Bringing Digital Curation to Archaeological Projects: Evidence from the BeArchaeo Project

Vincenzo Lombardo
Università di Torino
Torino, Italy
vincenzo.lombardo@unito.it

Tugce Karatas
Università di Torino
Torino, Italy
tugce.karatas@unito.it

Rossana Damiano
Università di Torino
Torino, Italy
rossana.damiano@unito.it

Claudio Mattutino
Università di Torino
Torino, Italy
claudio.mattutino@unito.it

Mariko Sasakura
University of Okayama
Okayama, Japan
sasakura@momo.cs.okayama-u.ac.jp

ABSTRACT

This paper addresses the importance of the curation of cultural heritage data that bear the content of the visual interfaces. Digital curation is an operational process that has been recently appointed as a practical method for the account of digital data that arise in cultural heritage projects. The goal of the paper is to state the importance of digital curation for the archaeological projects and to illustrate how it works in the case the Beyond Archaeology (BeArchaeo) project, an on-going European project, where all the interfaces that allow for the interactions with the digital objects will be based on a centralized database. Here we address all the phases that concern the archaeological activities, from the excavation to the exhibition of the findings, and we show how we develop the centralized data repository that undergoes the development of the communication interfaces.

CCS CONCEPTS

• **Human-centered computing** → **User interface management systems**; **Collaborative interaction**; **Human computer interaction (HCI)**; • **Information systems** → **Data management systems**.

KEYWORDS

Digital curation, Cultural heritage, Archaeology, Project workflow

ACM Reference Format:
Vincenzo Lombardo, Tugce Karatas, Rossana Damiano, Claudio Mattutino, and Mariko Sasakura. 2020. Bringing Digital Curation to Archaeological Projects: Evidence from the BeArchaeo Project. In *Proceedings of AVI²CH 2020: Workshop on Advanced Visual Interfaces and Interactions in Cultural Heritage (AVI²CH 2020).* ACM, New York, NY, USA, 4 pages.

1 INTRODUCTION

The development of visual interfaces for cultural heritage (CH) projects more and more relies on large repositories of data [1, 19]. Data originate in the digitisation of CH assets, which involve many different activities, such as data acquisition, data visualization, data analysis and interpretation, and the consequent dissemination of the results. Current technologies focus on providing digital tools for data representation, processing, and visualization, while mostly neglect the importance of analysis and sharing and fail to provide platforms for formal interoperability of researches [11, 20].

This is especially true for archaeological projects: the collaboration of diverse scientists onto the same digital data requires a unique virtual space alongside with targeted interfaces that help to document the irreversible processes of the excavations and creation of the metadata for the transparency of the reconstruction processes for physical CH objects; also, it would be desirable to provide accessibility during the processing, interpretation, dissemination, and communication of data [6].

Recently, there has been some effort in providing interoperable repositories of CH projects and collections. For example, Europeana is a web portal that features over 58 million cultural heritage items from around 4,000 institutions across Europe organized from crowdsourced collections [1]. Also, in the archaeological field, we can mention the Digital Archaeological Record (tDAR)[2], an international digital repository for the records of archaeological investigations. However, there is not an accepted systematic procedure for organizing the vast amount of digital data that are collected during an archaeological project and that are the content base for the several targeted visual interfaces.

Although most of the CH institutions have similar problems when it comes to bringing the archaeological heritage content to life, there are some one-off projects that use digital archaeological data through a variety of advanced interfaces such as web-based publishing of the collection [15], virtual environments for accessing the objects [13], or on-site museum installations [3]. Also, there are initiatives that provide an open infrastructure for the systematic creation of immersive environments for interaction purposes, such as Immersia project [7]. Yet, this occurs without an approach to the entire life cycle of the archaeological data from excavation to exhibition, although some research has improved access to archaeology libraries by using multilingual glossaries and ontologies [12], integrating and handling vast heterogeneous data [18], and using

AVI²CH 2020, September 29, Island of Ischia, Italy

[1] https://pro.europeana.eu/about-us/mission (last visited on 18 May 2020)
[2] https://www.digitalantiquity.org/wp-uploads/2011/07/20110930-Building-tDAR.pdf (last visited on 12 May 2020)

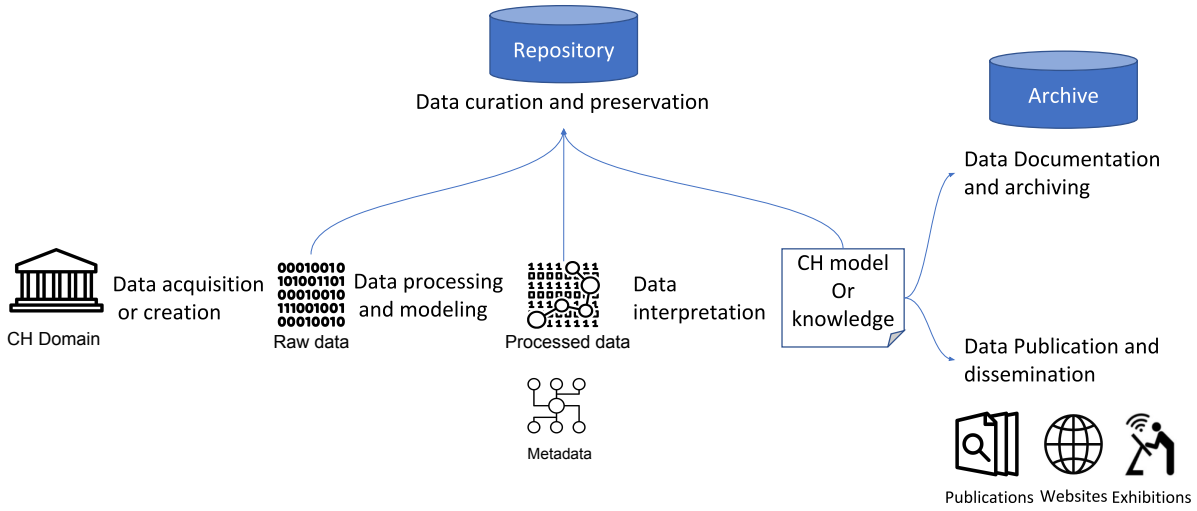

Figure 1: An abstract schema of the workflow of digital curation.

advanced user interfaces to present archaeology data [1, 19]. Finally, things are made harder by the fact that the successful completion of a project depends on interdisciplinary contributions, with different methodologies and terminologies [17].

Related to these needs, digital curation emerged recently as an operational schema for the management of CH information, addressing the challenges of the selection, preservation, maintenance, collection, and archiving of digital assets related to archaeological excavations as well as the added value for future exploitation [16, 17, 21]. However, with some noticeable exception (see below), digital curation has not been applied to archaeology. This paper proposes the systematic application of digital curation to archaeology, taking into account the activities that occur from the excavation process up to the exhibitions of the overall project as well as of the single items. There are at least three interfaces that are necessary for digital curation at various stages: 1) a back-end for each task, that allows scholars of several disciplines to input their data and metadata; 2) a front-end to allow project leaders in the several disciplines to check the database contents and to access materials related to other disciplines; 3) the interface designed for the exhibitions that present the outcomes of the project for a large audience. Here, we briefly sketch the digital curation process, with particular attention to archaeological activities, by referring to a well-known archaeological project, named the Çatalhöyük Living Archive. Then, we discuss the importance of a centralized database to support the implementation of the digital curation schema and the application of the digital curation schema to ongoing archaeological project BeArchaeo.

2 DIGITAL CURATION AND ARCHAEOLOGY

Digital curation involves "actively managing data [...] with the aim of supporting reproducibility of results, reuse of and adding value to that data, managing it from its point of creation until it is determined not to be useful, and ensuring its long-term accessibility and preservation, authenticity and integrity" (Digital Curation Center - DCC - website[3]). It supports bridging research, practice, and training across nations, disciplines, institutions, repositories, and data formats [8, 9], modelling workflows from different institutional contexts [17] and assessing the evolution of the e-resource management processes [2].

In the context of archaeology, curatorial work with digital technologies is often far less linear and scientific processes frequently cut across traditional organizational boundaries [5]. Although the tasks that compose the digital curation vary, we abstract the major archaeological activities together with their semantic interrelations, to describe how disciplinary communities can participate into the digital curation (such as, e.g., Archaeologists, Chemists, Historians, Earth scientists). These issues are mostly neglected by current approaches[14, 17].

Fig. 1 shows the major tasks that compose digital curation. For the sake of exemplification, we address the interdisciplinary, multi-team excavation of Çatalhöyük[4], the 9,000-year-old Neolithic settlement in the Konya plain of Turkey, because it has been, ante-litteram, an example of digital curation over the many years of its development. In digitally-enabled Çatalhöyük excavation, researchers have been engaging in a range of long-term activities that encompass the capture, description, annotation, classification, interpretation, knowledge enrichment and dissemination of digital documents. A number of databases have been created by diverse CH specialists who have worked in different areas during the different time periods alongside with different modalities of reporting [11]. The record of the archaeological site of Çatalhöyük was constructed dynamically, in-situ and ex-situ, as the excavation progressed, through a combination of centralized database records [4].

In general, the digital curation process of an archaeological project starts as soon as some asset is acquired (e.g., the identification of a significant stratum in the soil - a so-called *stratigraphic*

[3]http://www.dcc.ac.uk (last visited on 3 April 2020)
[4]http://catalhoyuk.stanford.edu, visited on 15 May 2020.

unit - or some *archaeological finding*, i.e. a container with a cup shape or an animal bone). As soon as this happens, some digital asset is built and recorded. The digital asset can be acquired from the asset itself (e.g., by photograph or scanning) or created from the scratch. This will produce some data, that we can generally name *raw data*, because they do not include an interpretation model yet. For example, the use of video recording equipment on Çatalhöyük allowed archaeologists to narrate their experience. These data were enriched with metadata that included one interpretation of the asset at some level (e.g., region of the video, identified via a path joining the pixels, label with the tag). Raw data and metadata are processed to enrich the data (*processed data*) and used for the interpretation process, which builds a coherent model and knowledge of the archaeological site history and findings (*CH model or knowledge* in the Figure1). All data and metadata are addressed through working interfaces that allow for the inserting and checking data in the database, while both the media produced and the interpretation results are the content that will be published through the interface designed for the final exhibition, both online and on site.

Given this prolonged usage of the data, also susceptible to revisions (theories change, and so are the interpretation models) and targeted to different goals (3d models for the interpretation are at a different resolution with respect to 3d models for the exhibition), there is a need for a centralized and active control of data and metadata.

Figure 2: The web app Çatalhöyük Living Archive.[10]

The data interpretation part of the digital curation process not only allows the researchers to use appropriate standards and working interfaces, but also supports the development and the evolution tracking of the database. For example, unit sheets and plans, once completed and checked by the area supervisor, were immediately digitized and inserted into the site database and became available to all on-site researchers. Once the field season was over, the database information was incorporated into the official Çatalhöyük website and then available to external research (Fig 2). The Living Archive project of Çatalhöyük has then been used in many publications (e.g. paper, conference presentations, field reports etc.) and exhibitions, e.g., "The Curious Case of Çatalhöyük" exhibition in Anamed (the Koç University Research Center for Anatolian Civilizations in Istanbul, Turkey)[5]. The latter is one example of the third type of interface required for a modern archaeological project.

3 THE BEARCHAEO PROJECT

The digital curation framework is being applied in an ongoing EU project named Beyond Archaeology (BeArchaeo)[6] which consists in the excavation of the archaeological site, the archaeometric analyses of the site and the excavated materials, and the interpretation/dissemination of the results about the Tobiotsuka Kofun (Soja city in Okayama Prefecture), together with other Kofun burial mounds and the related archaeological material in ancient Kibi and Izumo areas (present Okayama and Shimane Prefectures), in Japan. A preliminary achievement of this research has been the design and implementation of a semantic database for the encoding and storing of the digital data concerning the archaeological excavation and addressing the metadata belonging to the several disciplines though concerning the same cultural heritage object (typically, a stratigraphic unit or an archaeological finding)[7]. The project aims to develop expertise and skills under new perspectives, to enable a transdisciplinary research from the archaeological site to the museum display.

The data base design has followed a Semantic Web approach. A domain ontology has been realized for the major categories that appear in the archaeological excavations, namely stratigraphic units and archaeological findings. The knowledge has been acquired from the major documentation sources available with respect to the forms that have been in use for decades by the archaeological teams. These sources are mostly published by the national organizations of cultural heritage (see, e.g. the documentation records of the Italian Central Institute for the Catalogue and the Documentation[8]). Also, the BeArchaeo ontology has been aligned with international thesauri (e.g., Getty AAT[9]) and high–level domain ontology CIDOC-CRM and its collaboration family[10]). The major semantic properties were imported into an installation of the web-based Content Management System Omeka-S[11], which revealed to be useful for the definition of forms to be filled by the archaeologists in the field, deployed as "Resource Templates". These constitute the interface for inputting of data on behalf of the archaeologists (soon, other interfaces will be designed and implemented). For this interface, Omeka-S provides the possibility of the fast prototyping of several user interfaces for the back-end of the system. Omeka-S also provides an easy way to create the front-end (second type of) interface, where supervisors and stakeholders could explore the development of the archive and the related findings (Fig.3). Also, CMS Omeka-S has allowed for the uploading of rich media materials (currently photos and 3D models acquired from photogrammetry and scanning).

Considering the multi-cultural and multi-lingual issues of the BeArchaeo project, knowledge interoperability between Japanese and English researchers as well as data terminology is paramount.

[5] https://anamed.ku.edu.tr/en/events/bir-kazi-hikayesi-catalhoyuk-londrada/ (last visited on 25 May 2020).
[6] https://www.bearchaeo.com/ (last visited on 15 May 2020).
[7] https://bearchaeo.unito.it/omeka-s (last visited on 15 May 2020).
[8] http://www.catalogo.beniculturali.it/sigecSSU_FE/, in Italian.
[9] https://www.getty.edu/research/tools/vocabularies/aat/
[10] http://www.cidoc-crm.org/collaborations
[11] https://omeka.org/s/ (last visited on 15 May 2020)

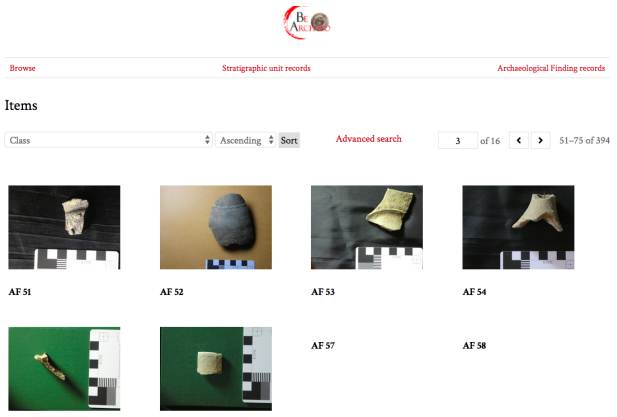

Figure 3: View of BeArchaeo Omeka-S built website.

In order to ease the insertion of data and metadata from Japanese scientists and collaborators as well as avoid misinterpretations due to linguistic terminology, we have provided Japanese Resource Templates, i.e. the back-end interface for the Archaeological Finding and Stratigraphic Unit records, respectively. For checking purposes, the current development website distinguishes between data inserted by the Japanese and European scientists[12]. This because, we are still in the phase of checking the input data and aligning descriptions. However, we are going to unify the descriptions, in order to allow for a terminological search beyond the linguistic and cultural barriers.

4 CONCLUSIONS

We have presented an account of digital curation for archaeological projects and we have given a preliminary account of digital curation as employed in an ongoing project, named BeArchaeo. Claiming that the development of a functioning digital curation pipeline is of benefit for the field of archaeology, we take the holistic perspective of digital curation from the excavation process to the organisation of the exhibition, by implementing and maintaining a centralized database.

The BeArchaeo project, having a notable inspiration from a previous pioneering work, is going to systematize this approach and to provide a viable system for the global supervision of digital data that are generated during the very long lifetime of an archaeological project. The convergence of many disciplines on single objects as well as the usage of the same data in many interfaces provides us a stimulating challenge for the future of cultural heritage management and communication.

In particular, we are going to address a deep analysis of the potential disciplinary targets, to propose advanced specific user interfaces to the centralized data in relation to cultural heritage interpretation and for exhibition purposes.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 754511. This research is also related to a project called Beyond Archaeology (BeArchaeo) which is funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie, Grant Agreement No 823826.

[12]https://bearchaeo.di.unito.it/omeka-s/s/jtoppage/page/welcome

REFERENCES

[1] Daniel Acevedo, Eileen Vote, David Laidlaw, and Martha Joukowsky. 2001. Archaeological Data Visualization in VR: Analysis of Lamp Finds at the Great Temple of Petra, a Case Study. https://doi.org/10.1109/VISUAL.2001.964560
[2] Janna Quitney Anderson, Lee Rainie, Martha M. Batorski, and D Hadden. 2010. The impact of the Internet on institutions in the future. Pew Research Center's Internet American Life Project.
[3] Kim Bale, Daisy Abbott, Ramy Gowigati, Douglas Pritchard, and Paul Chapman. 2011. Linking Evidence with Heritage Visualization using a large Scale Collaborative Interface. 121–128. https://doi.org/10.2312/VAST/VAST11/121-128
[4] Asa Berggren, Nicolo Dell'Unto, Maurizio Forte, Scott Haddow, Ian Hodder, Justine Issavi, Nicola Lercari, Camilla Mazzucato, Allison Mickel, and James Taylor. 2015. Revisiting reflexive archaeology at Catalhoyuk: Integrating digital and 3D technologies at the trowel's edge. *Antiquity* 89 (04 2015), 433–448. https://doi.org/10.15184/aqy.2014.43
[5] Maria Collins. 2009. Evolving Workflows: Knowing when to Hold'em, Knowing when to Fold'em. *The Serials Librarian* 57, 3 (2009), 261–271. https://doi.org/10.1080/03615260902877050
[6] Maurizio Forte, Nicoló Dell'Unto, Justine Issavi, L Onsurez, and Nicola Lercari. 2012. 3D ARCHAEOLOGY AT ÇATALHÖYÜK. *rnal International Journal of Heritage in the Digital Era* 1 (02 2012). https://doi.org/10.13140/2.1.5120.0649
[7] Ronan Gaugne, Valérie Gouranton, Georges Dumont, Alain Chauffaut, and Bruno Arnaldi. 2014. Immersia, an open immersive infrastructure: doing archaeology in virtual reality. *Archeologia e Calcolatori* (01 2014), 180–189.
[8] Anna Gold. 2010. Data Curation and Libraries: Short-Term Developments, Long-Term Prospects. *Office of the Dean (Library)* (01 2010).
[9] Ray Joyce. 2009. Sharks digital curation and the education of information professionals. *Museum Management and Curatorship* 24, 4 (2009), 357–368. https://doi.org/10.1080/09647770903314720
[10] Dominik Lukas, Claudia Engel, and Camilla Mazzucato. 2018. Towards a Living Archive: Making Multi Layered Research Data and Knowledge Generation Transparent. *Journal of Field Archaeology* 43 (10 2018), 19–30. https://doi.org/10.1080/00934690.2018.1516110
[11] Michael Ashley López, Ruth Tringham, and Cinzia Perlingieri. 2011. Last House on the Hill: Digitally Remediating Data and Media for Preservation and Access. *Journal on Computing and Cultural Heritage (JOCCH)* 4, 109–116. https://doi.org/10.1145/2050096.2050098
[12] Carlos Monroy, Richard Furuta, and Filipe Castro. 2010. Using an ontology and a multilingual glossary for enhancing the nautical archaeology digital library. 259–262. https://doi.org/10.1145/1816123.1816162
[13] Colleen Morgan. 2009. (Re)Building Çatalhöyük: Changing Virtual Reality in Archaeology. *Archaeologies-journal of The World Archaeological Congress* 5 (12 2009), 468–487. https://doi.org/10.1007/s11759-009-9113-0
[14] Maureen Pennock. 2007. Digital curation: A life-cycle approach to managing and preserving usable digital information. *Library and Archives* 1 (01 2007).
[15] Sofia Pescarin, Enzo d'Annibale, Bruno Fanini, Mohamed Farouk, Daniele Ferdani, Wim Hupperetz, Mohamed Maguid, Adnan Muftaveric, Niall O'hOisin, Breffni Malley, Dries Nollet, Alfonsina Pagano, Augusto Palombini, Daniel Pletinckx, Leonardo Rescic, Patrick Reuter, Selma Rizvić, Daria Ruggeri, Lucrezia Ungaro, and Paolo Vigliarolo. 2014. *Keys To Rome. Roman Culture, Virtual Museums.* https://doi.org/10.13140/2.1.2631.5046
[16] Alex Poole. 2016. The Conceptual Landscape of Digital Curation. *Journal of Documentation* 72 (09 2016). https://doi.org/10.1108/JD-10-2015-0123
[17] Colin Post, Alexandra Chassanoff, Christopher A Lee, Andrew Rabkin, Yinglong Zhang, Katherine Skinner, and Sam Meister. 2019. Digital Curation at Work: Modeling Workflows for Digital Archival Materials. In *Proceedings of ACM/IEEE Joint Conference on Digital Libraries*. ACM Press, 39–48.
[18] Unni Ravindranathan, Rao Shen, Marcos Gonçalves, Weiguo Fan, Edward Fox, and James Flanagan. 2004. ETANA-DL: a digital library for integrated handling of heterogeneous archaeological data. 76–77. https://doi.org/10.1145/996350.996370
[19] Stan Ruecker, Milena Radzikowska, and Stéfan Sinclair. 2011. *Visual interface design for digital cultural heritage: A guide to rich-prospect browsing.* Ashgate. 1–197 pages.
[20] Weiqi Shi, Eleni Kotoula, Kiraz Göze Akoğlu, Ying Yang, and Holly Rushmeier. 2016. CHER-Ob: A Tool for Shared Analysis in Cultural Heritage.
[21] Elizabeth Yakel, Paul Conway, Margaret Hedstrom, and David Wallace. 2011. Digital Curation for Digital Natives. *Journal of Education for Library and Information Science* 52 (01 2011), 23.