Proceeding Paper

SH.AR.P.P. (SHared ARchaeological Platform for Prehistory): Building an Informative System for Italian Prehistoric and Protohistoric Sites [†]

Varinia Nardi [1], Claudia Sabbini [1] and Saverio G. Malatesta [2,*]

[1] Independent Researcher, 00185 Roma, Italy; nardivarinia9@gmail.com (V.N.);
claudia.sabbini24@gmail.com (C.S.)

[2] Interdipartimental Research Center Digilab, Sapienza Università di Roma, 00185 Roma, Italy

[*] Correspondence: saveriogiulio.malatesta@uniroma1.it

[†] Presented at the ArcheoFOSS XIII Workshop—Open Software, Hardware, Processes, Data and Formats in Archaeological Research, Padova, Italy, 20–22 February 2019.

Abstract: Italy's prehistoric and protohistoric heritage is subjected to more threats than any other for reasons that go beyond its antiquity: if it is true that the record is often less imposing in its traces, this cannot justify the widespread general ignorance surrounding it. Such ignorance is mirrored and amplified by the lack of systematic recording of all kinds of evidence regarding this timeframe. Even though there is more than one platform available for such recording, its use is rarely considered accessible, and their features appear to be more oriented towards cataloguing rather than research or trying to capture the attention of a non-specialized audience. In this article, we pinpoint what seems to be missing and propose a model that can deal with the challenge.

Keywords: prehistory; archaeological research; database; GIS; cultural heritage; digitalization; open-source

1. Introduction

The abundance of Italian archaeological heritage goes without saying and so does the compelling need for digital management of old and new data. The last two decades have seen worldwide archaeological research deeply involved in establishing new ways of recording and standardizing data at every level, a necessary effort to preserve it and achieve the possibility to share it in order to increase everybody's insight into our past. As a result, countless projects have been launched since then that aim to set a standard for both excavation recording needs and spatial inquiries, ranging from proper landscape archaeology to comparative analysis. In this paper, we set aside the former approach to focus on possible mapping strategies for regional and over-regional archaeological evidence within a specific timeframe: the prehistory and protohistory of the Italian peninsula.

The peculiarities of such contexts call for a project tailored to the needs of the researchers, oriented towards the interaction between landscape, natural resources and human activities, as well as mindful of the possibility of developing informative strategies aimed at the general public. As an independent research group born into the Archeo&Arte3 dLab of Sapienza University in Rome, we have chosen to focus our efforts in that direction and started project SH.AR.P.P. (SHared ARchaeological Platform for Prehistory).

2. Background and Motivation

As previously mentioned in the introduction, there are already many projects to catalogue data of archaeological sites with relation to their territorial features, so why SH.AR.P.P.? Why add another proposal to the existing ones instead of enriching them? Our aim is not that of adjusting an alternative to older ongoing projects, but to build a solid branch that can boost their efficiency if linked or become a starting point for different

approaches to our cultural heritage if left on its own. To explain how this can be, it is necessary to start with some reference on what is already in use, to analyse what has efficiently served its purpose and what still seems to be missing.

The most influential webGIS actually in use is probably ArkeoGIS, developed by the Université de Strasbourg since 2009, which has among its strong points a worldwide perspective, the support offered by French and German institutions and agreements with Universities, private companies and independent researchers who contribute to increase and update its dataset, not to mention the importance conferred to geological features to be inferred with anthropological considerations [1] (p. 401) [2] (p. 1). Although it is an exceptional open source and a completely free tool, access to the platform is still limited to researchers whose profile has to be approved. The idea that limited access may guarantee safety from plundering is still widespread for obvious reasons, but it has a drawback: it necessarily confines an important part of our heritage to professionals, excluding the public who has every right to benefit from it. The matter is of no little importance in an age where institutions frequently ask culture to provide for itself: here public interest becomes pivotal in raising the necessary resources. Furthermore, as far as Italy is concerned, ArkeoGIS is still encountering difficulties in acquiring a relevant dataset.

Another interesting project of considerable merit is FastiOnline—FOLD&R run by AIAC (Associazione Internazionale di Archeologia Classica) and CSAI (Centro per lo Studio dell'Italia Antica by the University of Texas and the University of Austin) since 2000. Though it also has worldwide ambitions, this project differs from the ArkeoGIS experience for its less detailed database structure but offers more than an online webGIS that records research activities and the institutions involved. Linked to the spatial visualization of each team research area is an online peer-reviewed journal, Fasti Online Documents & Research (FOLD&R) where researchers can publish their results. The site embraces three main fields: excavations, surveys and archaeological conservation, each one corresponding to a proper independent section [3]. Here, archaeological sites and research areas can be understood—where recorded—through a summary of the findings, a few searchable fields and possibly through links to a specific bibliography uploaded on the site or provided through direct citations. The effort is enormous and the amount of data impressive, but queries still lack the flexibility needed for many purposes of research and the dataset is far from complete.

Data exchange between such broad projects and those conducted by public institutions is continuous. Italy has been working on its own database, known as SIGECweb: a platform that aims to set standards for data entry, verify the accuracy, and allow the distribution of data for both tangible and intangible forms of cultural heritage on national boundaries. The project has been growing for years through the tireless collaboration between MiBACT (Ministero per i Beni e le Attività Culturali e per il Turismo) and the ICCD (Istituto Centrale per il Catalogo e la Documentazione): their results have brought forth a significant amount of specific data sheets and thesauri to be employed for the standard classification of evidence. The management of relationships between data sheets and the inference between data is guaranteed by a semantic ontology developed by the ICCD in collaboration with the CNR (Centro Nazionale di Ricerca): ArCO has been built following LOD logic in order to make efficient reuse of existing catalogued data and enable the fair sharing of its own contents [4].

If the task of elaborating a national standard for terminologies and descriptive features is indeed to be appointed to national organs, it is unreasonable to think that they could handle a data entry as massive as the one required on their own. Hence the need for collaboration between the central government and regional efforts in order to build a comprehensive database. The challenge was issued back in 2001 through the "Accordo Stato-Regioni, art. 3"—an agreement between the central government and the regional authorities—with different outcomes [5].

The response has shown a variety of distinct approaches to cultural heritage with regard to the features considered, the degree of in-depth analysis, the field research strategies

and the access to users. The experiences mentioned will surely contribute significantly to the amount of data that will be accessible on the SIGECweb, but it is clear that much still needs to be done in order to guarantee data consistency and substantial coverage. It is also evident that—apart from some exceptional instances—the prehistoric and protohistoric record is the most penalized. This is due to several reasons, starting with the very establishment of the thesauri realized by the ICCD, where sites and material culture are organized and defined under the scope of functional classes. The latter may be well suited to historical evidence—which, we agree, is preponderant—but are less fitting for materials so distant from our reality or any written testimony, a fact that should not be ignored. Prehistorians know the lengthy and difficult debate to assess the actual use of a tool, the high chance of encountering multi-purpose objects, and the relevance of specialistic analysis to account for an answer. To simplify for the sake of standardizing different realities may lead to an undesirable result: an enormous dataset lacking the useful indexes to employ such data. The dilemma is recognizable from various angles and introducing the issue of functionality is but to lift the lid of Pandora's box. Similar issues emerge when we confront the idea of defining land-use, assessing site size or the very extension of human occupation, not to mention the relevant fluctuations among chronologies. We are by no means arguing that colleagues working on historical contexts do not face the same occurrences, but we must remember that the more we go back in time, the more blurred the picture becomes, the harder the effort of defining what is before our eyes through fixed categories.

From this, the idea of SH.AR.P.P. was born: to become a bridge among different experiences and expertise, with the goal of bringing prehistory to the same level of resonance that other stages of our history have. As a bridge it has no intention of replacing other existing platforms, but wishes to extend their scope of action where it is otherwise necessarily limited.

The objective is ambitious and requires different steps to be undertaken. We planned at first to collect published data in order to analyse what input their bibliography had to offer and the difficulties we would encounter in trying to digitize textual information into searchable fields. On this basis we started structuring a database that could interact with ICCD's data sheets and still be enriched with details useful to researchers who are willing to link sites following different patterns of analysis, among which no small role is reserved to the interaction with natural resources. To this aim, we also provided a GIS system to match our database: its final form should allow the user to visualize queries based on a remarkable amount of structured data as well as support archaeological research and territorial awareness.

3. Materials and Methods

At the outset, our priority was to collect all available bibliographic documentation for a number of sites of different chronological phases (Table 1). While reading, we tried to record the basic knowledge of each site in an xml. file to keep track of numbers and also assess an average estimate of the quality and quantity of data available.

Table 1. Number of sites taken into account while planning the fields for the database.

Paleolithic	Mesolithic	Neolithic	Eneolithic	Bronze Age	Iron Age	Total
46	7	341	29	81	153	657

This time-consuming operation was necessary for us to understand what kind of information we were meant to transform into actual data to be stored in a database. We had a clear perception of the changes to archaeological research in the last century, the evolution of the specialized lexicon as well as the gap emerging between sites: some are known merely by name or little more than casual hints, others have been repeatedly published in detail. We also happened to notice disagreements between publications and assessing

chronologies may pose a challenge when it is uncertain if opportune calibration has been applied or not.

This first step made us aware of which categories were the most relevant to describe such realities, what was already enclosed in the ICCD's data sheets (and what prehistory could not offer to it, even when included) and what was still missing or could not be added to the present structure but that we nevertheless deemed noteworthy. An essential asset to this step was the coordination with the "P.A.S.T. in Coast" experience [6] that was reaching its conclusion and offered us an excellent example of recent data available from surveys focusing on human–landscape interaction.

As we were approaching the difficult task of building a solid structure for SH.AR.P.P., one that would be effective but also detailed and open to further improvement, we decided to focus our efforts on open-source tools whenever possible. This is not the place to describe the benefits of FOSS (*Free and Open Source Software*), but we cannot avoid underlining the reasonableness of that choice, while carving out a complex system that should constantly communicate with other platforms. As a result, we focused on *MariaDB* to build our database, although we encountered never-ending difficulties as humanists approaching a different world and frame of mind. We have been supported in this journey by a professional expert and by the software *SQLdbm* that—although not open source—has still been an irreplaceable ally, generating SQL scripts from graphic representation of our database tables and relationships, taking in account instructions concerning indexes and column constraints. What we obtained from SQLdbm cannot be regarded as the actual output, but it still represented the starting point for our modelling, which would have been impossible otherwise for archaeologists working in digital humanities.

In order to provide territorial information for archaeological data, we built a base map using Qgis 3.4.2 (Madeira) software. Landscape and geographical features were adopted from national and regional open data platforms, among which were the AGID (*Agenzia per l'Italia Digitale*) and the *Geoportale Nazionale*. More popular repositories such as "*OpenStreetMap*" (OSM) and "*OpenMapTiles*" (OMT) were also generously employed. The informative layers realized aim to embrace detailed data on political and natural features.

Establishing a nexus between cartographic data and those stored in the database should allow us to visualize complex queries aimed at solving archaeological questions and offer a different view on old debates. The system output will consist of a structured and linked open data collection in order to allow the user to make SPARQL queries.

4. Results

The strategy adopted allowed us to build a database structure modelled through eight main entities and twenty-six thesauri. The data recorded embrace a wide range of elements from site definition with spatial–chronological assessment to structural evidence known, from material culture useful to assess the site identity to osteological remains and archaeobotanical finds, from museums and other places connected to the site to research and promoting activities concerning the archaeological sites (Figure 1). All fields are searchable and indexed with a few exceptions concerning the actual visibility of and access to the archaeological location and the detailed interpretation of the site functionality. Every site will be linked to a bibliographical sheet with all known references.

Among the cited thesauri, six are directly linked to ICCD's definition of site (SI sheets), monument (MA sheets), and material culture (RA sheets) and allow for our work to be easily transferred into in the SIGECweb system. Data concerning the placing of archaeological sites has also been structured in a cooperative form with the field provided by the national platform for the same purpose. Other fields and thesauri are planned to offer an in-depth description of relevant sections or to translate ICCD's definitions to a nomenclature more suited to a prehistorian researcher. An example of how we have operated on this perspective comes from the already mentioned definition of material culture finds. We do not foresee at present the possibility of inserting a catalogue of every find discovered for a site, but we wish to incorporate specific elements pivotal to site

assessment, for example in the absence of absolute dates. Still, the actual nomenclature proposed on ICCD's sheet RA ("Reperti Archeologici") reasonably focuses on functionality, which may stretch definitions concerning prehistoric finds where usually the first step is oriented toward raw materials classification. We have then decided to propose a double attribution for such finds: every object is linked both to its raw material class and to the RA sheets' definition. Such methodology also actually simplifies other relationships to be built, as it makes it easier to connect samples of material culture to the specific analysis they were subjected to. For example, a researcher may have performed trace analysis on all osteological finds, including tools, but may not have done the same for those functionally equivalent tools realized in metal.
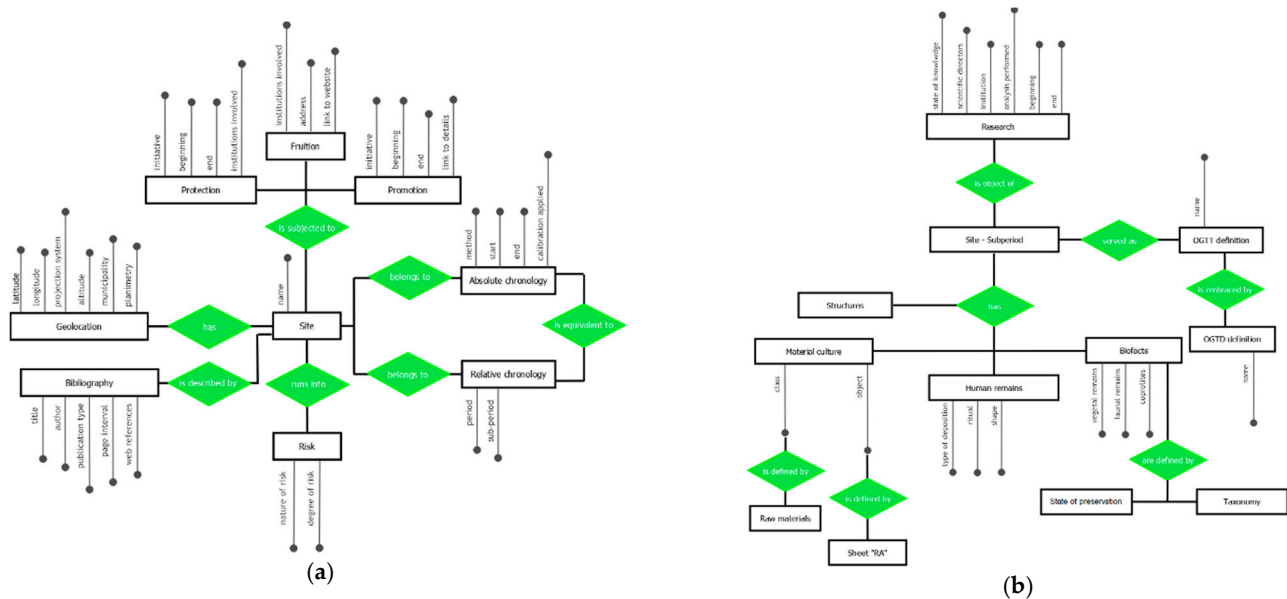

Figure 1. Main structure of SH.AR.P.P. database: (**a**) Elements related to the site definition in terms of spatial-chronological assessment, state of preservation and promotion; (**b**) elements related to the site description, concerning archaeological and bioarchaeological record, and the research activities undertaken.

If sometimes adapting existing thesauri to prehistory only requires adding some item to the institutional framework—as is the case of SI sheets focusing on site definition—other points have been more difficult to tackle. Such is the case of chronological attribution, due to the very uneven record available. Absolute dating is less employed than material culture seriation and the subsequent relative chronologies, which brings much uncertainty to the interpretations: even though this is something we are used to in archaeological practice, it becomes harder to incorporate its shades into a digital framework. We have chosen to realise a descriptive structure with absolute dates (when available) specifying calibration range and dating method, but we have also selected regional chronologies among the most shared to allow for comparison between most archaeological sites, thus establishing relationships between the two sequences.

To date, the structure here introduced is being put to the test. We are realizing each entity one at a time and running independent testing, making use of the data collected in our first survey among bibliographical sources. The material selected for testing focuses on three South Italian regions (Campania, Basilicata, Calabria).

A great deal of work has been carried out and much more is still needed, but the first results are encouraging and the perspectives for independent research are numerous.

All the while we proceeded with the QGIS project modelling, searching for useful open data informative layers to add. A first step regarded political features as boundaries, but also ancient and modern road networks, trails and toponymy.

Great care was also taken in laying out configuration layers such as contours, orography and hydrographic networks. Furthermore, a more detailed level of information can be attained by placing other natural features such as springs, rocky outcrops and pedological distribution of soils.

A rich geographic informative model, once linked to the archaeological data, can allow us to conceive predictive strategies for observing site distribution within the landscape: raw materials availability, as well as the presence of specific physical and environmental characteristics, can shed light on cultural choices and settlement strategies. To these land descriptors are added all archaeological features, such as sites or other evidence, with their GPS coordinates. When such detailed information is unknown, however, we find a solution by defining the area of interest at a local or municipal scale. The source of these types of data is always reported in order to discern between punctual information and the one referring to a general area.

5. Discussion

Developing a suitable structure for SH.AR.P.P. required us to visualize not only the ends behind the project, but also a public to answer to. What we never wavered on was the willingness to bring SH.AR.P.P. within the reach of almost everyone: as we pointed out earlier, we live in an age when only wishful thinking can believe in protecting our heritage by hiding it. Today, if we want to preserve our heritage, we must make it known, lived, ancient but still present in everyday life for the community that has the duty and the pleasure to participate in its safeguarding and promotion. We will therefore release all published data with their due references on a webGIS platform as soon as possible. For unpublished documents, we expect instead the possibility of adopting different policies, which are to be decided in agreement with the authors. As seen through our presentation, our main audience is composed of researchers and students, but SH.AR.P.P. also possesses some features of interest to others. The reflections and methodologies employed in this project can find a wide number of applications not only for archaeological research, but also in the promotion of cultural heritage, involving a non-specialized audience. SH.AR.P.P has been designed to become a multi-faceted platform offering support to public institutions, independent and structured researchers and people with an interest in promoting their heritage. It may offer storage space for new research, simplifying the exchange between it and older data, it may prove useful to reconstruct the history of previous research in a limited area and it may visually express the bond between human choices and environmental features.

The very interaction evidenced between natural landscape and human activities may encourage different ways to experience our land, merging different channels of tourism, or promoting local initiatives connected to museal structures or private enterprises working in the field of cultural heritage. The possibilities are endless, but everything starts with the awareness of what is ours to treasure and uphold, which is why we believe in this project.

We strongly believe that the upgradability of SH.AR.P.P. makes what we have introduced here a starting point for needed flexible applications, and we are confident that its call for data sharing puts it in line with other initiatives currently underway.

References

1. Bernard, L.; Ertlen, D.; Schwartz, D. ArkeoGIS, merging geographical and archaeological datas online. In Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Paris, France, 22–25 April 2014; Giligny, F., Djindjian, F., Costa, L., Moscati, P., Robert, S., Eds.; Archaeopress, Publishing Ltd.: Oxford, UK, 2015.
2. Bernard, L. Ten Years of ArkeoGIS: Feedback on the Challenge and Benefits of Digitally Sharing Archaeological Data. In Proceedings of the Workshop #2 DARIAH-CH, Neuchatel, Switzerland, 5–6 December 2019; Available online: https://univoak.eu/islandora/object/islandora:93356 (accessed on 15 March 2021).
3. Fasti Online. Available online: http://www.fastionline.org/index.php (accessed on 15 March 2021).
4. Carrero, V.A.; Gangemi, A.; Mancinelli, M.L. ArCo Ontology Network and LOD on Italian Cultural Heritage, in Open Data and Ontologies for Cultural Heritage. In Proceedings of the First International Workshop on Open Data and Ontologies for Cultural Heritage, Rome, Italy, 3 June 2019; Poggi, A., Ed.; CEUR Workshop Proceedings. pp. 97–102. Available online: http://ceur-ws.org/Vol-2375/short3.pdf (accessed on 15 March 2021).
5. Information Is Derived from the Conference "La Catalogazione dei Beni Culturali in Ambito Regionale: I Sistemi Informativi Regionali e Il Rapporto con Il Catalogo Nazionale", held in Livestream by the ICCD (Istituto Centrale per Il Catalogo e la Documentazione) on Its Youtube Channel. Available online: https://www.youtube.com/watch?v=B7QM2fisNTI&t=507s (accessed on 15 March 2021).
6. Di Bianco, L.; Albore Livadie, C.; Malatesta, G.S. Il Progetto "P.A.S.T. in Coast" e l'insediamento protostorico dello Scalandrone di Scala (Costa d'Amalfi—Campania). *Territ. Cult.* **2017**, *27*, 22–29.