Bridging Digital Approaches and Legacy in Archaeology

Markos Katsianis 1,*, Tuna Kalayci 2 and Apostolos Sarris 3

1 Department of History and Archaeology, University of Patras, GR26504 Patras, Greece
2 Department of Archaeological Sciences, Faculty of Archaeology, Leiden University, 2333 CC Leiden, The Netherlands
3 Digital Humanities Geoinformatics Lab., Archaeological Research Unit, Department of History and Archaeology, University of Cyprus, Nicosia 1678, Cyprus

* Correspondence: mkatsianis@upatras.gr

1. Introduction

The emergence of the ubiquitous digital ecosystem has provided new momentum for research in archaeology and the cultural heritage domain. Digital methods have changed the way we conduct archaeological research and opened new paths for thinking and writing about the past. Digitalization, however, is informed by the available technology and the ever-changing socio-technical landscape. Such circumstances fill the ecosystem with drawbacks, promises, and possibilities. Conversely, archaeological research and knowledge generation are historically situated events. “Digital archaeology” has embraced both ends of the spectrum and created novel challenges.

One of these challenges relates to the speed of change in digitalization, which is creating outdated data, antiquated machinery, and obsolete workflows in increasing proportions, alongside new data production practices harvesting alternative forms of archaeological labour (Casaroto, this issue). Datasets, data collections, documentation materials, and research outputs have been created using different technological solutions that encompass variable levels of traditional and digital methods and equipment. The constraints in analogue recording, the numerical grounding of early computing, and the limitations in early and more recent hardware and software have resulted in datasets that may or may not be acceptable for use by the digital methods of the present and the future. These factors contribute to what is broadly known as the “legacy” issue.

These “digital legacies” have been piling up faster than they are being integrated, affecting our ability to reflect on the sustainability of our digital products and re-use potential [1]. Therefore, we wonder to what extent new forms of data processing can maintain, utilize, or even enhance existing datasets and open new paths to creative digital representations and interpretations of the past. Are there any ways to escape the rigidity of legacy or at least transform legacy to work in the so-called “fourth paradigm” of data-intensive scientific discovery [2,3]?

This Special Issue invited contributions that described successes or failures when dealing with already-compiled research datasets and documentation materials in both analogue and digital formats. Themes included the tackling of data absence and uncertainty in traditional research archives, computational approaches to harvesting analogue or digital data, issues in using legacy data with big data analytics, the prospects and limitations in AI legacy processing, examples of establishing provenance when moving from analogue to digital modes, digital data reuse in research and beyond, and legacy data augmentation procedures. We further attempted to widen the scope of “legacy” and initiate a discussion around alternative facets, such as digital material culture including obsolete machinery and discontinued peripherals, political legacies in and of digital technology, the economics of sustaining digital legacies, and missed or disrupted digital legacies in the form of workflows, technologies, personal experiences, and proficiency.
2. Studying Digital Legacies

Digital legacies are abundant in every facet of the research conducted in the digital realm. Although archaeology is a discipline that quickly embraced information technology and computers [4], and the digital has infiltrated almost all aspects of archaeological practice, like in the wider world, we still experience a palimpsest or a mix-and-match of analogue and digital data or practices that can be linked to skeuomorphic processes of technological adoption [5].

This was best summarized in the presentations at the Computer Applications and Quantitative Methods in Archaeology conference, which turns 50 next year. Looking through the different topics that were addressed during this particular conference, one is able to see the progress in how the information technologies (ITs) and methods have been employed and how they have evolved in archaeological research. These include digital archiving, database schemas and metadata standards, virtual reconstructions and museum exhibitions, GIS spatial analyses, computational methods and modelling, web-based and mobile applications, satellite and aerial remote sensing, and the emerging machine learning paradigm, all of which contribute immensely to the way that archaeological research and cultural heritage management is practiced today.

Although we define digital legacies as a wide array of things including data, machinery, software, workflows, technologies, collective understanding, and personal experiences, the first idea to receive academic attention was that of data. Internet Archaeology, as early as 2008, provided a volume on legacy data where several publications discussed their findings (see [6]; the entire volume is accessible in https://intarch.ac.uk/journal/issue24/index.html accessed on 10 October 2022). Analogue data became digital and is perhaps the most researched and theorised aspect of archaeological legacies so far, e.g., [7]. On most occasions, these data transformations are made with a view to enhancement, re-analysis or re-interpretation using novel IT tools and methods. Such aspects include studies in the grey literature [8], excavation datasets [9–13], survey and CRM records [14–19], archival material [20], retrospective photogrammetry [21,22], and data harvesting and modelling [23,24].

However, even with this background, a substantial body of digital data has been generated with the prospect of becoming legacy itself due to the constant development of digital procedures, tools, and methods. The quick evolution of digital tools creates research archives that are “open and in flux” [25] or “legacy in the making” [26]. The problem of digital obsolescence has been highlighted for a long time [27], and efforts have been directed towards the creation of digital repositories for the storage, achieving, and accessibility of older (and current) datasets that could provide reuse potential from a wide range of users (see, for example, national repositories such as the ADS (Archaeology Data Service: https://archaeologydataservice.ac.uk/, accessed on 10 October 2022) in the UK, DANS-KNAW (Data Archiving and Networking Services, Royal Netherlands Academy of Arts and Sciences: https://dans.knaw.nl/en/, accessed on 10 October 2022) in the Netherlands or the SND (Swedish National Data Service: https://snd.gu.se/, accessed on 10 October 2022) in Sweden. Archaeological data survival and availability has since become an acknowledged research objective, and several transnational projects have been commenced in order to coordinate and integrate relevant efforts by individual researchers and existing policies by institutions or national bodies, such as ARIADNE and ARIADNEplus (https://ariadne-infrastructure.eu/, accessed on 10 October 2022) [28], PARTHENOS (https://www.parthenos-project.eu/, accessed on 10 October 2022) [29] and SEADDA (https://www.seadda.eu/, accessed on 10 October 2022) [30]. Recently, the preservation of software code has also received attention, as it is considered to increase methodological transparency, computational reproducibility, and the verifiability of research results [31,32].

These actions and wider research policies in Europe and beyond have been actively supportive of the FAIR data principles (https://www.go-fair.org/fair-principles/, accessed on 10 October 2022) that aim at the findability, accessibility, interoperability, and reuse of digital outcomes from different research projects [33]. These principles act as a guide for data
publishers and curators, helping them to assess whether their data can be compatible with Open Science [34] and foster reusability. Of course, these principles consider many technical aspects, requiring considerable effort in their application and underscoring the need for formal training, see also [35] (pp. 27–31). Furthermore, the results of the on-going ERC-funded project CAPTURE indicated that the lack of interoperability of archaeological data and information may be the outcome of techno-politics. To that end, the *paradata*, namely the creation method of digital research outputs, are receiving further attention [36,37].

Despite the progress that has been made with respect to data and code sustainability, little critical attention has been directed towards their usability within the framework of new data processing opportunities, broadly brought together under the so-called fourth scientific revolution and big data paradigm. Datafication processes evoke optimism in the processing powers of the new tools without much consideration of the actual informational capacity involved in already-collected data or the effort and compromises that need to be made in order to use them under this new data-intensive paradigm [38,39]. In many ways, they remind the positivist declarations about the power of the computer-aided statistical analyses of the 1960s, reflected in the theoretical paradigm of processual archaeology [40].

In this respect, efforts in documenting the history of digital practice in archaeology may be a place to look for obtaining an overview of the current situation. The gradual retirement of the first generation of pioneers and the onset of the subsequent generations of digital practitioners were confronted over time through special events, dedicated conferences and virtual exhibition projects that coincided with the 40-year anniversary of the CAA in Archaeology Conference [41,42] (see also the *Virtual Museum of Archaeological Computing*: https://archaeologicalcomputing.lincei.it/, accessed on 10 October 2022, and the *Personal Histories from the 40 Years of Computer Applications and Quantitative Methods in Archaeology Conference*: https://www.sms.cam.ac.uk/media/1357554, accessed on 10 October 2022). These events tried to document the history of digital practice in archaeology from the points of view of the actual researchers that lived through it. Early digital practitioners led the transition from a largely analogue practice environment and were succeeded by a still-active generation of scholars (digital migrants) that witnessed the transition onto the web, the mobile environment, and the IoT. From the point of view of the current generation of digital natives, a link must be maintained to showcase the elements of practice that halted their development in earlier periods. If we accept that “we are all digital archaeologists” ([43]: 523 emphasis in original), then such a link is possible. However, to complement this endeavour, a similar focus should also be directed to policies regarding the execution of archaeological fieldwork, as well as on educational curricula that will empower the digital archaeologists of the future.

We further need to realize our legacies in terms of our actual practices that include a network of actors. Our digital progress would not have been possible had it not been for several more actors, including progress in hardware and software. In the wake of the passing of Bruno Latour, we cannot forget his legacy in the description of scientific production through his *Actor–Network Theory* (ANT) [44], and we should direct our attention to the benefits and shortcomings of the digital tools we use to make sense of the past, e.g., [45]. These tools, be them equipment or software, are products of dynamic economical and technical environments and prone to change themselves including ripening or end of life, e.g., [46]. The very choice of digital tools shapes our methodologies and knowledge building practices. A facet of this discussion could include the intricate relationship between commercial and open-source practices, which has often been linked either to ethical or politically informed dimensions of research conduct or, at a more practical level, to the economic capacities of entire institutions or individual research projects (for a broad overview of OS, see [47]). Although commercial frameworks continue to expand both their functionality and customer base, and open-source solutions are increasingly supported as part of the Open Science movement, the implied boundary between commercial and open source is also dissolving; Microsoft purchasing GitHub is one of the better-known examples (https://news.microsoft.com/announcement/microsoft-acquires-github/, accessed on...
10 October 2022). We must also keep in mind that a similar discussion with respect to hardware is much more underdeveloped (although see, for example, [48]).

In many ways, it is still good to remember that certain palimpsests of practice exist to this day—not as relics of the past, but as different concurrent processes that may either be the result of or lead to digital inequalities [49]. Inequalities in accessibility, use, and proficiency are present to this day, despite our increasing evocation of the benefits of ubiquitous computing and mobile technologies. Critical and genealogical studies in the digital humanities highlight certain aspects of digital legacies, including post-colonialism, globalism, and white supremacy [50,51]. Within the field of archaeology, there have been people advocating the need to slow down and reflect on the wider ecosystem of digital practice [52], while links have been made to the importance of putting legacy data collections to use in empowering indigenous stakeholders and formalize their involvement in heritage protection and planning [53]. Decolonization of the digital approach [54,55] also brings potential for a better future. An effort towards the dissolution of existing norms and biases is especially relevant in this era as the data-intensive paradigm is making a noticeable comeback; what is embedded in the legacy may be reproduced.

3. The Current Volume

The studies contained in this Special Issue highlight different aspects of the above discussion.

Dawson et al.’s contribution titled “Temporal Frankensteins and Legacy Images” provides a thought-provoking piece that combines a strong epistemological concern with a useful art/archaeology take, using illustrative examples from Avebury. They argue that the proliferation of digital image processing technologies facilitates new potentialities with respect to reusing or repurposing legacy images from separate eras and in different original materialities (analogue/digital). The creative employment of standard image-based modelling procedures may allow stitched or mixed-together outputs, or imageries that co-include multiple temporalities. These, as the article title suggests, create some kinds of messy or Frankenstein representations that essentially suppress proximate viewpoints from different eras into a composite temporal object. The notion of a “timeshed”, extends a “viewshed” in time, indicating the changes from a specific viewpoint within specific temporal brackets or timespans. Rather than thinking of legacy images as partial snapshots of artefact biographies, in the digital environment, they can be augmented or combined into new digital visualizations that may be both revealing and captivating. These observations are especially relevant to the volume of digital images that form part of image-based documentation processes providing the mass legacy data of tomorrow.

Huggett’s review article on “Data Legacies, Epistemic Anxieties, and Digital Imaginaries in Archaeology” provides a theme-fitting contribution that combines questions not only of the very nature and characteristics of digital data, but also of our false or misleading expectations stemming mainly from our growing fascination with the possibilities offered by evolved, smart and progressively automated digital processing systems. The author provides insightful analyses of several themes that are pertinent to current or enduring views of data, such as data openness, data storehouses, data connectivity, data sizes, data neutrality and data remoteness. In all cases, it appears that the concept of data is not only under-theorised, but usually even idealised with respect to its perceived capacities. The discussion moves on to provide thorough and literature-informed considerations charting the various journeys of digital data and their encountered frictions over the course of several stages that include initial capture, processing, usage, and deposition, effectively demonstrating the fact that in many cases the end result is very difficult to re-task and re-use. It seems then that before harvesting the new opportunities provided by new technologies of data processing, we need to reflect on our data making and data (re)using practices.

Stamati et al.’s work on the “Virtual Reconstruction of the Temple on the Acropolis of Kymissala in Rhodes” outlines the entire operational methodology of the 3D virtual reconstruction of a ruined archaeological monument, providing an insightful case for the
usability of legacy data in such modelling exercises. The authors describe the objectives of digital architectural reconstructions, explaining the difficulties stemming from the lack of evidence on which to rely for making modelling decisions of past structures and forms. The study begins by capturing the geometric characteristics of the entire monument and its detached architectural elements in their present condition through photogrammetric surveys. Next comes the study of archaeological documentation and relevant archival sources to detect additional clues or connecting evidence. This is followed by formal stylistic comparisons with other sites that consolidate lost information and provide improbable, plausible and probable reconstruction proposals. Although the authors do not engage with methodologies that attempt to keep track of the reconstruction process, such as the extended matrix [56], they demonstrate how 3D modelling procedures can combine and enhance surviving building parts with legacy information into elements of experimentation within a virtual research environment.

Schmidt, Thiery and Trognitz’s paper on “Practices of Linked Open Data in Archaeology and their Realisation in Wikidata” exhibits the usefulness of LOD approaches in archaeological data connectivity. They provide an introduction of LOD and discuss their fundamentals in terms of structure, creation and publication through concrete examples. Using the open knowledge base of Wikidata, real case studies are described to showcase practical or guided workflows for preparing existing data to be realised as LOD, including existing tools to facilitate the process. Through this multifaceted presentation, the benefits and limitations of LOD are revealed, and although the authors seem, to some extent, to consider all non-LOD as some form of legacy data, when it comes to data linkages, they make a valid argument. Regardless of the possibilities inherent in LOD, we are still quite far from actually obtaining semantically aware machine-based data retrievability, computer-aided processing and inference or large community participation.

Finally, in “Digitising Legacy Field Survey Data: A Methodological Approach Based on Student Internships”, Casarotto discusses legacy data from archaeological surveys, taking into account the rich survey record (from the 1970s onwards) in the Mediterranean region. This older data are especially unique as in some cases they comprise the only available informational resources of landscapes that have since undergone considerable transformations. On the other hand, critical issues in building up legacy data are also noted. For instance, considerable data are still kept in obsolete media or software, while survey results are kept locally, and their access requires in-person visits; in both cases, significant costs are involved in their examination and consolidation. The example presented includes datasets of variable quality from 73 projects, corresponding to a territory exceeding 2 million hectares, that are curated and integrated. The core of Casarotto’s method for consolidating these data in an open and accessible digital format is based on niche sourcing principles. For data collection, students engage in activities ranging from domain knowledge acquisition to annotation of (meta)data, and finally, its deposition to open access repositories. The gain is two-fold as the practicalities of legacy data integration provide the means to question and investigate these datasets. However, this very process is also a reminder of the labour intensiveness of building up legacy data for potential combined dissemination and research employment.

4. Discussion

In the end, this volume has managed to only slightly investigate the broad area that we tried to define under the notion of digital legacies. Despite the opportunity for an open access publication without article processing charges, several planned contributions were unfortunately not realised. The reasons for this are likely related to the recent COVID-19-linked digital transformations that have deprived researchers of both funding and time. Indeed, the current research environment has gone through a major transition where the digital pace has been accelerated, evoking the even faster adoption of digital practices and tools (online meetings, collaborative authoring, virtual labs, cloud computing, social media
presence) as well as the re-appreciation of existing or already-collected archaeological resources [57].

In this volume, legacy data were still the main focus, albeit further explored and linked to new conceptualisation or utilisation methods. We do believe that this topic should be further investigated to inform current practices and provide a critical reflection on what is actually done rather than what is only reportedly done [58,59]. Within the area of the archaeological community, we need to accept that not all are well aware of the potential of digital technologies and even more well-trained in digital techniques, as well as acknowledge the fact that in many cases the resources or means towards such goals are simply absent. It is increasingly important that digital practitioners in archaeology acquire a critical understanding of the outcomes of the digital techniques and means that they are employing. This has to be especially stressed in a period where we see the rise of more automated and semi-automated digital tools (such as ML, DL and AI techniques) that are employed in different aspects of archaeological research. We conclude that perhaps more than ever, the entire community of digital archaeologists requires time and reflection on digital practices, the nature of data, and digital power dynamics that are formed on the basis of pre-existing ‘legacies’.

Funding: This research received no external funding.

Acknowledgments: We would like to acknowledge all authors that provided submissions for this Special Issue. Additionally, the editorial office and the editor-in-chief for providing guidance and help during the article processing stage.

Conflicts of Interest: The authors declare no conflict of interest.

References
5. Taylor, J.; Dell’Unto, N. Skeuomorphism in Digital Archeological Practice: A Barrier to Progress, or a Vital Cog in the Wheels of Change? Open Archaeol. 2021, 7, 482–498. [CrossRef]
12. Loy, M.; Stocker, S.R.; Davis, J.L. From Archive to GIS: Recovering Spatial Information for Tholos IV at the Palace of Nestor from the Notebooks of Lord William Taylour. Internet Archaeol. 2021, 56, 5. [CrossRef]


18. Brancato, R. How to access ancient landscapes? Field survey and legacy data integration for research on Greek and Roman settlement patterns in Eastern Sicily. *Groma* 2019, **4**, 1–32. [CrossRef]


21. Wallace, C.A.B. Retrospective Photogrammetry in Greek Archaeology. SDH 2017, **1**, 607–626. [CrossRef]


32. Strupler, N.; Wilkinson, T.C. Reproducibility in the field: Transparency, version control and collaboration on the project panoramos survey. *Open Archaeol.* 2017, **3**, 279–304. [CrossRef]


37. Huvila, I. Improving the usefulness of research data with better paradata. *Open Inf. Sci.* 2022, **6**, 28–48. [CrossRef]


45. Morgan, C. Analog to Digital: Transitions in Theory and Practice in Archaeological Photography at Çatalhöyük. Internet Archaeol. 2016, 42, 7. [CrossRef]


47. Wilson, A.T.; Edwards, B. (Eds.) Open Source Archaeology: Ethics and Practice; De Gruyte Open Poland: Warsaw, Poland, 2015. [CrossRef]


58. Huggett, J. A Manifesto for an Introspective Digital Archaeology. Open Archaeol. 2015, 1, 86–95. [CrossRef]