues (eds) *IMaTTe2017- Intangibility Matters - International conference on the values of tangible heritage*, Lisbon, May 29-30, 2017, LNEC: Lisboa. 407-412. http://repositorio.lnec.pt:8080/jspui/handle/123456789/1009617

Santos Silva, A., Paiva, M., Ricardo, J., Salta, M., Monteiro, A.M. and Candeias, A.E. 2006 'Characterisation of roman mortars from the archaeological site of Tróia (Portugal)', *Materials Science Forum* **514-516**, 1643-1647. https://doi.org/10.4028/www.scientific.net/MSF.514-516.1643.