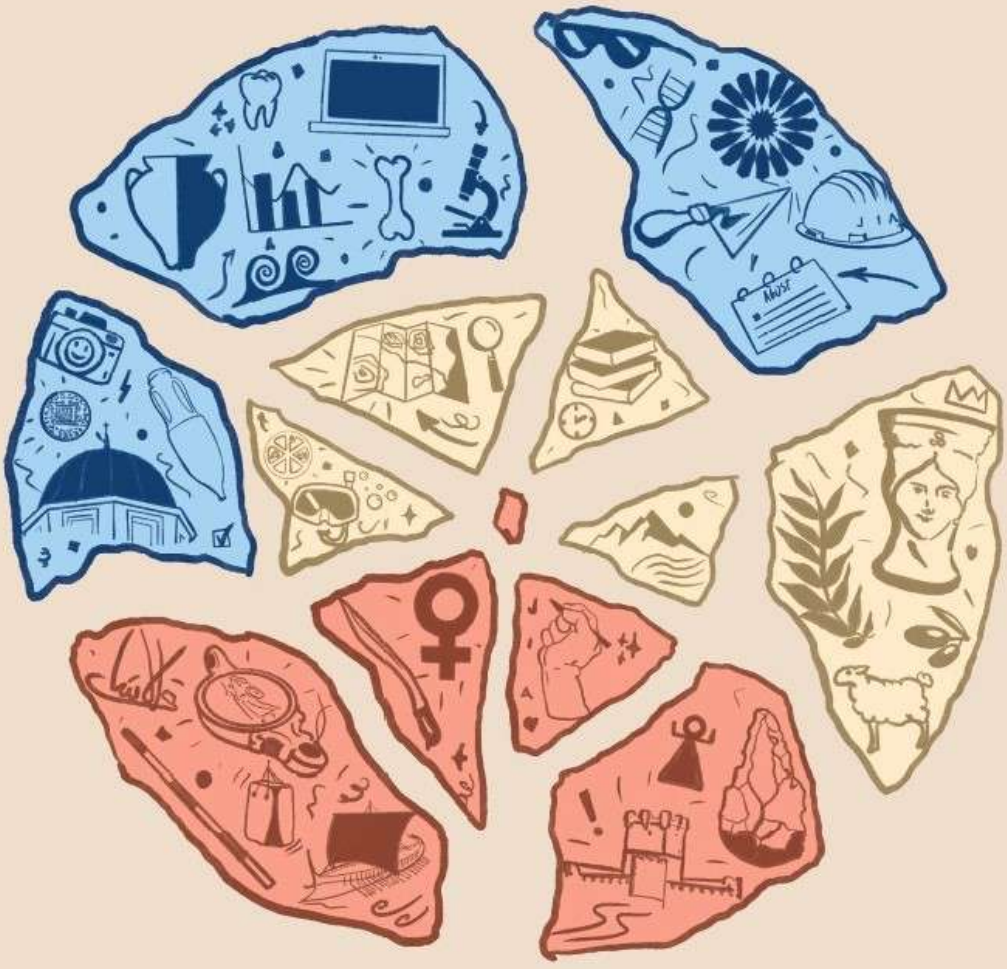
inaph

COLECCIÓN PETRACOS

Periferias: desde los márgenes de la arqueología

Sonia Carbonell Pastor, María Fructuoso Cárcel, Arturo García López, Paula Martín de la Sierra Pareja, José Luis Martínez Boix, Ester Moya Soriano, Nicolás Pastor Alameda, Pedro Ramón Baraza, Israel Serna Martínez (Eds.)



PUBLICACIONES INAPH
Universidad de Alicante

INAPH
COLECCIÓN *PETRACOS* 12

Periferias: desde los márgenes de la arqueología

SONIA CARBONELL PASTOR, MARÍA FRUCTUOSO CÁRCEL,
ARTURO GARCÍA LÓPEZ, PAULA MARTÍN DE LA SIERRA PAREJA,
JOSÉ LUIS MARTÍNEZ BOIX, ESTER MOYA SORIANO,
NICOLÁS PASTOR ALAMEDA, PEDRO RAMÓN BARAZA,
ISRAEL SERNA MARTÍNEZ
(EDS.)

Periferias: desde los márgenes de la arqueología

JIA-LACANT 2022
Actas de las XIII Jornadas de Jóvenes en Investigación Arqueológica (Universidad de Alicante, 23-26 de marzo de 2022)

PETRACOS es una publicación de difusión y divulgación científica en el ámbito de la Arqueología y el Patrimonio Histórico, cuyo objetivo central es la promoción de los estudios efectuados desde el Instituto Universitario de Investigación en Arqueología y Patrimonio Histórico de la Universidad de Alicante –INAPH–. *Petracos* también pretende ser una herramienta para favorecer la transparencia y eficacia de la investigación arqueológica desarrollada, transfiriendo a la sociedad el conocimiento generado con la mayor rigurosidad posible. Esta serie asegura la calidad de los estudios publicados mediante un riguroso proceso de revisión de los manuscritos remitidos y el aval de informes externos de especialistas relacionados con la materia, aunque no se identifica necesariamente con el contenido de los trabajos publicados.

Edita: Instituto Universitario de Investigación en Arqueología, Patrimonio Histórico (INAPH) y Publicacions Universitat d'Alacant

Imagen de cubierta:
José Luis Martínez Boix

ISBN: 978-84-978-84-1302-220-8
Depósito legal: A 246-2023
Diseño y maquetación: Marten Kwinkelenberg
Imprime: Byprint Percom S.L
Impreso en España

New perspectives on Digital Archaeology: From production to usability of data

Sonia Medina Gordo[1], Xènia Fructuoso Fernández[2], Guillem Domingo Ribas[3], Tania Freixas Roig[4] and Sabina Batlle Baró[5]

Abstract

The study of the material past through current digital methods and tools is not only highlighting the challenges of handling them correctly, but also how this context is influencing research in Archaeology, and especially the data produced within this discipline. In this paper, we introduce this topic discussing various issues, such as the information digitisation, the relevance of free and open-source software, the responsibility of data opening, and the perception of data for knowledge acquisition.

Keywords: digital data; data preservation; Open Archaeological Science; FOSS; archaeological knowledge

Resumen

El estudio del pasado material a través de los métodos y las herramientas digitales actuales no solamente está resaltando los retos que supone el correcto manejo de los mismos, sino también cómo este contexto está influyendo la investigación en Arqueología, y especialmente con los datos que esta disciplina produce. En este trabajo presentamos dicha situación discutiendo diversas cuestiones, como la digitalización de la información, la relevancia de los programas libres y de código abierto,

1. Universitat de Barcelona. Departament d'Història i Arqueologia. Institut d'Arqueologia. sonia.medina@ub.edu
2. Fundació UAB. Escola Superior d'Arxivística i Gestió de Documents (ESAGED). xenia.fructuoso@autonoma.cat
3. Newcastle University. School of History, Classics and Archaeology. McCord Centre for Landscape. g.domingo-ribas2@newcastle.ac.uk
4. Grup de Recerca en Arqueologia Protohistòrica (GRAP). tfreixas.histarq@gmail.com
5. Universitat de Barcelona. Departament de Biblioteconomia, Documentació i Comunicació Audiovisual. sabina.batlle@ub.edu

la responsabilidad en la apertura de los datos, y la percepción de los mismos para la adquisición de conocimiento.

Palabras clave: datos digitales; preservación del dato; ciencia arqueológica abierta; FOSS; conocimiento arqueológico.

1. Introduction

In the last decades, the employment of computational tools for the quantification of data, together with the use of algorithmic techniques or models for exploration and analysis of data has significantly modified the way we observe and work with the past in Archaeology. Indeed, the study of material vestiges within digital environments has helped us to planning projects and becoming more familiar with past realities from multiple perspectives, or even more cross-disciplinary and collaborative approaches.

These *other forms* of approaching and digitally handling the archaeological record involve a whole process of data production that comprises different stages. Aspects such as the capture, representation, preservation, access, exploitation, and dissemination of the information generated are only a small sample of a wide-ranging procedure that is indispensable during the processing of the data. Accordingly, we consider that in our task as generators and consumers of an ever more datified knowledge we need to address and debate new approaches and trends that are being projected in the field.

In the following sections, we will provide a general overview of the principal topics discussed during the second roundtable organised within the 13th JIA. The data life cycle represents the sequence that particular units of information undergo, ranging from their production to the end of their useful life. Hence, and taking into consideration the context of the so-called Digital Archaeology, the creation of digitalized archaeological record for its preservation and renovation, the benefits of using free and open-source software, the implications, and responsibilities for creating Open Data scenarios, along with the challenges of capturing data for knowledge gain will be the main topics of this paper.

2. Digital transformation of the archaeological fieldwork

As generally assumed, one of the most essential parts of archaeological data production comes from the archaeological record produced during fieldwork (Parcero-Oubiña et al.1999). This sort of evidence is mainly obtained during the survey or excavation of an archaeological site, being its material testimonies, the context, and the process of information management its main components, as well as the principal elements for the subsequent analysis and interpretation of data.

It seems common for many archaeologists that the textual and graphic recording work that will form part of this archaeological record is still done by hand. Such is the case with the plan and section drawings or the recording sheets of the Units of Stratigraphy. Therefore, the issue that is important to highlight at this point is the very preservation of all this knowledge that we produce by analogy, more so when we are aware of the destructive nature of archaeological data.

Thus, preserving these records optimally is a matter of necessity. Moreover, much of the produced information is as important, or even more, as the actual material recovered straight from the field. And it is so due to the context, which enables us to get an in-depth insight of and value each piece of the site.

Despite digitisation being the best choice for turning those analogous records into digital formats, and thus improving the chances of data preservation, it is a process in which not all research groups or private companies invest. This is probably owing to lack of money or time constraints. The shift from one to another format may result in a series of affordances and constraints, which define a whole data journey and frictions (Huggett, 2022: 282-284) that are not probably well-suited to everyone's necessities or options.

Nonetheless, we believe that dedicating time to exploring the possibilities for *re-contextualising* analogue archaeological records can make a worthwhile contribution to make the most of the digitization work. Richards-Rissetto and Landau's (2019) work is a good example of this conviction. These authors expose their perspectives on translating analog data to geospatial digital data, georeferencing scanned sketch maps and moving beyond a mere vectorization. Likewise, in other recent publications, R. Opitz (2018) communicated her experience in producing an excavation report that not only reconsiders the way we face so-called Public Archaeology by making digital content more appealing and user-friendly, but it also reformulates the idea of the excavation monograph.

3. Free and open-source software in Archaeology: Relevance and implications

Following the digital transformation of the archaeological fieldwork and research, both have benefited from the development of geospatial tools. Geographical Information Systems (GIS) and remote sensing technologies have become the standard tools for recording and managing relevant natural and cultural features in the landscape. In this line, the study of past social behaviours and its spatial implications has led professionals to the use of GIS tools, which have overtaken data collection and management and have become one of the most relevant tools in archaeological works.

However, GIS platforms are not exempt from issues. The usage of geospatial tools can be hindered by software companies' licences due to the limited budget of

most archaeological research or commercial archaeological companies, particularly in developing nations. Besides, research can also be limited by the functionalities offered by a particular commercial GIS platform that sometimes cannot enable customisable and functional enhancements (Steiniger and Hay, 2009). In turn, this type of software can act, as named by Morin *et al.* (2012), like a 'black box' in research: users might not be able to know how the data processing is performed because of the unavailability of the software's source code, replicating implementation errors, while going against the reproducibility principle from most publicly funded research.

Alternatively, free and open-source software (FOSS) represent an outstanding alternative to software licensing, data acquisition and tool development, breaking barriers for underfunded archaeologists (Carter, 2019). Moreover, FOSS does not only imply saving money. *'Free'* also stands for the liberty of each user to install, modify, and share the software when necessary, making a significant contribution to the sustainable development of software and good scientific practice in Archaeology (Ducke, 2012). Users can now share and reproduce existing computational tools and analysis, building new approaches from already developed and tested algorithms, and they can also have different copies of the software in different computers at no extra cost.

The transparency of FOSS provided by the full availability of the source code also contributes to the transparency required by good scientific practice. The existence of a vast online community of users implies a thorough revision of the source code, which allows the release of new versions fixing programming errors and software limitations. Moreover, users will benefit from free learning materials and online tutorials that can pave the way for the use of new tools: the ease of accessibility has also led to the development of numerous extensions or modules that broaden the capabilities of existing senior geospatial platforms, even enhancing cross-platform approaches in some cases (Orengo, 2015).

4. Who owns Archaeology? Towards an openness of data

Several researchers advocating for data openness emphasise the many benefits that this new model can bring. These include improvements in the scientific methodology (more transparency, more collaboration), increased visibility of research and results, and, most relevant to archaeology, chances for data reuse and preservation (Kansa et al. 2013; Richards and Moore, 2015).

Considering such an idyllic situation, one wonders why it is not already the prevailing model in the scientific world. According to the results of a survey conducted by the European project ARIADNEplus in 2019 (Geser, 2019), the main reasons that slow down the sharing of archaeological data are the lack of reward (unlike publications), the effort required to prepare those data for sharing, the fear of infringing intellectual property rights, and the lack of adequate repositories. Additionally,

other aspects hindering data openness include the fact that it is not a mandatory nor common practice amongst archaeologists, or the fear of losing the scoop on the data.

However, one aspect that most of the studies leave aside is the responsibility in the sense of defining who should be responsible for making sure that these data are shared (ensuring access) correctly (ensuring their quality) and maintained over time (ensuring their preservation). In a way, the mandate of transparency (submission of archaeological reports to the public administration) already ensures some level of data sharing, and if we accept that data are not ours but the administration's, they could make clear regulations that would solve our doubts about sharing rights.

On the other hand, the dilemma gets clearer when we address issues such as the long-term preservation of data. Due to the destructive nature of excavation methodologies, data are the last remaining evidence of Archaeology's subject of study (Huggett, 2015). This situation makes data preservation and reuse crucial for the present and future of the field. And despite not being a first need for the science advance (Borgman, 2015), open data seems to be able to help pave the way for both issues (Kansa et al. 2013; Richards and Moore, 2015).

However, if we are to learn from the barriers identified until now, we believe that it might be the public administration the one taking the leading role in this movement for data sharing. Policy development for the promotion of archaeological data openness would be desirable, with available funding accordingly; implementing reliable and quality public repositories, and ensuring the long-term curation and preservation of those data that are part of the public historical and archaeological heritage.

5. The role of cognition in data capture and knowledge discovery

The shift from traditional to digital capture methods has resulted in many scholars pointing out how archaeologists involve with digitised information. One of the most striking debates are being conducted around how we are *perceiving* this type of data for dealing archaeologically with a past whose only vestiges are its material remains, yet it may be presented in different formats. By attending this issue, some experts acknowledge a dichotomy in digital archaeological data, thus distinguishing between two basic notions; what is understood as *data,* and what is conceived as *capta* (Chippindale, 2000; Huggett, 2022: 274-276).

This difference has also led to recognise two distinct ways of processing information so as to approaching the past. On the one hand, archaeologists that try to handle a great deal of information as *raw data*, collected right away from either archaeological sites, legacy data, remote sensed models, or grey literature reports, will likely think quantitatively. This does not mean them using quantitative methods for enquiring data but having the tendency to analyse and think about each phenomenon through a lens of measurable information. In such a way, prospects for the

use of data are encouraged whilst limited to those that can only be mathematised, together with the issues these outlooks entail (Ammerman, 1992: 242-245).

Conversely, perspectives on the management of a *theory-laden data* seem to put effort into the application of the data modelling approach. In Information Science, the scope of this practice ranges from the schematisation of taxonomies for categorising information, to the creation of ontologies for representing abstractly the reality under examination, being CIDOC CRM the best-known and most widely used conceptual model among humanists and heritage curators.

In the light of this panorama, we might well go one step further by assessing the very complexity of information. For instance, Tobalina-Pulido and González-Pérez (2020) brought forward the imperfection of archaeological data, thus proposing a theoretical framework based on the concept of *vagueness*, and the way we portray and exploit it. Elsewhere, Huggett (2020) illustrated the different dimensions of ignorance, discussing how we can deal with gaps or inconsistencies in sources of information. Analogously, Thibodeau (2021: 3) delved into the issue by stating the cognitive disparity that emerges from a misconception between information *about* the past and information *from* the past, as the employed data are conceptually different.

6. Concluding remarks

Throughout this work we have been able to explore different issues, such as the conversion of archaeological data into digital formats, not only for its conservation, but also to broaden its possibilities of use within the so-called "digital turn".

In this line, we also exposed how the availability of functional free and open geospatial computational tools, covering most of the needs of professional archaeologists and researchers, is easing the sharing and reproducibility of software and analysis within the field.

In the same way, these perspectives of digitisation of data and their exploitation through free, open-use tools lead us to consider other equally relevant questions, such as the openness of this information. More specifically, we tried to elucidate on whom this task should fall.

Finally, considering a new research framework in which the archaeological record and digital tools converge and have become commonplace in the daily lives of many archaeologists, we addressed the perception of data in archaeology for the acquisition of knowledge. In doing so, we had to consider the disparity of points of view depending on the theoretical background of researchers, and the gaps that still exist in order to manage this knowledge as well.

7. Acknowledgements

This paper is an outcome of different PhD grants, supported by the Secretariat for Universities and Research of the Ministry of Business and Knowledge of the Government of Catalonia and the European Social Fund (2022 FI_B1 00021); Newcastle University (United Kingdom) through the NUAcT's Pastoralism and Landscape Sustainability (PLAS) Project; and the Ministerio de Ciencia e Innovación (PRE2019-091286).

The authors wish to thank Dr Esther Travé Allepuz for her helpful comments and suggestions.

8. Bibliography

Ammerman, A. J. (1992): Taking stock of Quantitative Archaeology. *Annual Review of Anthropology*, 21: 231-255.

Borgman, C. L. (2015): If Data Sharing is the Answer, What is the Question?. *ERCIM News*, 100: 15-16.

Carter, B. P. (2019): Identifying Landscape Modification using Open Data and Tools: The Charcoal Hearths of the Blue Mountain, Pennsylvania. *Historical Archaeology*, 53: 432-443. https:// doi:1007/s41636-019-00171-1.

Chippindale, C. (2000): Capta and data: On the nature of Archaeological Information. *American Antiquity*, 65, 4: 205-612. https://doi:10.2307/2694418.

Ducke, B. (2012): Natives of a connected world: free and open source software in archaeology. *World Archaeology*, 44(4): 571-579. https://doi:10.1080/00438243.2012.743259.

Geser, G. (ed.) (2019): ARIADNEplus. Community Needs Survey 2019. Report. [URL: http://ariadne-infrastructure.eu/wp-content/uploads/2019/11/ARIADNEplus-Survey-2019-Report.pdf]. Accessed on 17/05/2022.

Huggett, J. (2015): Challenging Digital Archaeology. *Open Archaeology*, 1: 79-85.

Huggett, J. (2020): Capturing the Silences in Digital Archaeological Knowledge. *Information*, 11: 278. https://doi:10.3390/info11050278.

Huggett, J. (2022): Data Legacies, Epistemic Anxieties and Digital Imaginaries in Archaeology. *Digital*, 2: 267-295. https://doi:10.3390/digital2020016.

Kansa, E. C., Kansa, S. W. & Goldstein, A. L. (2013): On ethics, sustainability and open access in archaeology. *The SAA Archaeological Record*, 13(4): 15-22.

Morin, A., Urban, J., Adams, P. D., Foster, I., Sali, A., Baker, D. & Sliz, P. (2012): Shining light into black boxes. *Science,* 336: 159-60. https://doi:10.1126/science.1218263.

Opitz, R. (2018): Publishing Archaeological Excavations at the Digital Turn. *Journal of Field Archaeology*, 43:sup1, 68-82. https://doi:10.1080/00934690.2018.1505409.

Orengo, H. A. (2015): Open source GIS and geospatial software: towards their integration into everyday archaeological practice. In Wilson, A. T. and Edwards, B. (eds.): *Open source archaeology. Ethics and Practice*. De Gruyter: 64-82. Warsaw. https:// doi:10.1515/9783110440171-006.

Parcero-Oubiña, C., Menéndez-Fernández, F & Blanco-Rotea, B. (1999): El Registro de la Información en intervenciones arqueológicas. *CAPA 9*. Laboratorio de Arqueoloxía e Formas Culturais. Santiago de Compostela.

Richards, J. & Moore, R. (2015): Here Today, Gone Tomorrow: Open Access, Open-Data and Digital Preservation. In A. T. Wilson and B. Edwards (eds.): *Open source archaeology. Ethics and practice*. De Gruyter: 30-43. Warsaw.

Richards-Rissetto, H. & Landau, K. (2019): Digitally-Mediated Practices of Geospatial Archaeological Data: Transformation, Integration, & Interpretation. *Journal of Computer Applications in Archaeology*, 2(1): 120-135. https://doi:10.5334/jcaa.30.

Steiniger, S. and Hay, G. (2009): Free and open source geographic information tools for landscape ecology. *Ecological Informatics*, 4 183-195. https://doi:10.1016/j.ecoinf.2009.07.004.

Thibodeau, K. (2021): Discerning Meaning and Producing Information: Semiosis in Knowing the Past. *Information*, 12, 363. https://doi:10.3390/info12090363.

Tobalina-Pulido, L.; González-Pérez, C. (2020): Valoración de la calidad de los datos arqueológicos a través de la gestión de su vaguedad. Aplicación al estudio del poblamiento tardorromano. *Complutum*, 31(2): 343-360. https://doi:10.5209/cmpl.72488.