ARBOREA – A new georeferenced database for plant macro-remains from archaeological sites in Central Italy

Claudia Moricca[1], Erasmo Di Fonso[2], Rachele Nicolini[1], Laura Sadori[1]

[1] Department of Environmental Biology, Sapienza University of Rome, P.le Aldo Moro 5, 00185 Rome, Italy, claudia.moricca@uniroma1.it, rachele.nicolini@uniroma1.it, laura.sadori@uniroma1.it
[2] Department "Italian Institute of Oriental Studies – ISO", Sapienza University of Rome, Circonvallazione Tiburtina 4, 00185 Rome, Italy
erasmo.difonso@uniroma1.it

Abstract – **ARBOREA (ARchaeoBOtanical geoReferenced rEcords of Ancient Italy) is a new digital tool designed to compile, harmonize, and georeference archaeobotanical data from Central Italy. Developed within the NRRP PE5 CHANGES Spoke 8 and aligned with the BRAIN database, ARBOREA addresses key issues of data dispersion and terminological inconsistency. The open-access database integrates information from over 150 sites and 1400 contexts, cataloguing more than 7000 samples and nearly 500 plant taxa. Structured relationally and supported by dedicated thesauri for plant nomenclature, parts, preservation types, and chronology (from prehistory to the modern era), it enables advanced, multi-factor queries and comparative analyses. ARBOREA enhances data accessibility and fosters interdisciplinary collaboration, offering a robust platform for reconstructing human-environment interactions across time.**

I. INTRODUCTION

Recent progress in archaeobotany, the discipline responsible for the study of plant remains from contexts affected by human presence [1], has greatly helped acquire information about human-environment interactions in Italy.

Although archaeobotanical research has become more widespread, access to the resulting data remains limited due to its dispersion across a wide array of scientific publications and often inaccessible archaeological reports. This fragmentation, coupled with the frequent focus on specific sites or isolated findings, significantly hinders the development of a comprehensive and diachronic historical reconstruction.

Established in 2015, the Botanical Record of Archaeobotany Italian Network (BRAIN; https://brainplants.successoterra.it) represents a key development in the systematic collection and dissemination of archaeobotanical data across Italy [2][3]. This collaborative platform provides a rigorously structured and continuously updated database that catalogues archaeological sites investigated for botanical remains, detailing their geographical coordinates, chronological range, cultural context, and categories of plant evidence, all supported by complete bibliographic documentation. The BRAIN database serves as a fundamental tool for enhancing data accessibility and interdisciplinary research within the archaeobotanical domain.

Developed within the framework of the National Recovery and Resilience Plan (NRRP) PE5 "CHANGES: Cultural Heritage Innovation for Next-Gen Sustainable Society" Spoke 8, this study presents ARBOREA (ARchaeoBOtanical geoReferenced rEcords of Ancient Italy), a newly designed georeferenced tool aimed at compiling, standardizing, and comparing data on plant macro-remains (seeds/fruits, wood). The tool harmonizes diverse datasets and organizes them according to chronological criteria.

II. MATERIALS AND METHODS

A. Study area

For our project, we define Central Italy as our study area. Although Eurostat [4] classifies Central Italy as comprising only Tuscany, Marche, Umbria, and Latium, we have chosen to also include Abruzzo and Molise. This decision was made to maintain consistency with the BRAIN database and to facilitate potential future interoperability between datasets. Nonetheless, the project is designed to be progressively expandable to cover the entire national territory.

B. Bibliographic research

We then collected published archaeobotanical studies on plant macro-remains from our study area. We began by identifying relevant publications listed in the BRAIN database that matched our criteria (plant macro-remains, Central Italy). Since BRAIN does not provide access to full texts, we obtained the publications through various channels. Open access articles or those available via journals with agreements with our university were downloaded directly from the publishers' websites. For the remaining publications, we used three methods: contacting the authors, consulting physical copies in libraries, or requesting PDFs through NILDE (Network Inter-Library Document Exchange), a web-based service for document supply and interlibrary loans (https://nildeworld.bo.cnr.it/).

C. *Data extrapolation*

One of the aspects we consider fundamental is the inclusion of bibliographic references for each data point we extracted, to properly credit the researchers responsible for the identification and interpretation of the material. With a view to interoperability and connection with the BRAIN database, we have included the associated BRAIN code [3] for each entry, which identifies the corresponding site (e.g., CLA26 – Santi Quattro Coronati Complex, Rome).

Data extrapolation was performed following the fields presented in section II D - Database structure.

D. *Database structure*

Data were organized in a relational database designed to manage hierarchical layers of information and the relationships linking them. The archaeological and archaeobotanical information was primarily divided into a context-sample structure, with the context having an internal hierarchic structure. This partitioning of information made it possible to enhance and systematize the data associated with both samples and contexts, thus increasing the analytical and comparative potential of the dataset. To further support the harmonization of data and improvement of data entry efficiency, two *thesauri* were added to this main corpus, both designed to fully leverage structured data in their respective domains: one for the management of plant *taxa*, the other for chronological/archaeological phases. The entire database was designed in a GIS platform to seamlessly integrate geographic data, a key element for achieving the objectives of the project (Fig. 1).
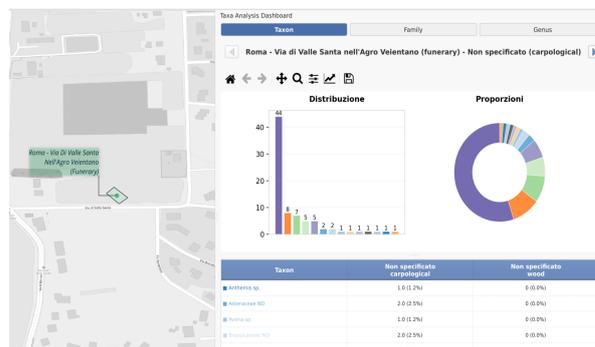

Fig. 1. Example of a context (Necropolis of Valle Santa nell'Agro Veientano [5]) showcasing the GIS database's internal comparative capabilities between sub-contexts.

The custom 'switch' tool also enables automatic and dynamic grouping across different levels of information.

E. *Botanical nomenclature*

Due to, among others, the advancement in molecular biology and the acquisition of new morphological data, numerous plants found in archaeobotanical assemblages have changed names in the last years/decades. For this reason, comparing data from different publications may not always be straightforward. To address this issue, we developed a dedicated *thesaurus*, adopting the nomenclature proposed by the Portal of Italian Flora (http:/dryades.units.it/floritaly) [6][7] as our reference system. Nevertheless, we also included a field to retain the original terminology as reported in the source publication.

F. *Plant parts and modalities of preservation*

In developing our database, we also standardized the terminology used to describe plant parts (e.g., seeds, fruits, teguments, etc.) recovered from archaeological contexts. The terminology varies significantly depending on factors such as plant family, preservation method, and the descriptions adopted by individual authors. To address this, we established a controlled vocabulary, primarily based on the classification proposed by Neef et al. [8].

We also harmonized the categorization of preservation methods, which was essential for the consistency of the database structure. Accordingly, for each specimen (when specified in the original publication), we indicate the type of preservation - whether through charring (Fig. 2a), mineralization (Fig. 2b), waterlogging (Fig. 2c), desiccation (Fig. 2d), impression, or a co-occurrence of these processes.
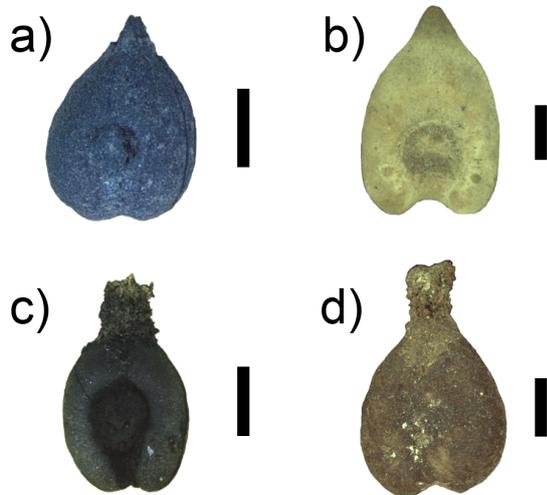
Fig. 2. Vitis vinifera L. seeds from archaeological sites in Italy preserved by: a) charring; b) mineralization; c) waterlogging; d) desiccation. Scale bar = 2 mm.

G. *Chronology*

The dataset includes a broad chronological range, from the Neolithic to the early modern period. To enhance the search efficiency and comparison of the data, a dedicated *thesaurus* was developed, allowing a better categorization of information. This approach has made it possible to effectively manage the complex relationships among phases, sub-phases, and archaeological *facies*, preserving the original published information, and optimizing the attribution of multi-phase classifications for each context. For the disambiguation of absolute chronological terms, reference was made - wherever possible - to the online repository PeriodO - A gazetteer of periods for linking and visualizing data (https://perio.do/en/)[9], while the standardization of labels assigned to the various chronological spans followed ICCD standards, where applicable.

III. RESULTS

We have applied the proposed protocol and built the database considering over 150 archaeological sites, for a total of approximately 1400 isolated contexts. These allowed us to catalogue more than 7000 samples, belonging to almost 500 different taxa (Fig. 3).

The developed dictionaries cover a chronological period of over six millennia (with the possibility of inserting extending this range further back), with a depth of three distinct levels of hierarchical information.
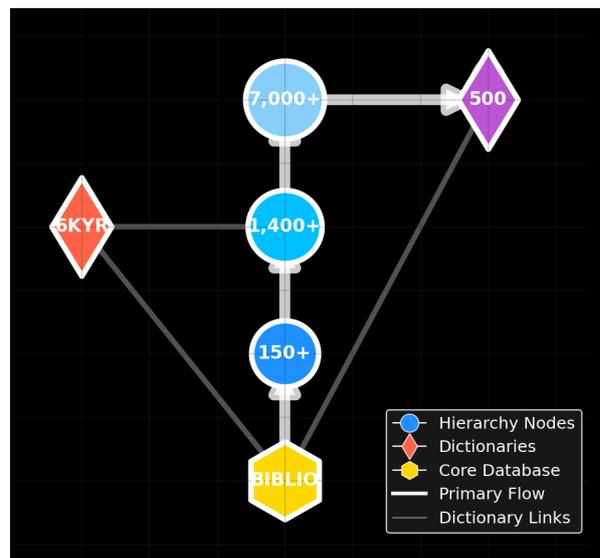

Fig. 3. Simplified ARBOREA database structure

IV. DISCUSSION

The ARBOREA project is conceived as a central hub that responds to the pressing need for such a tool within the Italian research landscape. Working in harmony with the existing BRAIN project, this open-access database will soon provide an optimized foundation for online querying and researching archaeobotanical data in central Italy, using the relevant bibliography. At the same time, it will offer a framework for retrieving optimized raw data and making available dedicated tools that will enable users to compare data in real time through multi-factor queries. This framework will provide users with advanced instruments not only to interact with data through complex and customizable queries, but also to perform in-depth analyses, explore information across multiple layers of detail, refine research approaches dynamically, and gain a more immediate and coherent understanding of the available data.

Although we firmly believe it is essential to consult the full texts of publications authored by experts in the field, the presented tool will facilitate the work of archaeobotanists in interpreting their unpublished data. They will be able to easily compare the occurrence of certain taxa in similar archaeological contexts and nearby geographic locations. ARBOREA also opens the possibility of review articles, for example concerning the introduction of archaeophytes (e.g. *Prunus persica* (L.) Batsch - peach [10]) or neophytes (e.g. *Cucurbita pepo* L. - squash [11]) to Italy.

Furthermore, gathering data in a single infrastructure also makes it easier to create collaborations between researchers working in the same field, creating a starting point for multi-site studies.

Archaeobotanists are not the only professional figures that could benefit from our database. In fact, ARBOREA

could be an incipit for multidisciplinary studies, highlighting the potential of archaeobotany in the narration of past populations.

V. CONCLUSION

The ARBOREA project represents a significant advancement in the structuring and accessibility of archaeobotanical data from Central Italy. By compiling, standardizing, and georeferencing a wide range of datasets, ARBOREA addresses long-standing challenges of data fragmentation and terminological inconsistency within the discipline. The integration with *thesauri*, hierarchical structures, and chronological harmonization tools makes this platform not only a repository, but also a dynamic research environment. Designed to be interoperable with existing resources like the BRAIN database, ARBOREA enhances the potential for diachronic, multi-scalar, and interdisciplinary analyses. It supports archaeobotanists in contextualizing new data, facilitates comparative studies across time and space, and fosters collaborative research. Beyond its technical and scientific contributions, ARBOREA lays the groundwork for broader narratives on human-plant interactions, opening new pathways for both academic exploration and public engagement with the environmental histories of ancient Italy.

VI. CITATIONS AND REFERENCES

[1] D.Q. Fuller, L. Lucas, "Archaeobotany", in: C. Smith (ed.), Encyclopedia of Global Archaeology, Springer, New York, USA, 2014.

[2] A.M. Mercuri, E. Clò, J. Zappa, G. Bosi, E. Furia, C. Ricucci, M. Di Lena, F. Camerini, A. Florenzano, "BRAIN – Holocene archeo-data for assessing plant-cultural diversity in Italy and other Mediterranean regions", Scientific Data, vol. 11, No. 1, May 2024, 520.

[3] A. M. Mercuri, E. Allevato, D. Arobba, M. Bandini Mazzanti, G. Bosi, R. Caramiello, E. Castiglioni, M. L. Carra, A. Celant, L. Costantini, G. Di Pasquale, G. Fiorentino, A. Florenzano, M. Guido, M. Marchesini, M. Mariotti Lippi, S. Marvelli, A. Miola, C. Montanari, R. Nisbet, L. Peña-Chocarro, R. Perego, C. Ravazzi, M. Rottoli, L. Sadori, M. Ucchesu, R. Rinaldi, "Pollen and macroremains from Holocene archaeological sites: a dataset for the understanding of the bio-cultural diversity of the Italian landscape", Review of Palaeobotany and Palynology, vol. 218, August 2015, pp. 250-266.

[4] Eurostat, NUTS – Nomenclature of territorial units for statistics. European Commission, 2024. https://ec.europa.eu/eurostat/web/nuts/ [retrieved April 29, 2025].

[5] C. Moricca, A. De Cristofaro, L. Ambrosini, "Archaeobotanical Evidence of Funerary Plant Offerings at the Southern Etrurian Necropolis of "Valle Santa nell'Agro Veientano" (Rome, Italy)", Interdisciplinaria Archaeologica, vol. 14, No. 1, January 2023, pp. 63-70.

[6] F. Bartolucci, L. Peruzzi, G. Galasso, A. Alessandrini, N.M.G. Ardenghi, G. Bacchetta, E. Banfi, G. Barberis, L. Bernardo, D. Bouvet, M. Bovio, G. Calvia, M. Castello, L. Cecchi, E. Del Guacchio, G. Domina, S. Fascetti, L. Gallo, G. Gottschlich, R. Guarino, L. Gubellini, N. Hofmann, M. Iberite, P. Jiménez-Mejías, D. Longo, D. Marchetti, F. Martini, R.R. Masina, P. Medagli, S. Peccenini, F. Prosser, F. Roma-Marzio, L. Rosati, A. Santangelo, A. Scoppola, A. Selvaggi, F. Selvi, A. Soldano, A. Stinca, R.P. Wagensommer, T. Wilham, F. Conti, "A second update to the checklist of the vascular flora native to Italy", Plant Biosystems, vol. 58, No. 2, April 2024, pp. 219-296.

[7] G. Galasso, F. Conti, L. Peruzzi, A. Alessandrini, N.M.G. Ardenghi, G. Bacchetta, E. Banfi, G. Barberis, L. Bernardo, D. Bouvet, M. Bovio, M. Castello, L. Cecchi, E. Del Guacchio, G. Domina, S. Fascetti, L. Gallo, R. Guarino, L. Gubellini, A. Guiggi, N. Hofmann, M. Iberite, P. Jiménez-Mejías, D. Longo, D. Marchetti, F. Martini, R.R. Masin, P. Medagli, C.M. Musarella, S. Peccenini, L. Podda, F. Prosser, F. Roma-Marzio, L. Rosati, A. Santangelo, A. Scoppola, A. Selvaggi, F. Selvi, A. Soldano, A. Stinca, R.P. Wagensommer, T. Wilham, F. Bartolucci, "A second update to the checklist of the vascular flora alien to Italy", Plant Biosystems, vol. 58, No. 2, March 2024, pp. 297-340.

[8] R. Neef, R. T. Cappers, R. M. Bekker, "Digital atlas of economic plants in archaeology", Barkhuis, Groningen, The Netherlands, 2012.

[9] A. Rabinowitz, R. Shaw, R., P. Golden, "Making up for Lost Time: Digital Epigraphy, Chronology, and the PeriodO Project", in: A. De Santis e I. Rossi (eds.), Crossing Experiences in Digital Epigraphy: From Practice to Discipline, De Gruyter Brill, Berlin, 2018.

[10] L. Sadori, E. Allevato, G. Bosi, G. Caneva, E. Castiglioni, A. Celant, G. Di Pasquale, M. Giardini, M. Mazzanti, R. Rinaldi, M. Rottoli, F. Susanna, "The introduction and diffusion of peach in ancient Italy", in: J. P. Morel and A. M. Mercuri (eds.), Plants and culture: seeds of the cultural heritage of Europe, Edipuglia, 2009, pp. 45-61.

[11] F. Alhaique, C. Moricca, L. Barelli, A. Masi, R. Pugliese, L. Sadori, G. Romagnoli, L. Piermartini, L. Brancazi, F. Gabbianelli, G. Chillemi, A. Valentini, "Elite Food Between the Late Middle Ages and Renaissance: Some Case Studies from Latium", Environmental Archaeology, vol. 30, No. 2, January 2021, pp. 89–103.