**Archeologia e Calcolatori**
**Supplemento 13, 2025, 29-36**
**doi 10.19282/ACS.13.2025.03**

# NEW TECHNOLOGIES FOR THE PRESERVATION OF PRE-COLUMBIAN ARCHAEOLOGICAL HERITAGE IN SOUTH AMERICA: THE MAPHSA PROJECT

## 1. Introduction

The archaeological heritage of South America is facing increasing threats due to the expansion of industrial-scale agricultural activities, cattle ranching, infrastructure expansion, illegal wood harvesting, and the current fire emergency plaguing the Amazon and other biomes of the continent. Those threats to the natural environment go hand in hand with the damage to the archaeological heritage. Biomes such as the Amazon, particularly its southern fringes, the Cerrado, spanning a vast expanse of wooded savanna covering approximately 2 million km² in central Brazil, and the montane forests in Colombia, have witnessed accelerated deforestation in recent years, as corroborated by real-time remote sensing observations.

For instance, the Cerrado biome has merely half of its original vegetation preserved due to the expansion of agricultural and other economic activities, accentuating its vulnerability to climate change and calling for immediate surveys and conservation interventions (Siqueira, Peterson 2003; Morandi *et al.* 2020). Similarly, regions covered by montane and Amazon forests in Colombia encounter gaps in data pertaining to the location, nature, and threats confronting archaeological heritage. The Congo micro-basin in the Sierra Nevada de Santa Marta region in Colombia is renowned for its profusion of archaeological heritage, comprising approximately 820 pre-Columbian archaeological features encompassing terraces, burials, retaining walls, pathways, and staircases. Yet, several of these features are inadequately documented and seriously threatened by agricultural expansion, looters, and natural causes (Duran-Izquierdo, Olivero-Verbel 2021).

Documenting such pre-Columbian sites before significant heritage loss, in addition to finding new sites through remote sensing, and increasing knowledge about their distribution and significance, are critical steps in preserving the cultural heritage and in changing attitudes towards their destruction. With these imperatives in mind, the Mapping the Archaeological Pre-Columbian Heritage in South America (MAPHSA[1]) project in its initial stage aims to document the archaeological heritage of about 3.5 million km² of endangered forest and savanna biomes across Brazil and Colombia. This will be accomplished by a combination of archival work (revision of legacy

[1] https://www.upf.edu/web/maphsa.

data in existing databases as well as unpublished work), remote sensing, machine learning, and ground survey. Supported by Arcadia – a charitable fund established by Lisbet Rausing and Peter Baldwin – this initiative represents a collaborative effort led by Universitat Pompeu Fabra, in partnership with Universiteit Leiden, Universidade de Sao Paulo, Universidad del Magdalena, the Brazilian National Institute for Space Research, Instituto Colombiano de Antropología e Historia, and Instituto Catalán de Arqueología Clásica.

## 2. Database development

Part of the goal of MAPHSA, and specifically of its database, is to provide a state of the art, public open access database that provides information on the pre-Columbian archaeological sites found in the target extent, which comprises roughly 3.5 km² of Brazil and Colombia. For those purposes, a common standard previously adopted by other Arcadia-funded projects, the CIDOC-CRM (Bekiari *et al.* 2021), as well as technology that relies on it and supports the previous principles (Arches) were used. Arches is a free, open-source software platform designed specifically for cultural heritage organizations to manage their data (Sheldrick, Zerbini 2017). Developed by the Getty Conservation Institute and World Monuments Fund, Arches provides a flexible and customizable database solution for documenting and inventorying cultural heritage sites, artifacts, and related information.

The integrated database should be capable of incorporating both existing legacy data and new entries from various sources, maintaining coherence and scalability. It should also prioritize open access and adherence to FAIR principles, ensuring transparency and accessibility of deployed technology and data. Ultimately, MAPHSA's database will cater to researchers, domain experts, and public users with an adaptable user experience.

Our proposed schema is designed to effectively organize and manage diverse heritage data. At its core, the schema comprises four main entity types: Heritage Location, Information Source, Project Area, and Condition Assessor (Fig. 1). Heritage Location represents pre-Columbian archaeological sites, Information Source accounts for the origin of data, Project Area delineates geographical boundaries, and Condition Assessor evaluates site conditions.

The Heritage Location Resource Model, based on the CIDOC ontology, offers a structured framework for data organization within the MAPHSA database. This model comprises several branches, each serving a distinct purpose. The Administrative Division branch delineates geographical boundaries, while the Heritage Location Summary branch provides essential site information. Detailed observations on archaeological dimensions are housed in the Archaeological Assessment branch, while Built Component branches link physical structures with construction details. Feature branches capture mostly architectural elements

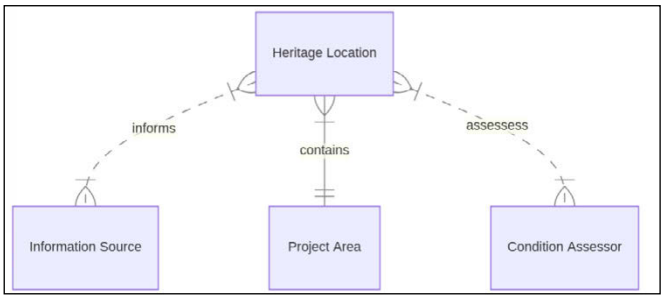


Fig. 1 – Proposed schema for MAPHSA's database.

such as post holes, ovens, hearths and others present in the archaeological sites, and Find branches catalogue small archaeological objects. The Environment Assessment branch offers insights into the site's physical surroundings, while the Condition Assessment branch documents disturbances and recommendations.

To promote standardization and interoperability, controlled vocabularies play a crucial role within the schema. These controlled vocabularies ensure data consistency and facilitate collaboration with other projects following CIDOC ontology standards. Covering various concepts relevant to heritage preservation and management, the employed thesauri contribute to MAPHSA's objective of establishing a robust and interoperable cultural heritage database.

## 3. User interface for data input and editing

To enhance user-friendly data interactions, MAPHSA incorporates QGIS, an open-source geographic information system, into its workflow. QGIS offers a versatile platform for researchers and project staff to interact with the MAPHSA database and is seamlessly interoperable with Postgres/PostGIS.

In the initial phase, QGIS serves as the user interface for data insertion and editing. Upon connection to the Postgres database, tables with their internal relationships are loaded into the software's main interface, with this configuration saved in the QGIS project file. Utilizing QGIS internal field attribute masks, known as widgets, a comprehensive UI for data entry and editing has been developed, embedding database relationships directly into these widgets. Adding an element from the general mask of the her_maphsa table cascades the fields of the macro areas of the model, such as Archaeological Assessment and Environmental Assessment.

The data entry interface prioritizes simplicity and intuitiveness, aiming for usability across all project teams with minimal training. Most fields feature 'Value Maps' and 'Relationship Value' combo boxes, drawing values from the project's vocabulary repertoire. The read-only QGIS project file is stored in

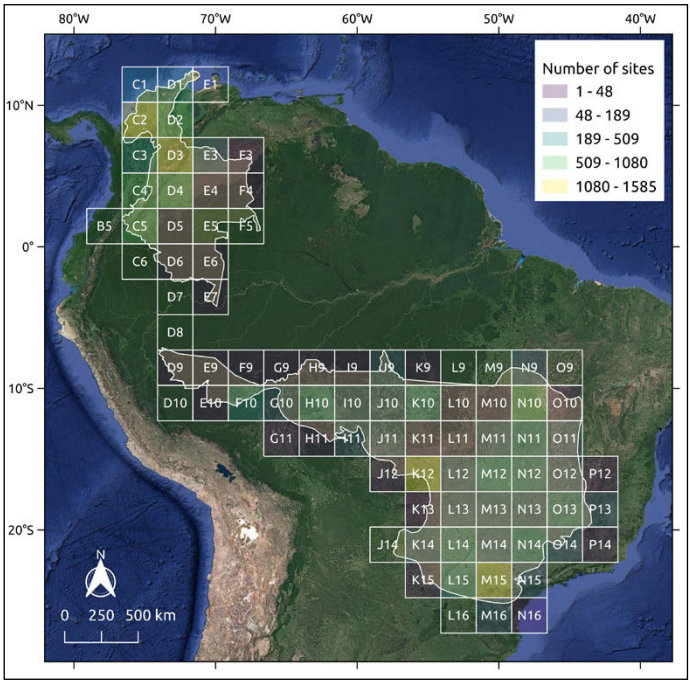


Fig. 2 – Number of legacy database sites in each of the project's grid squares (including pre-Columbian and historical sites).

a schema of the Postgres database, containing not only the UI for data entry but also layer styles for maps and basemap layers. This ensures uniformity in data visualization while allowing customization when specific basemaps or cartography are required.

A first batch of data was obtained from Brazil through SICG – the Integrated System for Knowledge and Management (https://sicg.iphan.gov.br/) which makes publicly available the latest data held by the National Record of Archaeological Sites (CNSA) and the National Institute of Historical and Artistic Heritage (IPHAN). The data, encompassing the entirety of the study area in the Brazil part of the project, contains 9342 archaeological sites. In parallel to the automated ingestion of data from the Brazilian database, an examination of archaeological legacy data in Colombia has been completed in 12 out of the 21 departments encompassed within the project's study area (Fig. 2). The dataset from the study area in Colombia comprises 20377 locations with archaeological findings. The revision process has been completed for archaeological legacy data in both the Caribbean Region and the Eastern Cordillera. These regions are particularly significant, as they have been

identified as areas with frequent pre-Columbian findings. Moreover, these departments are deemed to be at the highest risk for heritage-related events due to the intensity of modern activities within their boundaries.

## 4. Remote site identification

Remote sensing has been successfully applied to identify pre-Columbian earthworks in the forested and savanna regions of South America, changing our perception of the scale of landscape modification in regions that were previously thought to be pristine (Schaan *et al.* 2012; De Souza *et al.* 2018). Furthermore, the remote identification of sites under the forest canopy started to become a reality also through the application of LiDAR technology (Peripato *et al.* 2023). However, the full impact of past human occupation in many forested and savanna regions of South America, and consequently the heritage connected to them, is very much underestimated.

The available data and preliminary experiments suggest that remote, automated identification of archaeological sites is possible even in areas currently covered by vegetation. Modern botanical data point to a pronounced effect of past human management in forest composition throughout the Amazon and Atlantic forests (Levis *et al.* 2012, 2017; Clement *et al.* 2015). This suggests that forest patches with an anthropogenic signal could be distinguished using satellite imagery. As a confirmation of that, maximum entropy (MaxEnt) models suggest that the occurrence of *terra preta* (anthropogenic dark earth) sites can be predicted from remotely-sensed indices such as normalized difference vegetation index (NDVI) (Palace *et al.* 2017). The results show that terra preta sites have less green canopy, lower canopy water content, increased drought susceptibility, lower biomass and lower tree height compared to surrounding forests (Palace *et al.* 2017). Such characteristics of the vegetation serve as an important indicator that will allow adjusting the predictive model for the distribution of archaeological sites in the future.

## 5. Digitalization of training data

Delimited archaeological sites, extracted from scientific publications, were employed as foundational data to generate polygonal representations within Google Earth. Subsequently, the data crafted in Google Earth were imported into QGIS. A meticulous collation of archaeological site coordinates and plans taken from several academic scientific literature formed the bedrock for delineating the boundaries of each archaeological site. The georeferencing and digitization into polygons took place within Google Earth, with subsequent refinement in QGIS. The mapping described above is derived from existing site plans that have been published. In addition to that, we

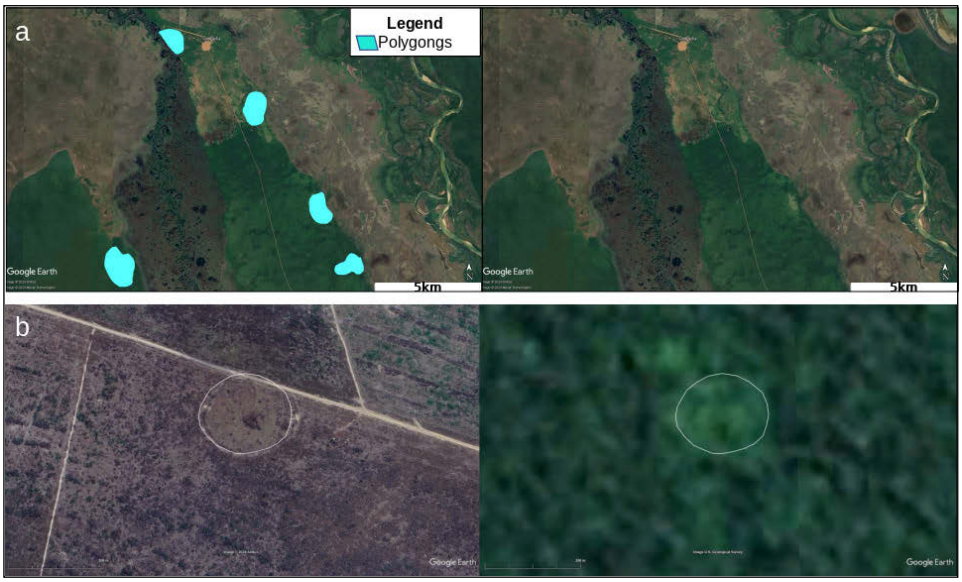


Fig. 3 – a) Archaeological sites at the Upper Xingu, Southern Amazon, Brazil (Heckenberger *et al.* 2003; Heckenberger 2004, 2014); b) example of a geometrical earthwork (ring ditch) in the State of Mato Grosso, Brazil, showing the current situation (left) and historical satellite imagery (Google) previous to the deforestation of the site.

have integrated previously published data on sites located in areas that have undergone deforestation. These sites primarily include earthwork structures, regionally recognized as 'geoglyphs' or ring ditches.

The primary objective of this integration is to develop a comprehensive dataset to facilitate the training of machine learning algorithms for automated site detection. The inclusion of sites currently situated in open terrains is meant to be used in tandem with historical satellite imagery – capturing the time when these areas were still densely forested.

## 6. Conclusion and future directions

By employing a multidisciplinary approach that integrates archival research, remote sensing, machine learning, and ground surveys, MAPHSA is establishing a comprehensive database of archaeological sites across a vast and ecologically sensitive region. The open-access nature of this database, adhering to FAIR principles, fosters transparency and international collaboration in heritage preservation efforts.

Looking ahead, several avenues hold promise for further research and development. Expanding the dataset by incorporating additional archaeologi-

cal publications will enrich the analysis and refine the effectiveness of automated site detection methods. Integrating more advanced machine learning algorithms alongside cutting-edge remote sensing and LiDAR technologies is likely to enhance both the accuracy and efficiency of identifying archaeological sites, particularly in forested areas. However, a crucial element will be the development of methodologies specifically tailored to the challenges presented by dense vegetation cover in the Amazon basin. By overcoming these challenges, MAPHSA can pave the way for a more nuanced understanding of archaeological landscapes across the region. Ultimately, this forward-looking initiative has the potential to propel archaeological preservation analysis to a new level of sophistication, providing invaluable insights that will guide future conservation efforts.

Jonas Gregorio de Souza, Marco Madella, Marco Moderato
Alan Tapscott Baltar, Shaddai Heidgen
CASEs, Departament d'Humanitats
Universitat Pompeu Fabra
jonas.gregorio@upf.edu, marco.madella@upf.edu, marco.moderato@upf.edu
alan.tapscott@upf.edu, shaddai.heidgen@upf.edu
Sebastian Fajardo
Leiden University
s.d.fajardo.bernal@liacs.leidenuniv.nl
Juan Carlos Vargas Ruiz
Universidad del Magdalena
jvargasr@unimagdalena.edu.co
Eduardo Góes Neves, André Strauss
Museum of Archaeology and Ethnology
University of São Paulo
edgneves@usp.br, strauss@usp.br

## REFERENCES

Bekiari C., Bruseker G., Doerr M., Ore C.-E., Stead S., Velios A. 2021, *Definition of the CIDOC Conceptual Reference Model v7.1.1 Model* (https://doi.org/10.26225/FDZH-X261).

Clement C.R., Denevan W.M., Heckenberger M.J., Junqueira A.B., Neves E.G., Teixeira W.G., Woods W.I. 2015, *The domestication of Amazonia before European conquest*, «Proceedings of the Royal Society B: Biological Sciences», 282.1812, 20150813 (https://doi.org/10.1098/rspb.2015.0813).

De Souza J.G., Schaan D.P., Robinson M., Barbosa A.D., Aragão L.E.O.C., Marimon B.H., Marimon B.S., Da Silva I.B., Khan S.S., Nakahara F.R., Iriarte J. 2018, *Pre-Columbian earth-builders settled along the entire southern rim of the Amazon*, «Nature Communications», 9.1, 1125 (https://doi.org/10.1038/s41467-018-03510-7).

Duran-Izquierdo M., Olivero-Verbel J. 2021, *Vulnerability assessment of Sierra Nevada de Santa Marta, Colombia: World's most irreplaceable nature reserve*, «Global Ecology and Conservation», 28, e01592 (https://doi.org/10.1016/j.gecco.2021.e01592).

Heckenberger M.J. 2004, *The Ecology of Power*, Routledge, London-New York (https://doi.org/10.4324/9780203486627).

Heckenberger M.J. 2014, *Bio-historical diversity, sustainability and collaboration in the Xingu*, «Anuário Antropológico», 39.2, 69-96 (https://doi.org/10.4000/aa.1250).

Heckenberger M.J., Russell J.C., Fausto C., Toney J.R., Schmidt M.J., Pereira E., Franchetto B., Kuikuro A. 2008, *Pre-Columbian urbanism, anthropogenic landscapes, and the future of the Amazon*, «Science», 321.5893, 1214-1217 (https://doi.org/10.1126/science.1159769).

Levis C. *et al.* 2017, *Persistent effects of pre-Columbian plant domestication on Amazonian forest composition*, «Science», 355.6328, 925-931 (https://doi.org/10.1126/science.aal0157).

Levis C., Souza P.F.D., Schietti J., Emilio T., Pinto J.L.P.D.V., Clement C.R., Costa F.R.C. 2012, *Historical human footprint on modern tree species composition in the Purus-Madeira Interfluve, Central Amazonia*, «PLoS ONE», 7.11, e48559 (https://doi.org/10.1371/journal.pone.0048559).

Morandi P.S. *et al.* 2020, *Tree diversity and above-ground biomass in the South America Cerrado biome and their conservation implications*, «Biodiversity and Conservation», 29.5, 1519-1536 (https://doi.org/10.1007/s10531-018-1589-8).

Palace M.W., McMichael C.N.H., Braswell B.H., Hagen S.C., Bush M.B., Neves E., Tamanaha E., Herrick C., Frolking S. 2017, *Ancient Amazonian populations left lasting impacts on forest structure*, «Ecosphere», 8.12, e02035 (https://doi.org/10.1002/ecs2.2035).

Peripato V. *et al.* 2023, *More than 10,000 pre-Columbian earthworks are still hidden throughout Amazonia*, «Science», 382.6666, 103-109 (https://doi.org/10.1126/science.ade2541).

Schaan D., Pärssinen M., Saunaluoma S., Ranzi A., Bueno M., Barbosa A. 2012, *New radiometric dates for precolumbian (2000–700 b.p.) earthworks in western Amazonia, Brazil*, «Journal of Field Archaeology», 37.2, 132-142 (https://doi.org/10.1179/009346901 2Z.00000000012).

Sheldrick N., Zerbini A. 2017, *A heritage inventory for documenting endangered archaeology in the Middle East and North Africa*, «ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences», IV-2-W2, 237-241 (https://doi.org/10.5194/isprs-annals-IV-2-W2-237-2017).

Siqueira M.F.D., Peterson A.T. 2003, *Consequences of global climate change for geographic distributions of cerrado tree species*, «Biota Neotropica», 3.2, 1-14 (https://doi.org/10.1590/S1676-06032003000200005).

ABSTRACT

The first stage of the Mapping the Archaeological Pre-Columbian Heritage in South America (MAPHSA) initiative integrates transregional datasets from Brazil and Colombia, encompassing a broad spectrum of cultural heritage. MAPHSA aims to construct an integrated database of legacy and new information pertaining to the pre-Columbian archaeological heritage of South America. It aims to provide an open-access digital repository grounded on the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles, thereby facilitating resource accessibility and fostering collaborative research. This database will benefit researchers within the surveyed countries, the international archaeological community, and national heritage agencies, consequently fortifying heritage preservation and promoting cross-border studies. MAPHSA adopts a multidisciplinary approach, combining archival research, remote sensing techniques, machine learning methodologies, and ground surveys to identify and evaluate the preservation status of archaeological sites. The workflow integrates diverse datasets into a Postgres/PostGIS geodatabase, complemented by the Arches information management platform incorporating a CIDOC-CRM ontology data model, drawing insights from analogous large-scale projects. By accentuating the heritage significance of regions such as the Amazon and the often-neglected Cerrado savanna, MAPHSA aims to alter public perceptions and stimulate concerted efforts towards their protection.