ARCHAIDE

ARCHAELOGICAL
AUTOMATIC INTERPRETATION
AND DOCUMENTATION
OF CERAMICS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 693548

Inside the ArchAIDE project

Objectives

Pottery classification is of fundamental importance for the comprehension and dating of the archaeological contexts, and for understanding the dynamics of production, trade flows, and social interactions. Unfortunately, classification requires complex skills and since it is heavily dependent on human inspection and interpretation it is a very time consuming activity, both for researchers and professionals. The digital automation of pottery classification would revolutionise archaeological practice, behaviours and expectations, would meet real user needs and generate economic benefit, reducing time and costs, would create societal benefits from cultural heritage, improving access and exploitation of the digital cultural heritage in a sustainable way, and would afford a deeper knowledge of the context in which cultural heritage exists and is formed. Our project is aimed at providing an innovative approach and infrastructure to support the archaeological interpretation process (during both fieldwork and post-excavation analysis). Our idea is to support the classification and interpretation work of the archaeologist with innovative computer-based tools, able to provide the user with features for the semi-automatic characterization and matching of each discovered sherd over the huge existing ceramic catalogues. This will include the design, development and assessment of a new software platform offering applications, tools and services for digital archaeology. This framework, that will be available through both a mobile application and a desktop version, will be able to support archaeologists in recognising and classifying pottery sherds during excavation and post-excavation analysis. It will contribute to decentre cultural heritage away from institutional structures towards the individual. The system will be designed to provide very easy-to-use interfaces (e.g. touch-based definition of the potsherd profile from a photograph acquired with the mobile device) and will support efficient and powerful algorithms for characterization, search and retrieval of the possible visual/geometrical correspondences over a complex database built from the data provided by classical 2D printed repositories. Our approach is driven by archaeologists' needs; since we are aware of the caution of the discipline in front of the replacement of well-established methods, we plan to support this specific Humanities domain by exploiting what is already available in the Archaeology domain in terms of good practices and representation paradigms. We will investigate both 2D- and 3D-based characterization of shapes for search and interpretation, always trying to be compliant with the pre-existing non-digital methodologies. For instance, profile-based approaches, heavily used in archaeological practice, are an ideal approach to support classification. We thus plan to deliver efficient computer-supported tools for drafting the profile of each sherd and to automatically match it with the huge archives provided by available classifications (currently encoded only in drawings and written descriptions contained in books and publications). Instruments for 3D acquisition of classified potsherd will also be supported, to allow to

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 693548

automatically enrich the repository and to contribute to digital documentation. The system will also be able to support the production of archaeological documentation, including data on localisation provided by the mobile device (GPS).
This platform will allow to access tools and services able to enhance the analysis of archaeological resources, such as the open data publication of the pottery classification, or the data analysis and data visualisation of spatial distribution of a certain pottery typology, leading to a deeper interpretations of the past. The integration of cultural heritage information from different sources, together with faster description, cataloguing and improved accessibility can be exploited to generate new knowledge around archaeological heritage. Data visualisation, for instance, would stimulate new research perspectives, and could enable new interpretation and understanding of Europe's history, and would bring archaeological storytelling to new audiences in a novel ways. By means of a wider dissemination of user-generated content, the framework would permit to develop the culture of sharing cultural resources, research and knowledge.
The project will involve archaeologists, mathematicians, IT scientists and archaeology professionals; it will be tested on different pottery classes (with an extensive on-the-field test and assessment experiences), especially having large-scale production and distribution, so as to promote and foster a new shared culture among European archaeologists. In fact, the employment of a common platform/environment would stimulate cross-border, cross-lingual multi-disciplinary research, enabling collaboration, partnerships and co-production of knowledge across sectors and communities of researchers and users.
The project will develop a new infrastructure and the related algorithms and tools to support archaeological investigation, documentation and reasoning.

Overall concept

The rationale of the project is based on the assertion that we urge digital tools that should increase our insight capabilities, increasing accessibility to the huge corpus of knowledge laying in books and grey literature and make this available to archaeologists directly on the field, in a structured and semi-automatic way. Current reasoning and interpretation is based mostly on an extensive use of personal experience, backed by the consultation of paper-based resources. Therefore, we plan to make a proficient use of all this knowledge, bringing it to the mass of professional archaeologists working on the field. On the other hand, the method for describing archaeological potsherds is in itself a scientifically valid model, but affected in terms of effectiveness by a lack of sustainability, so our proposal is not to change the current overall approach and reasoning process of archaeological community, but to support the current practice in a radically innovative way by means of an integrated set of cutting edge ICT technologies. Furthermore pottery recognition requires complex skills and, since it is heavily dependent on human inspection and interpretation, it is a high-priced and a time consuming activity, both for researchers and professionals. The consequence of this is the fact that this operation requests heavy public or private funding, that many times are not available condemning important pieces of our cultural heritage and history to oblivion. Current ICT technologies permit to innovate substantially the practice of on-the-field archaeology (speeding up the analysis and documentation process, reducing the required level of expertise), without a thorough phase of fundamental research, and to achieve a major technological breakthrough that will further consolidate and strengthen Europe's scientific and technological leadership in a market that could show a strong and stable growth for professional applications. On the other hand, the policy of the European

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 693548

Commission about a free circulation of knowledge in terms of open access to cultural data, more than the ICT development, allows us to pursue the aim to bring cultural content to new audiences in novel ways.

Digital pottery catalogue: moving from paper to digital schematic representation

Current catalogues describe a large number of classes of pottery using a quite standardized and consolidated format, which includes a sketch of the internal and external profiles of the exemplar, and also a description of morphological, typological, and (if available) decoration features. The first contribution of this activity will be an as-automatic-as-possible procedure to transform the paper catalogues in a digital description, to be used as a data pool for an accurate search and retrieval process. This will entail: scanning (2D digitization) of the paper catalogue(s); segmentation and vectorialization of the graphical drawings proposed in those printed catalogues; and linking the graphical representation with the metadata reported in the catalogues. Since we are interested in designing automatic matching and retrieval features, digital description does not means here only digitization of the paper catalogues, but includes understanding the meaning of the graphic representation and its conversion in a format that includes shape (in vectorial format, not raster) and semantic. A set of metadata will be added to the extracted data.
This digital catalogue will be a deliverable by itself, being produced in a format that will allow immediate publication on the web and free access by our community.

A tool for on-site documentation of pottery sherds

A second basic resource of our project is a system that should support the on-site digital documentation of the sherds. This tool (mainly designed for tablets, to allow easy deployment on the excavation site) will work as follows: the user will be able to take an image of a fragment (using the on-board camera) and will provide only a few input data to the classification system, to characterize the sherd. A record will be created for each sherd, holding both a graphical representation (the image(s) taken), plus the vectorial representation reconstructed from the image(s) and the user input. For the production of the latter semantic graphics data, a simple sketch-based interface will give the possibility to indicate portions or lines on the acquired image (i.e. the internal or external profile, a decorated area) so that the system will be able to categorize and store those vectorial data and use them in the subsequent classification phase. The interactive mobile-based system will also allow selecting specific metadata tag, in case the user wants to add some preliminary knowledge that might be used (and checked) later on in the classification stage.
This sherd record will be transferred to a remote server for inclusion in the semantic database of the current excavation and will be used for classification purposes.

Semi-automatic classification of pottery sherds (on-site or remotely)

Given a sherd record produced by the tool, the system (a software component running on a remote server) will be able to perform a search over the classification knowledge stored in the digital catalogue. The search results could be refined by adding more metadata tags (the system will offer all resources for easy metadata selection, such as ontologies and dictionary of terms). The result of each single classification action will be added to the sherd record, together with the added metadata tags (if any) used by the user in

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 693548

the search & classification action (to keep track of the specific decision process implemented to get to the classification decision).

Automated production of the sherd's identity card

In addition to the sherd classification, another time-consuming operation is the production of the findings catalogue. Therefore, we plan to derive a complete identity card of the sherd (following consolidated archaeological documentation procedures and practice) by transforming the data we have stored in the sherd record into a formatted electronic document, printable or visual. This will be produced in an automatic manner and will also allow implementing a few functionalities to add any other data or visual assets the archaeologist would like to have in the identity card.

Data visualisation

The system will also generate a knowledge based system for archaeological management, research and knowledge generation. Since the comprehension of the assemblage, namely collection of objects of every single non empty stratum, of the context, namely the relationships among different ceramic classes in the same location, of the relationships between the location of the finding and the productive centre, and of the relationships with elements found in different locations is a fundamental operation in pottery analysis, all the information encoded in the pottery identity cards will be shared, visualised and integrated with cultural heritage information from different sources in order to produce a really significant impact in the advancement of the discipline and in the accessibility for professional and non-professional users. Real time comparisons between different archaeological sites and regions will be made possible, thus highlighting differences and commonalities in the economy of the ancient world. Furthermore a web-based visualization tool will improve accessibility to archaeological heritage and generate new understanding about the dynamics of pottery production, trade flows, and social interactions. Finally, once the data are geolocated, the system will provide mapping tool in order to visualise pottery diffusion, site location and so on.

Open data and reuse

The long term sustainability of the ArchAIDE project is also related to the long-term preservation and the reuse of the data collected. As open data best practice requires data be made available and easily accessible for re-use, the data created through this project will be preserved and disseminated according to the international data archiving standards set out in the Open Archival Information System (OAIS). The data will be disseminated online, and made freely available for use and re-use.

Expected impacts

We expected impacts on archaeological professional labour market, archaeological research, and training and education in archaeology.

As for the archaeological professional labour market, it is estimated that approximately 33,000 archaeologists now work across Europe as a whole. The largest estimated populations of archaeologists are in the United Kingdom, Germany, France and Italy, with more than 4,000 individual archaeologists working in each of these countries. On the whole, this group of professionals represents 0.006% of the combined

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 693548

total workforces of Europe. Moreover, it is calculated that a total of over €1 billion is spent on professional archaeology every year, with the majority of that expenditure being on the salary costs, which represent 60% of the costs of running an archaeological organization . Outside Europe, several studies have permitted to estimate a number of 11,350 people working as professionals in archaeology in the USA (2010), a number of 6,255 professional archaeologists in Japan (2008), and between 500 and 600 professional archaeologists working in Australia. This information underlines the world-wide potential market of ArchAIDE. Besides, the possibility to reduce time and costs will bring positive effect on the whole sector: archaeologists could be paid for produce new knowledge and archaeological organization could invest money in R&D.

The figure above give the potential market and economic size open for ICT tools supporting the field and documentation work of archaeologists. This domain is not new to the endorsement of new technologies.

Another point to underline is not just the potential size, but also the current lack of competitors on the specific domain of the ArchAIDE system: no existing commercial tool implements the features planned in ArchAIDE.

As for archaeological research, ArchAIDE will move archaeologists from spending time on routine tasks like drawing and classification to create background knowledge and enable new knowledge generation, permitting to develop new research field in archaeology.

As for education and training, students complete their studies with training activities leaded by Universities or Archaeological companies on archaeological sites. In such a context, ArchAIDE will become an extra training tool with two main advantages:

-a significant increase of App downloadings;

-consolidating the use of the tool, since the students will become archaeologists who apply those knowledge, skills and tools in their professional life.