



Project Gallery

Unlocking archaeological data online via the PEPAdb (Prehistoric Europe's Personal Adornment Database) initiative for Open Science

Galo Romero-García^{1,*} , Daniel Sánchez-Gómez⁴, José Á. Garrido-Cordero⁴, José M. Martínez-Blanes^{2,3}, Ana C. Sousa⁴ & Carlos P. Odriozola^{1,4}

¹ Department of Prehistory and Archaeology, University of Sevilla, Spain

² Department of Inorganic Chemistry, University of Sevilla, Spain

³ Institute for Material Science of Sevilla, Joint Centre of the University of Sevilla and Consejo Superior de Investigaciones Científicas (CSIC), Spain

⁴ Centro de Arqueologia da Universidade de Lisboa (UNIARQ), University of Lisbon, Portugal

* Author for correspondence ✉ galorom@us.es

PEPAdb (Prehistoric Europe's Personal Adornment Database) is a long-term, open-ended project that aims to improve access to archaeological data online. Its website (<https://pepadb.us.es>) publishes and analyses datasets about prehistoric personal adornment, drawing on the results of various research projects and bibliographic references.

Keywords: data analysis, Open Science, FAIR principles, personal adornment, Mediterranean Europe

Introduction

Access to publications and datasets generated by publicly funded R&D projects, either raw or processed, is limited for many scientists and citizens. This has caused a movement in academia toward the use of Open Science—defined by the Organisation for Economic Co-operation and Development (OECD) in 2015 as a way to make data and the methods used to analyse them openly available to professionals and public stakeholders to democratise science and enhance the reproducibility of research (Marwick *et al.* 2017; Robinson *et al.* 2019). The alignment of governmental policies with the Open Data Charter (Brandusescu & Lämmerhirt 2019)—signed by G8 leaders to promote transparency, innovation and accountability—and implementation of FAIR principles (Findable, Accessible, Interoperable and Reusable data) intend to make data accessible and reusable. In this way, Open Science has become a priority for the European Commission and national governments. FAIR and open-data sharing have become mandatory when publishing publicly funded scientific research results.

In light of these new policies, there has been a recent proliferation of Open Access publications that are available for free online. This move towards Open Access publications must

Received: 21 August 2023; Revised: 15 November 2023; Accepted: 28 November 2023

© The Author(s), 2024. Published by Cambridge University Press on behalf of Antiquity Publications Ltd. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (<https://creativecommons.org/licenses/by-nc-sa/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is included and the original work is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use.

also be made towards Open Access research data, according to government policies and best practice in the sector (Andreoli-Versbach & Mueller-Langer 2014).

In this context, archaeology is currently at a crossroads. Archaeology generates huge amounts of spatially indexed and non-indexed datasets, which in most cases cannot be published in traditional formats (e.g. journal articles, book chapters, conference proceedings) owing to the structure of the dataset and, therefore, they are not fully reusable because the raw data are not accessible, which compromises reproducibility.

Data sharing, both in terms of provision and access, has become a notorious problem in archaeological research (Sobotkova 2018). The Prehistoric Europe's Personal Adornment Data Base (PEPAdb) project aims to address this situation for research into its subject matter from across Europe by making its datasets available through web technologies, which are configured to provide feature-rich applications, to scholars and citizens.

The background for PEPAdb

PEPAdb is a multidisciplinary, open-ended and ongoing online project carried out by the QUANTA²S (Quantitative Archaeology and Archaeological Science Research Group). The project's data platform and webapps (available at <https://pepadb.us.es>) are hosted by the Department of Prehistory and Archaeology of the University of Seville and UNIARQ (Centro de Arqueologia da Universidade de Lisboa). Initially, in March 2013, PEPAdb aimed to enhance understanding of the geographic origin and spatiotemporal distribution patterns of greenstones used for personal adornment during the Neolithic and Bronze Ages (sixth to second millennia BC) within Mediterranean Europe.

In the past five years, the project has evolved and broadened its focus to encompass the geographic origin and spatiotemporal distribution patterns of diverse personal adornments—including beads, pendants, charms—of any colour and its scope now also includes the study of amber and coated beads. This expanded investigation involves a comprehensive inventory and characterisation of raw materials, detailing their geographic origins and the technologies involved in their production.

During the past 10 years, PEPAdb has been funded by the following agencies and budget programmes: Ministry of Economy and Competitiveness (MINECO); Ministry of Science and Innovation (MICIN); Fundação para a Ciência e a Tecnologia (FCT); Plan Andaluz de Desarrollo e Innovación, Junta de Andalucía (PAIDI); Juan de la Cierva Program of the MICIN, University Teacher Training Program (FPU); Plan for Transfer and Innovation of the University of Seville (PPTIUS); Youth Employment Plan of the University of Seville (PEJUS). It has a budget of more than €1.2 million that is allocated to personnel, research infrastructures and R&D expenses.

Over the past 10 years, the results of PEPAdb have been published in different media—including journal articles, book chapters, proceedings—and the processed PEPAdb datasets are partially accessible through these media. However, the responsibility and belief in Open Science motivates the project's move towards making all its datasets available through web technologies.

PEPAdb serves as the repository for over 90 000 records of personal ornaments from late European prehistory, with this number continually expanding—21 656 records were made

available on 1 August 2023 and the complete database will be available by the end of March 2024. The present coverage spans Mediterranean Europe and countries in Western Asia. The dataset within PEPAdb includes: a) information extracted from bibliographic resources and museum inventories—focusing on resolving contradictions and ambiguities between bibliographic references and museum inventories through a comprehensive cross-check of these sources; b) details related to the elemental and mineralogical composition of the analysed pieces; c) information concerning the elemental and mineral composition of the mineral sources used for provenance analysis; and d) spatial information linked to each of the records in the dataset.

Therefore, PEPAdb comprises a significant amount of spatially indexed data of various types (spatial, experimental, modelled) and formats (WMS, WFS, json, csv). The authors aim to make these data freely available, following the Open Science approach, ensuring easy discovery, access, indexing and reuse. Researchers have the capability to upload and work on their own data within the ‘cartographic viewer’, which is an interactive map on the website. Moreover, we are working on integrating machine-learning-based apps—such as MACLAS (a supervised multi-class framework for mineral classification of Iberian beads) and VORTEX (a supervised multi-class framework for provenance classification of variscite beads)—to enhance the website’s functionalities by the second quarter of 2024.

PEPAdb as an Open Science initiative

The PEPAdb Open Science initiative aims to fulfil the FAIR principles through the development of a Spatial Data Infrastructure (SDI) that complies with the EU Directive on open data and reuse of public sector information (Directive 2019/1024) and the FAIR policy. This SDI will allow the general public, the scientific community or policy makers to access and reuse the data and information recorded in the project for their own benefit.

The combination of spatial databases and GIS (Geographic Information Systems) software has opened the possibility of creating archaeological web tools and services for sharing indexed spatial data; however, this is not yet commonplace. PEPAdb intends to make its datasets publicly available through a web tool designed to model, analyse, visualise or generate new geospatial data, information and value-added resources—such as thematic cartographies—but also raw data and empirical datasets through an online cartographic viewer and a database-query application.

PEPAdb ensures its sustainability through a robust data-management model, incorporating regular updates and maintenance so it can accommodate evolving technologies. Documentation on methodologies and standards will assist future researchers. The web tool, along with all the data, operates under a CC4 licence, making it openly available and fostering cross-sector working. Researchers will benefit from the tool’s detailed instructions and support channels to enable seamless integration into their projects, promoting an open and collaborative research environment.

Putting FAIR into practice

Open Science has been considered in the design of the data-collection strategy of all the funding programmes described above. The focus is on creating a curated dataset that can be reused beyond the individual projects and ensure that all published work is reproducible.

PEPAdb website description

PEPAdb is designed to provide users with the ability to view, query and download datasets in a responsive web design. This tool has been developed to serve as an archaeological information resource for the scientific community. It includes:

- 1) A cartographic viewer (Figure 1) showing the frequency of minerals forming beads at each archaeological site and/or structure.
- 2) INSPIRE (a European Union directive that aimed to enhance the sharing of environmental spatial information among public sector organisations and better facilitate public access to environmental information across Europe) standard web services, such as WMS and WFS (Figure 2).

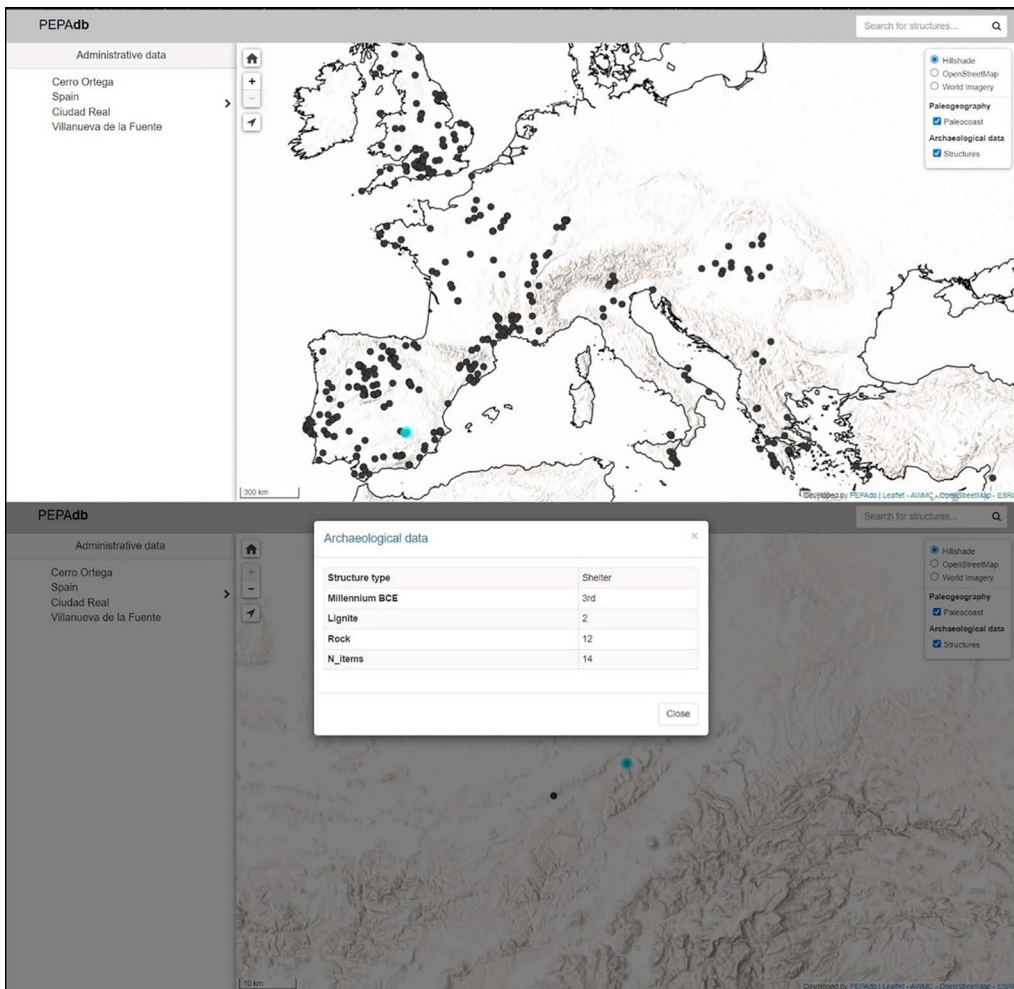


Figure 1. Screenshot of cartographic viewer showing the frequency of minerals forming beads at each archaeological site and/or structure (figure by authors).

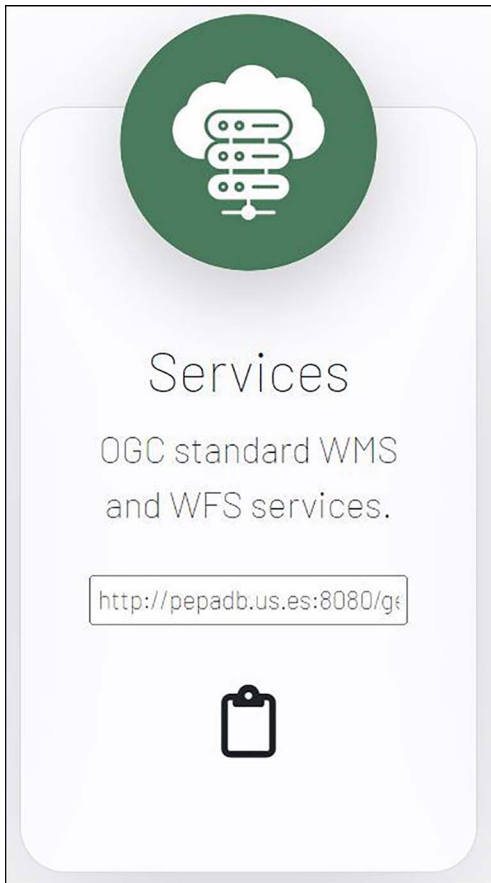


Figure 2. Screenshot of INSPIRE standard web services showing access options (figure by authors).

- 3) Accessible and downloadable compositional, mineralogical and metric data, allowing users to easily retrieve specific information of interest.
- 4) A csv file with all the project's raw data.

Regarding the online database (Figure 3), functionalities for consulting, filtering and downloading alphanumeric information in standard file formats (.json, .csv, .xml, .txt, .sql and .xls) are available to users. In relation to the FAIR principle of reusability, it is the responsibility of researchers to specify whether or not the downloadable dataset is raw or processed. For this reason, we include a section called 'raw data', which provides a dataset from which both spatial and non-spatial information (processed data) has been generated.

Conclusion and next steps

To build a robust analysis/sharing tool for the scientific and policy-making

Archaeological structures	Country	Administration (1st level)	Administration (2nd level)	Structure type	Millennium BCE	Item type	Raw m	JSON	XML	CSV	TXT	SQL	MS-Excel	ems
<input checked="" type="checkbox"/>	Adabrock	Scotland	Na h-Eileanan Siar	Ness (isle of Lewis)	Barrow	Undefined	Bead	Am						
<input checked="" type="checkbox"/>	Alalakh	Turkey	Hatay Province	Varsili		2nd	Undefined	Am						
<input checked="" type="checkbox"/>	Alicar 3	Portugal	Faro	Portimao - Mexilhoeira Grande	Tholos	3rd	Bead	Am						
<input type="checkbox"/>	Alicar 4	Portugal	Faro	Portimao - Mexilhoeira Grande	Tholos	3rd	Bead	Amber						1
<input type="checkbox"/>	Aldea del Terrifuelo	Spain	Salamanca	Aldeavejia de Tormes	Passage grave	4th-3rd	Bead	Lignite						2
<input type="checkbox"/>	Aldea del Terrifuelo	Spain	Salamanca	Aldeavejia de Tormes	Passage grave	4th-3rd	Bead	Rock						1475
<input type="checkbox"/>	Aldeagordillo	Spain	Ávila	Ávila	Domestic	3rd	Bead	Rock						16
<input type="checkbox"/>	Alto da Feteira	Portugal	Leria	Pombal - Arroiteia	Dolmen	3rd	Bead	Shell						22
<input type="checkbox"/>	Alto del Reinoso	Spain	Burgos	Monasterio de Rodilla	Chamber	4th	Bead	Lignite						2
<input type="checkbox"/>	Alto del Reinoso	Spain	Burgos	Monasterio de Rodilla	Chamber	4th	Bead	Rock						19
<input type="checkbox"/>	Ambérieu-en-Bugey	France	Ain	Ambérieu-en-Bugey	Undefined	Undefined	Bead	Amber						6
<input type="checkbox"/>	Ambérieu-en-Bugey	France	Ain	Ambérieu-en-Bugey	Undefined	Undefined	Spacer	Amber						1
<input type="checkbox"/>	Anafisi	Greece	Heraklion	Herossosinos		2nd	Undefined	Amber						3
<input type="checkbox"/>	Ancecy	France	Haute-Savoie	Ancecy		3rd	Bead	Amber						1
<input type="checkbox"/>	Anta Grande do Zambujeiro	Portugal	Évora	Évora - Nossa Senhora da Tourega e Nossa Senhora de Guadalupe	Dolmen	3rd	Bead	Amber						168

Figure 3. The database allows filtering and downloading in various standard file formats (.json, .csv, .xml, .txt, .sql and .xls) (figure by authors).

community, additional development is required. Although we have already introduced the cartographic viewer, an online database and utilised Open Geospatial Consortium services, we aim to enhance the website further. The functionalities due to be implemented by 2025, such as MACLAS and VORTEX, will significantly augment the website's value. We plan to integrate analysis and dissemination tools, enabling researchers to work with their datasets alongside the PEPAdb dataset.

Funding statement

Funding for this work is provided through research project 'PID2021-124421NB-I00' (https://investigacion.us.es/sisius/sis_proyecto.php?idproy=36407).

References

- ANDREOLI-VERSBACH, P. & F. MUELLER-LANGER. 2014. Open access to data: an ideal professed but not practised. *Research Policy* 43: 1621–33. <https://doi.org/10.1016/j.respol.2014.04.008>
- BRANDUSESCU, A. & D. LÄMMERHIRT. 2019. Open Data Charter Measurement Guide. <http://dx.doi.org/10.2139/ssrn.3331162>
- MARWICK, B. *et al.* 2017. Open Science in Archaeology. Preprint. SocArXiv. Available at: <https://osf.io/72n8g> (accessed November 2023). <https://doi.org/10.31235/osf.io/72n8g>
- ROBINSON, E., C. NICHOLSON & R.L. KELLY. 2019. The importance of spatial data to open-access national archaeological databases and the development of paleodemography research. *Advances in Archaeological Practice* 7: 395–408. <https://doi.org/10.1017/aap.2019.29>
- SOBOTKOVA, A. 2018. Sociotechnical obstacles to archaeological data reuse. *Advances in Archaeological Practice* 6: 117–24. <https://doi.org/10.1017/aap.2017.37>