EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2016)
C. E. Catalano and L. De Luca (Editors)

ArchAIDE - Archaeological Automatic Interpretation and Documentation of cEramics

M. L. Gualandi [1], R. Scopigno[2], L. Wolf[3], J. Richards[4], J. Buxeda i Garrigos [5], M. Heinzelmann[6], M.A. Hervas[7], L. Vila[8] and M. Zallocco[9]

[1] Dipartimento di CivilťL e Forme del Sapere, Universita' di Pisa, Italy
[2] CNR-ISTI, Pisa, Italy
[3] Blavatnik School of Computer Science, Tel Aviv University, Israel
[4] Archaeology Data Service, University of York, York, UK
[5] Prehistoria, Historia Antiga i Arqueologia, Universitat de Barcelona, Spain
[6] Archaeological Institute, University of Cologne, Cologne, Germany
[7] Baraka Archaeologos S.A., Ciudad Real, Spain
[8] Elements - Centre de gestio i difusio de Patrimonio Cultural, Palma de Mallorca, Spain
[9] Inera srl, Pisa, Italy

Abstract

The goals of H2020 project "ArchAIDE: are to support the classification and interpretation work of archaeologists with innovative computer-based tools, able to provide the user with features for the semi-automatic description and matching of potsherds over the huge existing ceramic catalogues. Pottery classification is of fundamental importance for the comprehension and dating of the archaeological contexts, and for understanding production, trade flows and social interactions, but it requires complex skills and it is a very time consuming activity, both for researchers and professionals.
The aim of ArchAIDE is to support the work of archaeologists, in order to meet real user needs and generate economic benefits, reducing time and costs. This would create societal benefits from cultural heritage, improving access, re-use and exploitation of the digital cultural heritage in a sustainable way. These objectives will be achieved through the development of:

- *an as-automatic-as-possible procedure to transform the paper catalogues in a digital description, to be used as a data pool for search and retrieval process;*
- *a tool (mainly designed for mobile devices) that will support archaeologists in recognizing and classifying potsherds during excavation and post-excavation analysis, through an easy-to-use interface and efficient algorithms for characterisation, search and retrieval of the visual/geometrical correspondences;*
- *an automatic procedure to derive a complete potsherds identity card by transforming the data collected into a formatted electronic document, printable or visual;*
- *a web-based real-time data visualisation to improve access to archaeological heritage and generate new understanding;*
- *an open archive to allow the archival and re-use of archaeological data, transforming them into common heritage and permitting economic sustainability.*

Those tools will be tested and assessed on real-cases scenarios, paving the way to future exploitation.

Categories and Subject Descriptors (according to ACM CCS): Computer Graphics [Computing methodologies]: Shape Modeling—Visualization [Human-centered computing]: Visualization systems and tools—

1. The project in the state-of-the-art, and the project approach

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.

Computer Science already took into account the issues of pottery classification. Basic image processing techniques were proposed [DLS95], while most of the approaches have been based on the analysis and comparison of profiles [KS02, GKSS04, MG05] or

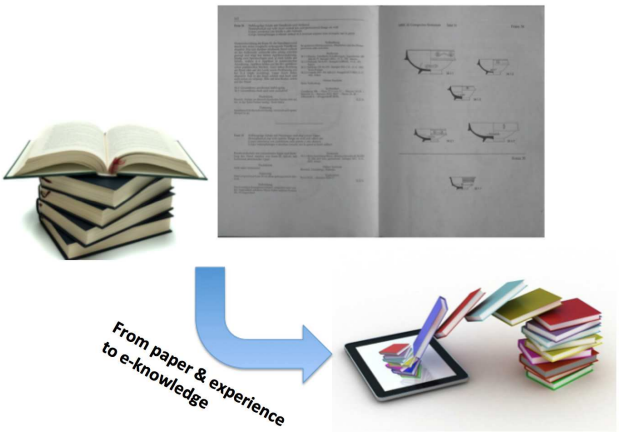

Figure 1: *The methodological advance planned with the ArchAIDE project.*

of texture and color [MD13].
Automated systems to collect and compare data have been proposed as well [KSM02, SKN*12] but their use limited by the lack of standardization of typology definition, and also by the heterogeneity of input data.
Nevertheless, the digital automation of pottery classification would have a major impact on the work of , especially those working in rescue excavations (Figure 1).

The work of the project 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 (Figure 2). 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 characterisation, 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 characterisation of shapes for search and interpretation, always trying to be compliant with the pre-existing non-digital methodologies. 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 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. The components of this platform are described in the following subsections.

2. Project methodology

The ArchAIDE project is structured in ten work packages. As usual, the activities include common activities (project management, specifications, dissemination, exploitation) and focused activities, each of the latter involving a subset of partners focusing on specific themes. They concern the development of the key technological components and their integration in a common system. The overall planned workflow, at the basis of our project vision, is depicted in Figure 2. Two structural components are at the base of our proposed renovation of the work of the archaeologists:

- The digitized corpus of catalogues;
- The overall data archive, storing both the digitized catalogues and the results of previous archaeological campaigns (including both raw data and final documentation produced).

According to our vision, the on-the-field activity of the archaeologists will be served by ArchAIDE following the pipeline presented in Figure 2, according to a cycle where all actions find support or contribute data to the data archive:

- For each a new fragment/sherd discovered on the field:
  - The fragment is documented (taking photographs and entering the required textual data) by using a mobile system, able to store the data locally or to directly transmit them to the centralised data base;

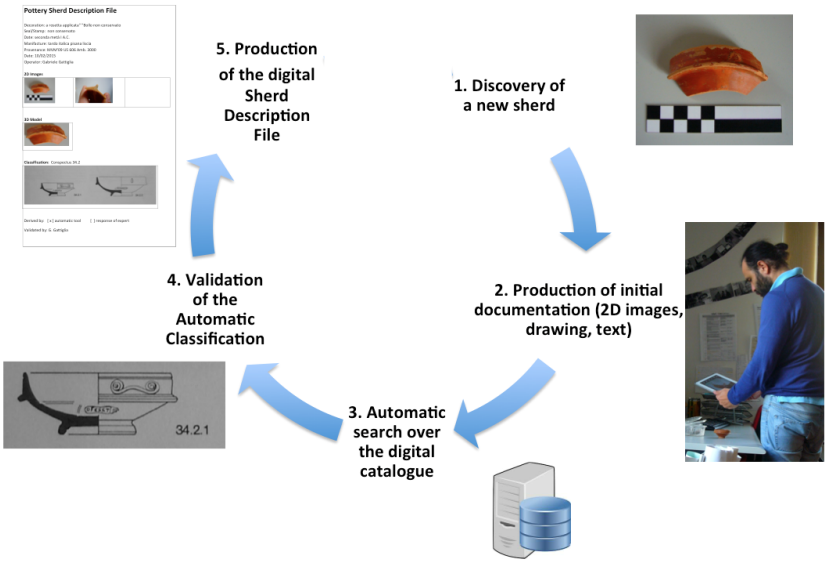

Figure 2: *ArchAIDE: towards an innovative classification and documentation cycle.*

  - For each fragment, the system offers a very intuitive and easy interface for drawing its profile (in a guided manner, following a sketch-based approach on the base of the photograph acquired); this profile will characterise the fragment;

- An automatic search for similarity is performed by comparing the fragment sketched drawing with the descriptors of the ceramic classes contained in the digitised corpus of catalogues
- The archaeologist validates the search results and takes an informed decision over the classification of the fragment;
- The result is stored on the data archive, producing a digital sherd description file (that becomes part of the archive but also a printable document).

To implement this framework we need some substantial research on algorithms:

- We will focus on technologies for the digitisation of paper catalogues, to convert paper document in digital format, firstly solving the easier problems (digitisation, OCR, detection of figures) and then focusing on the more complex issue, that is how to enable the semi-automatic conversion of raster drawings into in a geometric 2D description of each ceramic class (based on drawings and profiles);
- We will also focus on efficient and robust algorithms and data structure for supporting the search by similarity of fragments description file with respect to the database of profiles recollected from catalogues.

These algorithms, together with specific tools for interactive data presentation and annotation, will be integrated in a prototype software platform, which afterwards will be extensively tested by users and experts. This SW platform is composed by two main components:

- A system working on mobile devices (preferably on tablets) to be used as the preferred interface for data entry, sketch-based drafting of profiles and interface to the centralised data base;
- A data base front-end, giving access to the data and to the functionalities of the overall system (all but the image acquisition and profile drafting) to be used on desktop computers, with a GUI implemented as a web interface running in all common web browsers.

3. Project consortium

The project consortium includes a solid set of Human Sciences partners, who are highly representative of the academic world and of the institutions devoted to the research, investigation and archival of archaeological data. The archaeological partners of the consortium are the University of Pisa, the University of Barcelona, the University of Cologne and the University of York.
The University of Pisa (UNIPI, project coordinator) will join the consortium with the MAPPA Lab, a research unit of the Department of Civilisations and Form of Knowledge, which has an exceptional tradition of research on Mediterranean Archaeology. The MAPPA Lab has a relevant experience about archaeological method and theory, archaeological data recording and management, archaeological open data, and geo-archaeology, archaeological communication and storytelling.
The University of Barcelona (UB) has a long tradition in Archaeology. The Barcelona team Material Culture and Archaeometry is a research unit focused on promoting studies of material culture, and especially on archaeological ceramics, incorporating archaeometric approaches.
The Department of Archaeology at the University of Cologne (UCO) will join the consortium with the Cologne Digital Archaeology Laboratory, which hosts ARACHNE, a highly structured object database with more than 2 million scans representing more than 1 million datasets. A main experience of UCO is the structuring and

archiving of object metadata.
The Department of Archaeology at the University of York (UoY) hosts the Archaeology Data Service (ADS), the world-leading digital data archive for archaeology. ADS main skills are in long term preservation of digital data, and in the promotion and dissemination of a broad range of data in archaeology. The ADS promotes good practice in the use of digital data in archaeology, it provides technical advice to the research community, and supports the deployment of digital technologies. . The main strengths of the ADS are in the development of data standards and digital archiving. De facto, ADS will also play the role of linking Human Science partners with ICT partners.

In order to reach its technical goals, the consortium would need the contribution of many different technical backgrounds and skills, since we need to master the design of a complex ICT platform which will use: image processing, image understanding, semantic databases, 3D geometry representation and processing, sketch-based interfaces, shape-related search and retrieval from 2D/3D/semantic databases, development of GUIÕs for mobile platforms, web-based visualisation and cloud computing. Therefore, some key players in ICT design and development (both at the academic and industrial level) are part of the ArchAIDE consortium.
CNR-ISTI is an institute of Italian CNR devoted to research on ICT, active on several major domains that include Visual Media and Cultural Heritage. CNR-ISTI will join the consortium with its Visual Computing Lab (VCLab) research unit. Current research focuses on the design of 3D digitisation/scanning, geometric data processing, visualisation systems, multi-resolution representation and rendering of huge datasets, 3D Web Applications.
Tel Aviv University (TAU) will participate in the consortium with the school of Computer Science - Deep Learning Lab, which develops innovative research in face recognition, document analysis, image textual description, and action recognition, among other fields.

As for the industrial/commercial involvement in the project, the consortium comprises two archaeological companies and one ICT company, not only to ensure exploitation of the results, but also because their participation is consistent to achieve the expected results.
As for the latter, Inera s.r.l. (INERA) is a software company (computer technologies and services) founded in 2001 by a group of advisors and senior designers with a ten-year experience in the field of protocols and of web apps. The team includes professionals expert in the whole range of systems for the Internet and for the software products that aim at the growth and innovation of their clients.
As for the archaeological team of the project, the consortium involves also the people that most often work on the field, i.e. the private companies and professionals that are usually contracted to perform preventive and development-led archaeological investigations. These are also key components of this project, since we should consider also their view and expectations, as well as include them in the assessment phase.
Baraka ArqueÛlogos S.L. (BARAKA) is a young company (founded in 2012), but composed of archaeologists with a long career and a very solid experience in the development and practise of techniques of archaeological excavation, recording methods on excavations, analysis of stratigraphic sequences, and study of archaeological artefacts, especially ceramics.
Elements S.L. (ELEMENTS) is a private archaeological and cultural heritage consulting company founded in 2009, which has an accredited experience in development and application of digital technologies, in archaeological excavation recording, as well as on ceramic studies. In this regard, it is worthy to mention that the company has been responsible for the digital inventory of a set of ceramics (Museum of Mallorca collection), which included 3D recording and scan of potteries.

4. Project duration and plan

ArchAIDE is a three year RIA project, approved by EC under call H2020-REFLECTIVE-6-2015, grant agreement no. 693548.
The kickoff meeting has been held in Pisa on June 15-16, 2016.
The project is organised in ten WPs and the major deliverables will aim at the challenge of building a scalable and easy-to-use technology for documenting findings of archaeological excavations directly on the field, for supporting archaeologists in the subsequent interpretation phase, for the archival of all those data on remote repository and finally to providing easy access to those data to the professional and non professional community. The will focus on well-defined clear measurable objectives, challenging but realistically achievable within the duration of the project.

References

[DLS95] DURHAM P., LEWIS P. H., SHENNAN S.: Artefact matching and retrieval using the generalised hough transform. In *Proceedings of Computer Applications in Archaeology* (1995), Wilcock J., Lockyear K., (Eds.), pp. 25—30. URL: http://eprints.soton.ac.uk/250789/. 1

[GKSS04] GILBOA A., KARASIK A., SHARON I., SMILANSKY U.: Towards computerized typology and classification of ceramics. *Journal of Archaeological Science 31*, 6 (2004), 681 – 694. URL: http://www.sciencedirect.com/science/article/pii/S0305440303001808, doi:http://dx.doi.org/10.1016/j.jas.2003.10.013. 1

[KS02] KAMPEL M., SABLATNIG R.: Automated segmentation of archaeological profiles for classification. In *Proc. of 16th International Conference on Pattern Recognition, Quebec City* (2002), Kasturi R., Laurendeau D., Suen C., (Eds.), vol. 1, IEEE Computer Society, p. 57âĂŞ60. 1

[KSM02] KAMPEL M., SABLATNIG R., MARA H.: Automated documentation system for pottery. In *Proc. of 1st International Workshop On 3d Virtual Heritage, Geneva, Switzerland* (2002), Magnenat-Thalmann N., Rindel J., (Eds.), p. 14âĂŞ20. 2

[MD13] MAKRIDIS M., DARAS P.: Automatic classification of archaeological pottery sherds. *J. Comput. Cult. Herit. 5*, 4 (Jan. 2013), 15:1–15:21. URL: http://doi.acm.org/10.1145/2399180.2399183, doi:10.1145/2399180.2399183. 2

[MG05] MAIZA C., GAILDRAT V.: Automatic Classification of Archaeological Potsherds. In *The 8th International Conference on Computer Graphics and Artificial Intelligence, 3IA'2005, Limoges, France, 11/05/2005-12/05/2005* (http://www.xlim.fr, mai 2005), Dimitri P., (Ed.), Laboratoire XLIM - UniversitÃl' de Limoges, pp. 135–147. URL: ftp://ftp.irit.fr/IRIT/VORTEX/Maiza13.pdf. 1

[SKN*12] SMITH N. G., KARASIK A., NARAYANAN T., OLSON E. S., SMILANSKY U., LEVY T. E.: The <i>Pottery informatics query database</i>: A new method for mathematic and quantitative analyses of large regional ceramic datasets. *Journal of Archaeological Method and Theory* (Sept. 2012), 1–39. URL: http://dx.doi.org/10.1007/s10816-012-9148-1, doi:10.1007/s10816-012-9148-1. 2