

## **SUSTAINABILITY, UNSUSTAINABILITY AND OPPORTUNITY FOR ARCHAEOLOGICAL DATA**

**Theme:** 3. Sustainable archaeology and heritage in an unsustainable world

**Author:** Holly Wright

**Co-Author(s):** Julian Richards, Paola Ronzino, Franco Niccolucci

**Keywords:** data management, data stewardship, data sustainability, infrastructure, archive, grey literature

The rise of digital data in archaeology has created a sustainability crisis requiring urgent action, while also creating opportunities. The majority of archaeological interventions are non-repeatable, and the careful recording carried out to document the resource becomes primary data. As this data is increasingly undertaken using digital methods and tools, archaeological data is often born digital, and with no paper surrogates for the primary record. Archaeological researchers are creative and innovative in their methodologies; adopting, adapting and developing novel techniques and approaches, requiring stewardship of a far greater variety of data formats than other cultural and scientific domains, along with more complex understandings of data re-use. This combination of factors, along with the challenges created by development-led archaeology and a research environment focused on project-by-project funding models, makes moving to a sustainable model even more challenging.

At the same time, work around stewardship and management of archaeological data has contributed to the creation of persistent resources, including an increase in the availability of open access to 'grey literature', and data from diverse providers has been made interoperable and cross-searchable across national boundaries. Continued effort to make archaeological data open, persistent and sustainable are urgently needed, and archaeologists must work together to raise awareness and take action. The ARIANDEplus infrastructure (<https://ariadne-infrastructure.eu/>) and Saving European Archaeology from the Digital Dark Age (COST Action SEADDA, <https://seadda.eu/>), invites papers discussing the sustainability of archaeological data, exemplars of open data and data re-use, and technologies and initiatives that promote interoperability and persistent resources. This session also welcomes papers discussing challenges associated with the sustainability of archaeological data, to promote better understanding of how we may work together as a community to address them.

**Abstract book ISBN:**

978-80-907270-7-6

## **Abstracts for session #350**

## **DIGGING DIGITAL IN EPHESES - CHANCES AND CHALLENGES OF A LONG-TERM PROJECT**

Helmut Schwaiger<sup>1</sup>, Karl Burkhart<sup>1</sup>

<sup>1</sup> *Austrian Archaeological Institute*

Since 125 years archaeological research takes place at the site of ancient Ephesos. In this period a lot of data has been generated and published. As methods, possibilities and technology have been progressing, also the variety of data has significantly increased.

Although usually different excavation projects are conducted within a certain time period a long-term project like Ephesos bears the opportunity to develop strategies of an efficient life cycle of research data. The interdisciplinary approach offers a wide variety of research fields which is both, fruitful and challenging. Different groups of scientists produce data which have some elements in common but on the other hand their nature differ a lot when looking closer. The workflow has shown that a proper data management plan with strict routines are inevitable in order to avoid a 'Babylonian language confusion'. In order to provide data exchange with other research groups outside the Ephesian project a data standardization has to be developed. With such an - sometimes challenging - effort the interoperability and re-use of research data shall be ensured. At the same time the integration of the results of the institute's work in international frameworks has shown that certain challenges have to be solved.

The paper will present the status quo of an ongoing process. Solutions have to be constantly re-evaluated and their implementation in the international research data infrastructure is highly needed. Once lifted treasures should not remain reburied due to a lack of missing data curation.

### **Keywords**

data management, interoperability, interdisciplinarity, Ephesos, big dig, infrastructure

### **Note/comment**

## **SAVING DATA IN THE CZECH REPUBLIC? WHERE WE ARE AND WHAT TO EXPECT**

David Novak<sup>1</sup>, Olga Lečbychová<sup>2</sup>

<sup>1</sup> *Institute of Archaeology of the CAS, Prague*

<sup>2</sup> *Institute of Archaeology of the CAS, Brno*

The archiving of digital data was until recently a largely overlooked topic in the Czech Republic. The establishment of a specialized archaeological digital infrastructure, the Archaeological Information System of the CR (AIS CR; <http://www.aiscr.cz/>), provided a new impetus in 2016. Overall, this has channelled the flow of data in Czech archaeology but has not yet expanded the spectrum of information that is stored in principle. The scope of the infrastructure is limited to some extent by legal standards, as the obligation to permanently retain the results of field investigations is defined in very vague terms. As a result, most primary data remains unknown, usually on completely substandard storage sites and not backed up. While a substantial part of the documentation is included in the centrally registered excavation reports, it is not available in its best-use raw form. A complete paradigm shift is inevitable unless we are to lose decades of primary results from archaeological research. Poster describes the current situation and presents both the steps that are taken and needed in the future to adequately protect primary data. It is clear that general initiatives such as FAIR Data or EOSC will play an important role throughout the process, changing data policy from the top, as well as international disciplinary projects (SEADDA, ARIADNEplus) that help to set good practice within the community. But implementation will require a proactive approach by individual states and institutions.

### **Keywords**

digital archiving, research infrastructures, primary data, best practice, FAIR Data

### **Note/comment**

## **SUSTAINABLE ARCHAEOLOGICAL DATA IN NORWAY**

Espen Uleberg<sup>1</sup>, Mieko Matsumoto<sup>1</sup>, Christian-Emil Ore<sup>2</sup>, Jakob Kile-Vesik<sup>1</sup>

<sup>1</sup> *Museum of Cultural History, University of Oslo*

<sup>2</sup> *University of Oslo*

MUSIT (MuseumIT) is a shared infrastructure for the Norwegian university museums. Artefact metadata are stored in a database used for cataloguing and collection curation. By 2020, close to 1.5 million of these artefact entries are published online at unimus.no. As part of the MUSIT cooperation, the museums decided in 2011 to use a common solution for excavation documentation – the Swedish Intrasis (intrasis.com). The result is more than 1200 such projects. The e-infrastructure project ADED (Archaeological Digital Excavation Documentation) will migrate these to one system allowing simultaneous map-based and text-based queries. The new system will be based on PostgreSQL, and ensure that the data will be more sustainable. As part of this migration process the excavation data will be mapped to CIDOC CRMarcheo which will ease the integration with other linked data portals for archaeological in Scandinavia and on a general international level. The MUSIT cooperation at the University of Oslo will be responsible for curating the data after the project period.

The Museum of Cultural History publish excavation reports and other of its publications at the Research Archive at the University of Oslo (duo.uio.no). Excavation and artefact photos are stored in the MUSIT photo database and are published at unimus.no/foto, most of them with a CC 4.0 BY-SA license. The museum is creating more 3D-documentation of artefacts and in the field. In case of photogrammetry, files and models are stored in the MUSIT database. The files from lightscanning are stored in standard formats and as raw files at the university servers.

The museum is this way controlling the documentation process from data acquisition to long time, sustainable data storage. There are challenges connected to archaeological interventions conducted where the information pipeline from counties to museums are not always well defined, but this will also addressed by the ADED project.

### **Keywords**

infrastructure, Norway, MUSIT, CRMArchaeo

### **Note/comment**

Abstract #: 04

## **DIGITAL ARCHAEOLOGICAL DATA IN THE WILD WEST: THE CHALLENGE OF PRACTICING RESPONSIBLE DIGITAL DATA ARCHIVING AND ACCESS IN NORTH AMERICA**

Rachel Fernandez<sup>1</sup>

<sup>1</sup> *Center for Digital Antiquity*

Archaeology in the United States is managed by a variety of Federal, State, and Tribal agencies, universities, and private sector organizations. The decentralized nature of North American archaeology results in a lack of standard requirements for digital data documentation, accessibility, curation, and preservation. In contrast, European Union initiatives, such as Europeana, SEADDA, and ARIADNEplus provide useful examples of the advantages of implementing coordinated approaches to the responsible stewardship and management of archaeological data. Along these lines, the Center for Digital Antiquity at Arizona State University has developed a robust infrastructure that supports open access, reproducible science, and synthetic research through their domain repository, the Digital Archaeological Record (tDAR). Using case studies of projects that utilize tDAR for their archival needs, I discuss the advantages and pitfalls of archiving digital archaeological data in the context of North American archaeology, while looking towards potential collaboration and learning opportunities from our European colleagues.

### **Keywords**

Digital Curation, Repositories, North America

### **Note/comment**

## **CHALLENGES OF MAKING COLLECTIONS ACCESSIBLE. THE CENIEH COLLECTIONS' SYSTEM**

Joseba Rios-Garaizar<sup>1</sup>, Jesus Rodríguez-Méndez<sup>1</sup>, Cecilia Calvo<sup>1</sup>, Gonzálo Cuesta<sup>2</sup>

<sup>1</sup> *Centro Nacional de Investigación sobre la Evolución Humana (CENIEH), Paseo de la Sierra de Atapuerca 3, Burgos 09002, Spain.*

<sup>2</sup> *Universidad de Burgos*

The Centro Nacional de Investigación sobre la Evolución Humana (CENIEH) has had, since its foundation in 2009, the mission of curating and giving access to archaeological collections, particularly those recovered in the UNESCO site of Atapuerca (Burgos, Spain). In 2019 the CENIEH became officially the depository of Atapuerca collections in collaboration with the Museo de Burgos, Museo de Evolución Humana and Junta de Castilla y León. In this time huge efforts have been made not to improve the deposits and the facilities, but also to establish a framework for the digital archiving of the collections and to think how these collections will be accessible for researchers and public in the future. A dedicated software for managing the different collections has been developed. This includes not only the archeo-paleontological collections, but also reference collections as the COAC (Colección de Anatomía Comparada), Litho (Lithothèque), CET (Colección Experimental de Traceología), and Colección de Referencia de Antropología Dental.

Also, since January 2019, the CENIEH is part of the AriadnePlus Consortium. The role of the CENIEH is to provide access to the huge amount of digital information stored and generated by the CENIEH. Different types of data (digital collections, document repositories, raw data, reports) are susceptible to being stored, managed and shared. This action requires a reflexion about how to store this data and how can we make it persistent, accessible and valuable for different kinds of audiences (researchers, general public) without compromising the data itself, the ongoing research and the integrity of the archeological sites of provenance.

### **Keywords**

Collection, Software, Data, Paleoanthropology

### **Note/comment**

More authors will be added for the presentation

## **BEST PRACTICE IN FIELDWALKING DOCUMENTATION: FORMATION OF AN EAA COMMUNITY**

Martijn van Leusen<sup>1</sup>

<sup>1</sup> *Groningen Institute of Archaeology*

For at least the past 25 years archaeologists working around the Mediterranean have stressed, to little avail, the enormous importance - for heritage management and scientific purposes - of being able to merge the hundreds, if not thousands, of fieldwalking datasets generated since the 1950s (Barker & Mattingly eds., 1999-2000). But attempts to analyze multiple fieldwalking datasets (e.g., Alcock & Cherry eds., 2004; Launaro 2011) have so far failed to go beyond an uninformative 'least common denominator' approach. More-over, the data of most fieldwalking surveys, and the metadata for nearly all of them, remain unpublished or inaccessible. However, in recent years we have seen progress on four fronts: survey archaeologists themselves, in bi-annual meetings, are working towards agreement on good field and documentation practice; the FASTI Online Survey project is now actively promoting the submission of legacy survey datasets for online archiving and publication; a Dutch-UK-Italian research team is test-driving a merged survey database for the area around Rome (see paper by Attema in this session); and proposals have been prepared for a field survey extension to the CIDOC Conceptual Reference Model, an international standard for information exchange (De Haas and Van Leusen, forthcoming). Following a review of the status quo, it will be argued that standardisation of survey documentation practices should now be our first goal, along with a definition and justification of best practice in modern, systematic fieldwalking survey. Having the political and organisational weight of the EAA, as the largest gathering of professional archaeologists in Europe, behind this initiative is crucial; the formation of an EAA community for this purpose will be announced.

### **Keywords**

Fieldwalking survey, Data integration, Best practice guidance

### **Note/comment**

Links to the paper proposed by Attema for this session.



## **AND EVERY FAIR FROM FAIR SOMETIME DECLINES? TRACKING THE RISE OF PARADATA TO INCREASE DATA SUSTAINABILITY**

Olle Sköld<sup>1</sup>, Lisa Borjesson<sup>1</sup>, Isto Huvila<sup>1</sup>

<sup>1</sup> *Uppsala University*

Large-scale initiatives for archaeological data sharing (e.g. ARIADNE, SEADDA) and the many developing national and institutional data repositories and archives provide unprecedented access to archaeological research data. Access however does not equal impact. Multidisciplinary studies of data reuse (Pasquetto, Borgman & Wofford, 2019) show that infrastructures for data sharing become widely useful when deposited and archived research data are coupled with rich 'paradata'—contextual information detailing the processes of data production and repository curation (Faniel, Frank & Yakel, 2019). How to identify, document, and communicate paradata of purposeful quantities and qualities therefore makes up a burgeoning area of research in archaeological information science that requires committed inquiry into the practical work of archaeologists and other researchers heading for the archaeological data.

This presentation reports the initial results from a multi-national interview study of archaeologists with repeated experience of publishing data in state-of-the-art data repositories and archives. The presentation investigates archaeologists' efforts to counter the so called "data creators' advantage" (Pasquetto, Borgman & Wofford, 2019) by coupling their published research data with paradata with the intent to support data reuse. Particular focus is put on what we call the "publishing threshold": the liminal space that archaeologists navigate when deciding which paradata to supplement in order to make published data as reusable as possible. Interview data will be complemented by an extensive range of examples illustrating the frontiers of paradata publishing and principal paradata categories in archaeological research.

The study is part of the CAPturing Paradata for documenTing data creation and Use for the REsearch of the future (CAPTURE) funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme grant agreement No 818210.

### **Keywords**

Archive, Data publishing, Data reuse, Paradata

### **Note/comment**

## **RED PILL OR BLUE PILL? USING THE MATRIX TO INVESTIGATE THE FUTURE SUSTAINABILITY OF ARCHAEOLOGICAL DATA.**

Keith May<sup>1,2</sup>

<sup>1</sup> HE - Historic England

<sup>2</sup> University of South Wales

Over the last 25 years the archaeological and Heritage sectors among others have concentrated on how the digital data created and stored on computers can be preserved to the same degree that physical museum objects can be kept for the benefit of others. This interest in digital preservation has been especially strong in the archaeological world where excavation data sets are being increasingly gathered as "born digital" data, using the latest computer technologies, i.e. data created on, and only preservable on a computer. This focus has helped address the principle concerns about "how do we keep this stuff digitally".

The focus of digital archives and museums is now switching from not just simply providing better access to digital archives. Increasingly questions of sustainability that ask how are users in commercial archaeological units, curatorial organizations and academia, along with other members of the public, going to make best (re)use of this growing body of digital information and data.

Using the particular recording practices and life-cycles of stratigraphic and phasing data from archaeological interventions, the AHRC funded Leadership Fellow project The Matrix (<https://www.researchgate.net/project/The-Matrix-connecting-and-re-using-digital-records-and-archives-of-archaeological-investigations>) is being undertaken to investigate how digital data from archaeological excavations can be made more useful and interesting to a range of users and audiences.

The project has objectives in four main areas of activity

- 1) Digital Standards;
- 2) Heritage Data;
- 3) Stratigraphy Standards;
- 4) Search Tools;

that aim to address two key research questions:

1. How can we encourage the sharing, linking and interoperability of archaeological data and information, particularly information derived from the commercial sector in order to maximize public value and enhance the research potential of archaeological data?
2. How can we ensure the consistent development, application, encouragement and ultimately enforcement of technical information and data standards and their promotion to others?

**Keywords**

Heritage Data, data reuse, digital data, digital archives, Stratigraphy, Matrix

**Note/comment**