

Chapter 16

Preserving Our Digital Heritage: Information Systems for Data Management and Preservation

Julian D Richards, Kieron Niven and Stuart Jeffrey

Abstract It is essential that we develop effective systems for the management and preservation of digital heritage data. This chapter outlines the key issues surrounding access, sharing and curation, and describes current efforts to establish research infrastructures in a number of countries. It aims to provide a detailed overview of the issues involved in the creation, ingest, preservation and dissemination of 3D datasets in particular. The chapter incorporates specific examples from past and present Archaeology Data Service (ADS) projects and highlights the recent work undertaken by the ADS and partners to specify standards and workflows in order to aid the preservation and reuse of 3D datasets.

Keywords 3D laser scanning · Close range photogrammetry · Data management · Digital archiving · Digital repositories · Research infrastructure

16.1 Introduction

There is a pressing need to preserve and integrate existing archaeological research data to enable researchers to use new and powerful technologies. Large numbers of archaeological datasets span different periods, domains and regions; more are continuously created as a result of the increasing use of computer-based recording. They are the accumulated outcome of the research of individuals, teams and

J. DRichards (✉) · K. Niven · S. Jeffrey
Archaeology Data Service, University of York, YO1 7EP York, UK
e-mail: julian.richards@york.ac.uk

K. Niven
e-mail: kieron.niven@york.ac.uk

S. Jeffrey
e-mail: stuart.jeffrey@york.ac.uk

institutions, but they form a vast and fragmented corpus and their potential is constrained by difficulties of access and lack of integration. Furthermore, these data are fragile and they will be lost unless they are actively curated.

In particular, the tremendous growth in the use of 3D data in archaeology over the last 10 years can be seen not only in the increasing availability of services and technologies that allow the collection of such data but also in the way in which such datasets often play a key or uniting role in larger, more diverse projects. The generation of 3D data occurs at numerous different scales, from landscape or seabed survey, through the laser scanning or photogrammetric survey of buildings and monuments, all the way down to the digitisation of small objects.

With such a pervasive and important role, the issue of preserving such data for future reuse and reinterpretation comes to the forefront. This is particularly relevant where the data are expensive to acquire or where they are used to monitor or 'digitally preserve' sites or objects that are either inaccessible or subject to deterioration.

The objective of this chapter is to outline the key issues of data management and preservation, with particular reference to some of the large and 3D datasets developed by the applications described in many of the other chapters in this volume. The chapter begins by discussing the major issues and challenges, including digital preservation, access, synthesis and integration, and the increasing requirements and demand for open data. It describes existing efforts to establish research infrastructures in the heritage sector. Finally, it looks at work by the Archaeology Data Service (ADS) and partners to deal with the particular challenges of large 3D datasets.

16.2 Background: Major Issues and Challenges

The current situation in the heritage sector is characterised by a high degree of fragmentation and difficult access due to the fact that:

- There are different actors involved in data creation and management, including research groups, museums, scientific laboratories, cultural heritage administrations, contract excavators and others.
- Data are created and/or need to be consulted in different stages of the archaeological investigation from excavation or field survey to publication of data analysis and interpretation.
- Data may be embedded in, or attached to, monuments records, documentation of excavations or field surveys, scientific laboratory analyses, museum reference collections and others.
- Data types are varied and comprise, for example, textual descriptions, drawings, photographs, maps at diverse scales, 3D models derived from photogrammetry or laser scanning, grey literature (i.e. unpublished reports of contracted excavation work), as well as traditional academic publications.

- Data are increasingly born digital, and the functionality of a GIS or 3D model is not available in a traditional paper publication format.
- Data are fragile, and without adequate documentation and active curation they will not be available for future generations of scholars.

Furthermore, archaeology is unusual in that the creation of knowledge results from the physical destruction of primary evidence, making access to data all the more critical in order to test, assess, and subsequently reanalyse and reinterpret both data and the hypotheses arising from them.

16.2.1 Digital Preservation

The issues associated with the long-term preservation of digital data—together with the advantages of doing so—are becoming increasingly well-known across a wide range of fields. As a result, in recent years guidance and support have been developed at both national and international levels through a number of organisations and projects such as the Digital Preservation Coalition (DPC) and Digital Curation Centre (DCC) in the UK, the National Digital Information Infrastructure and Preservation Program (NDIIPP) in the US, and Digital Preservation Europe (DPE) and the Open Planets Foundation in Europe.

Within Archaeology, although awareness of the need to actively manage digital data is growing, practical developments towards doing so in a secure and standardised fashion are still some way behind other disciplines. As in the wider digital preservation sphere, the issues that are pertinent to the preservation of archaeological digital data revolve around the definition of standards and best practice, i.e. what should be preserved and how this should best be done. A significant element of digital archiving focuses on the use and suitability of data file formats for the preservation and dissemination of data and involves such considerations as binary versus ASCII data types, proprietary versus open file formats, and the management of data compression. In addition, all data requires some form of documentation in order to be understood, not only in terms of how it came into being, but also what it represents and how it can be used. The specification of documentation and metadata standards, and their applicability to archaeological data, remains a significant digital preservation issue.

One of the most widely acknowledged approaches to the practical matter of preserving digital data for the long term is the Open Archival Information System (OAIS) reference model. OAIS comprises hundreds of pages of guidance and good practice and makes clear the importance of open file formats, data migration, robust and distributed hardware infrastructure and the necessity of discovery, access and delivery systems (CCSDS 2012). It does not, however, define the practical implementations of the recommended processes. Actual digital preservation based on OAIS can be enormously complex. In archaeology, preservation processes may have to deal with hundreds of file types, from hundreds of types of

devices, using a plethora of software packages and the whole broad range of archaeological techniques. In addition, for a digital archive to be considered credible, and thereby attain ‘trusted digital repository’ status, it must be able to demonstrate well-documented preservation policies and processes as well as having a robust long-term sustainability plan. The accreditation of digital repositories is still in its early stages although the Data Seal of Approval (DSA) provides an internationally recognised kitemark for repositories and a new ISO standard was recently published (ISO 2012). In 2010, a number of European organisations signed a Memorandum of Understanding that links these into a wider European framework for certification (TDR 2010).

16.2.2 Access and Value

Despite some notable exceptions in one or two countries, most archaeological data is still not accessible because the traditional approach to research also protects the intellectual property rights of researchers—sometimes beyond any reasonable term, as in the case of excavations unpublished for decades and still ‘under study’ by the archaeologist. It does not favour, or even consider, the publication of primary data. By contrast, access to data and data sharing is generally perceived as important. In a survey undertaken by the ADS in 2007, 70 % of respondents had somehow reused old data and 80 % would allow access to their data; one commented that ‘having such data available will assist any longer-term monitoring projects or even cast new light on a previously recorded subject’ (Austin and Mitcham 2007, p. 36).

Nevertheless, and although initiatives to create public archives of heritage data such as ADS have existed for a long time, heritage data sharing is not yet common practice in Europe. Public data repositories and the related standardisation are also the best solution for long-term preservation. Reinforcement for this practice may come from implementing a recommendation that public funding agencies ‘should incentivize a scientific culture in which sharing of data becomes an accepted norm of professional behaviour’ (Kintigh et al. 2010, p. 4). In other words, researchers who want public money must share their data. It should also be noted that data (and not only scientific reports) are part of the EU Open Access initiative, as clarified by *Sharing Knowledge: Open Access and Preservation in Europe*, the conclusion of a 2010 EU strategic workshop (Swan 2011). Section 2.5.2 of the *Digital Agenda for Europe* states that publicly funded research should be widely disseminated through Open Access publication of scientific data and papers (European Commission 2010). In May 2011, the UK’s Engineering and Physical Sciences Research Council (EPSRC) gave research organisations 12 months in which to develop individual roadmaps to put policies and procedures in place to ensure the preservation and availability of digital research data for at least 10 years. Applicants for UK research council funding are also generally required to submit Data Management Plans as part of their proposals (Higgins 2008; Jones 2011).

Other European countries are also working individually, or in combination, to develop data preservation and access policies. ESFRI, the European Strategy Forum on Research Infrastructures, is a strategic initiative to develop the scientific integration of Europe and to strengthen its international outreach. One of its key goals is to facilitate multilateral initiatives leading to the better use and development of research infrastructures, at EU and international level. The ESFRI programme has provided funding for scientific research infrastructures across a range of disciplines. It provided start-up funding for the preparatory phase of Digital Research Infrastructure for the Arts and Humanities (DARIAH), which is now in the construction phase with support from a number of EU countries.

The primary nature of archaeological data makes it particularly vulnerable to data loss and the importance of heritage to cultural identity across many European nations should make it a key priority for support. But how well placed are European repositories to meet this challenge? In many countries it has been assumed that libraries and archives, the traditional custodians of records, will simply take on this additional role, although few are adequately resourced, or staffed, to deal with the scale and complexity of digital data, particularly with 3D data. Several studies have recognised the value of discipline-based repositories in developing stakeholder communities, avoiding fragmentation and establishing discipline-specific data preservation expertise (e.g. RIN 2011).

16.2.3 Synthesis and Integration

To date, synthetic research has often comprised the summarised results of specific research projects. In the words of Kintigh et al. (2010, p. 2), researchers: ‘desperately need to foster synthetic research that transcends the spatial and temporal scales of individual research projects’. This requires tools and methods to integrate and synthesise data collected by researchers in different investigations (Kintigh 2006; Snow et al. 2006). Researchers have to handle an enormous amount of information, but it is not storage resource or processing performance that are required. The purely technological approach of providing more petabytes or guaranteeing more teraflops is insufficient; the diversity of archaeology requires fundamental research encompassing many disciplines, and developing innovative approaches. Integration also represents a challenge when considering the diversity of contexts, collecting protocols, relevance and goals under which data are collected.

A special role in synthesising information is played by innovative visualisation technologies. 3D digitization applications in archaeology are eased by new low-cost devices, improved accuracy and speed, emergence of new image-based solutions to process raw data, more sophisticated algorithms and the consolidation of open source solutions (e.g. the Italian Research Council’s *MeshLab* platform, totalling several thousands of users worldwide). The recent introduction of HTML 5 and WebGL makes it possible to use 3D models on web pages and to distribute

those representations on the Internet. These technologies are now able to produce excellent digital replicas of heritage assets and can be considered as a mature resource. The availability of sophisticated digital clones may extend archaeologists' ability to use a number of computer-assisted tools in order to compare, measure, comprehend and gain new insights (e.g. Scopigno et al. 2011).

16.2.4 Increasing Demand for Data and Interest in Data Sharing

While the data landscape is fragmented, demand from archaeologists to access existing data for consultation, comparison and reuse in current research is widespread. For example, in the UK, the ADS had over one million downloads of unpublished fieldwork reports in the 12 months from February 2011, with an increase of interest as more data are made available. In the United States a recent survey by Archaeoinformatics.org among members of the Society for American Archaeology (SAA) shows that 94 % of respondents would use electronic data more, if they were accessible. The 2011 RIN/JISC report on *Data centres: their use, value and impact* revealed that 84 % of users believed that the existence of the ADS had made a positive impact on the culture of data sharing, 79 % reported that it had improved the efficiency of their research and 65 % said that it had reduced the cost of data acquisition (RIN 2011).

Data sharing has also gained momentum through the promotion of Linked Open Data. Several countries, including the USA, UK and France are moving towards open governmental data and this will inevitably have implications for data provision by state public heritage bodies. The development of an archaeological semantic web has been a long-held vision (Richards 2006) but, until recently, there were few working examples. A number of the basic building blocks are now in place, including mappings of data schemas to standard high level ontologies such as the CIDOC CRM and the provision of classification systems, thesauri and authority files as Simple Knowledge Organization System (SKOS) web services, through projects such as STELLAR and STAR (Binding et al. 2008; Binding 2010; Tudhope et al. 2011).

16.2.5 Research Infrastructures

In summary, there is therefore, a strong case for the development of research infrastructures, generally at a national level, to provide leadership in information management, data access and preservation, but with collaboration at an international level to facilitate the development of common standards and interoperability. Many of the big archaeological research questions transcend modern political

boundaries and research will be enhanced by integrated user access, whilst responsibility for preservation needs to be distributed (Kenny and Richards 2005).

The UK's Archaeology Data Service is the longest standing digital archive for archaeology, and recently enjoyed its fifteenth birthday. The ADS was established in 1996 as one of the five discipline-based service providers making up the UK Arts and Humanities Data Service (AHDS). It is hosted by the University of York. Funding for the ADS came initially from the UK Arts and Humanities Research Board (now AHRC) together with the Joint Information Systems Committee (JISC) but currently consists of elements of core funding from the AHRC together with the Natural Environment Research Council (NERC) alongside a range of project-based funding from a variety of UK and European project and organisations (Richards 2008).

The ADS is the mandated place of deposit for archaeological research data for a number of research councils and heritage organisations and makes all its holdings freely available for download or online research. At the last count, it provides access to over 17,000 unpublished fieldwork reports (the so-called grey literature) and over 500 data rich digital archives. The ADS was the first archaeological digital archive in Europe, and was only preceded by the now defunct Archaeological Data Archive Project (ADAP), in the United States (Eiteljorg 1994). In recent years, however, there have been related initiatives in several other European countries, although admittedly these are concentrated in Northern Europe and Scandinavia.

In 2007 the ADS was joined by EDNA, the e-depot for Dutch archaeology, which was established as part of DANS (Data Archiving and Networked Services), and funded by KNAW, one of the main Dutch Research Councils. Like the UK Data Archive, DANS originated as a social science data archive but from there it expanded into archiving historical data sets and then, in collaboration with Leiden University, into Archaeology through a 2004–2006 pilot study. As of 2007, agreements to deposit archaeological data at DANS were formalised in the quality standard for Dutch archaeology, making archaeology one of the largest components of the digital resources hosted by DANS. By the end of 2011, the EDNA provided access to over 17,000 reports and excavation archives, although some are only downloadable by registered archaeological users. EDNA employs two archaeological archivists, but also benefits from input from the much larger staff of DANS.

Recently, the Swedish National Data Service (SND), based at the University of Gothenburg, decided to extend its collection policy to focus on Archaeology. It has worked with the Department of Archaeology and History at the Uppsala University to archive a number of archaeological reports. At present, SND is starting the publication of over 200 GIS files with the excavation data from Östergötland. SND is a service organisation for Swedish research within the humanities, social sciences and health sciences. A second Swedish infrastructure initiative focuses upon access to data pertaining to environmental archaeology. The Strategic Environmental Archaeology Database (SEAD) is based at Umea University, in northern Sweden. The SEAD project is funded by the Swedish Research Council and

Council for Research Infrastructures. It aims to facilitate the online storage, extraction, analysis and visualisation of data on past climates, environments and human impact, by providing online tools to aid international researchers in these tasks, and by providing access to data that are currently not accessible online (Buckland et al. 2011).

The most recent initiative to establish a national archaeological digital research infrastructure in Europe has been led by the German Archaeological Institute (DAI), which is part of the DFG, funded via the German Foreign Ministry. In 2012, the DAI established a new project, IANUS, with an initial staff of two, to scope what would be required to set up a digital archive for German archaeology.

In North America, there have been a small number of significant initiatives which seek to provide cross-institutional support for digital archiving. Although seen primarily as a data publication tool, *Open Context*, based at the Alexandria Archive Institute, has developed a relationship with the California Digital Library to provide for long-term citation and preservation, and it is now one of two repositories mandated by the National Science Foundation (Kansa and Whitcher Kansa 2009, 2010). The other is *tDAR*, hosted at Arizona State University, and supported since 2009 by a 4-year start-up grant from the Andrew W Mellon Foundation to the Digital Antiquity consortium (McManamon 2010). In Canada, the Canadian Foundation for Innovation and the Ontario Ministry of Research and Innovation have funded *Sustainable Archaeology*, a 9.8 million Canadian dollar joint initiative between the Western University and McMaster University, with the initial aim to digitally consolidate archaeological collections that are currently scattered across the Province of Ontario, Canada. These initiatives make Northern American archaeology better placed to address growing pressure from governmental and research bodies to make the results of research and the data underpinning scientific research freely and publicly accessible.

In Australia too, there have been numerous attempts to develop a digital research infrastructure for archaeologists. The latest of these is Federated Archaeological Information Management System (FAIMS), a highly ambitious project led by the University of New South Wales, and funded by the Australian Government's NECTAR programme. FAIMS is a 12-month project which aims to 'assemble a comprehensive information system for archaeology. This system will allow data from field and laboratory work to be born digital using mobile devices, processed in local databases, extracted to data warehouses suitable for sophisticated analysis, and exchanged online through cultural heritage registries and data repositories'.

16.3 Dealing with 3D Datasets

3D datasets are now routinely generated by a range of techniques in many archaeological projects, as demonstrated by the case studies in this volume. They can be used to integrate data derived from multiple sources but they can raise new

or unique problems for data management, digital preservation and access. The following examples, derived from the ADS experience, highlight the development of archival processes that address a number of these problems.

16.3.1 The Big Data Project

Many datasets generated by techniques such as marine bathymetry, laser scanning and LiDAR present specific challenges in that the volume of data created is frequently very large and therefore has storage implications (including the cost of buying and maintaining hardware or purchasing separate storage). In addition, beyond the actual storage of data, the physical size of such datasets can often create problems in terms of data access or reuse. The 2006 Big Data Project, funded by English Heritage, looked specifically at the practical issues raised in storing and disseminating large 3D datasets through three case studies: marine survey data from Wessex Archaeology, laser scanning data from Durham University and LiDAR data from English Heritage (Austin and Mitcham 2007). The project started with a data audit and culminated with the deposit of data from each case study with the ADS. In addition, the project also carried out a questionnaire survey and workshop aimed at ‘Big Data’ creators in order to quantify and assess the types of data being created alongside the options available for dissemination and reuse. As a result of these activities, the project produced a final report aimed at raising awareness of the issues associated with creating, storing and accessing ‘Big Data’ as well as providing guidance in terms of both policy and practice.

The project report also provided a key set of recommendations for future research which has subsequently informed the recent Guides to Good Practice project (discussed) with the findings incorporated into the relevant individual guides (Austin et al. 2008).

16.3.2 The VENUS Project

In addition to the issues of storage and dissemination highlighted by the Big Data Project, ADS involvement in the 2006–2009 European VENUS project looked at the preservation of large, complex marine survey datasets, often featuring multiple streams of data combining to form various different data ‘products’ (Alcala et al. 2008). One key aspect of this project was to demonstrate how data selection plays a key role in producing a robust and reusable digital archive. The VENUS project itself aimed to develop scientific methodologies and deliver technological tools for the virtual exploration of deep underwater archaeology sites with the ADS role being focussed on the long-term preservation of the project’s digital outputs. The ADS specifically focussed on the publication of a VENUS Guide to Good Practice (Austin et al. 2009) alongside an exemplar digital archive (Drap 2009).

A significant outcome of the project was the identification of ‘Preservation Intervention Points’ (PIPs) in the data lifecycle of the project. The VENUS underwater missions themselves surveyed shipwrecks at various depths by employing a complex data acquisition process using remotely operated unmanned vehicles with innovative sonar and photogrammetry equipment. Subsequent data processing stages also included the plotting of archaeological artefacts and the production of 3D models. At various stages in the data lifecycle, ADS identified PIPs where data were transformed by processes such as decimation, aggregation, recasting or annotation, in addition to data being migrated from format to format. These stages were then evaluated in terms of whether, for the purposes of preservation, it might be appropriate to intervene and take a preservation copy of the data to be archived. The evaluation process itself was based on a number of broad criteria, seven in total, which allowed each point to be weighed up in terms of categories such as reuse potential, repeatability, value (cost) and available meta-data (in terms of both data reuse and the repeatability of specific processes). Interestingly, the PIP criteria highlights that, although it is generally considered good practice to archive data in as raw a state as possible—because often the subsequent transformations applied can be recreated—this is not—always the case for certain 3D datasets where data are subsequently merged (e.g. meshes) or processed (e.g. cleaned or decimated) to create ‘new’ interim datasets, often via a proprietary or automated processes (Fig. 16.1).

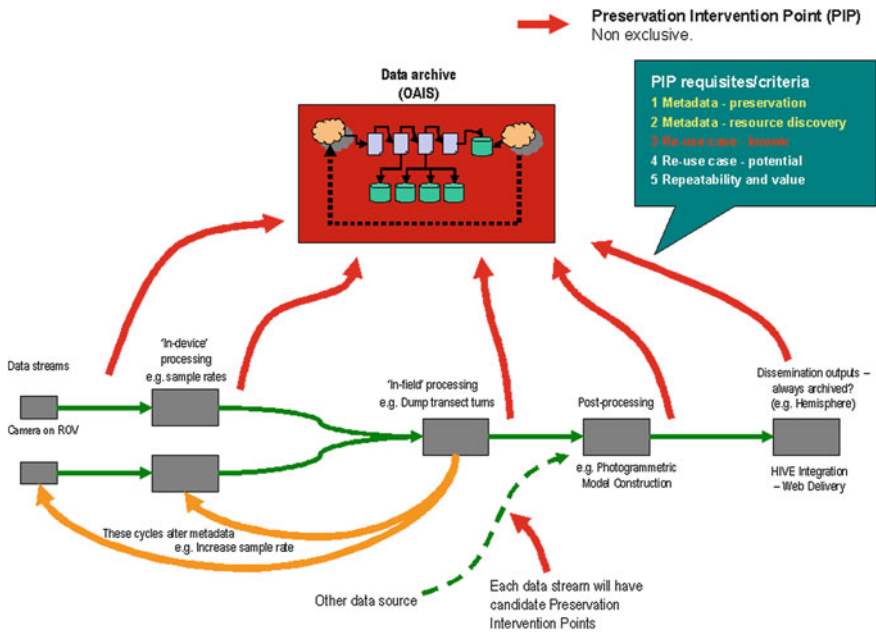


Fig. 16.1 Preservation intervention points in the digital workflow

As with the Big Data Project, the VENUS project has also made a significant contribution to the subsequent revision and expansion of the Guides to Good Practice through the further development of elements of the VENUS guide into a new general guide looking at marine survey data. In addition, the concept of PIPs is equally applicable to other datasets including laser scanning and photogrammetry and this conceptualisation of the process of data selection has been incorporated into these guides.

16.4 Guides to Good Practice

The research that emerged from the Big Data and VENUS projects, as mentioned above, was incorporated into a wider suite of guidance material in the 2009–2010 Guides to Good Practice project (Mitcham et al. 2010). The 2-year collaborative project, funded by the US Mellon Foundation and English Heritage, aimed to revise, update and expand the original ADS series of Guides to Good Practice. The six guides originally authored between 1998 and 2002 (Gillings and Wise 1998; Bewley et al. 1998; Richards and Robinson 2000; Schmidt 2001; Eiteljorg et al. 2002; Fernie and Richards 2002) were integrated and updated within an online wiki and extended to cover a number of additional techniques including close range photogrammetry (CRP), marine survey and laser scanning (Fig. 16.2).

**Archaeology Data Service / Digital Antiquity
Guides to Good Practice**

Log in

- Home
- Full Table of Contents
- Digital Archiving**
 - About these Guidelines
 - How to use these Guides
 - What is Digital Archiving?
 - Archival Strategies
- The Project Lifecycle**
 - Planning for the Creation of Digital Data
 - Project Documentation
 - Project Metadata
 - Data Selection: Preservation Intervention Points
 - The Project Archive: Storage and Dissemination
 - Copyright and Intellectual Property Rights
- Basic Components**
 - Documents and Texts
 - Databases and Spreadsheets

This new and revised series of Guides to Good Practice have been produced as the result of a two-year collaborative project between the UK Archaeology Data Service and Digital Antiquity in the US. The project has encompassed important revisions of the existing six ADS Guides as well as the development of entirely new documents covering areas such as marine survey, laser scanning, close-range photogrammetry, digital audio and digital video. The project has involved previous Guides authors revising existing content alongside new authors, from both Europe and the US, also contributing to the development of the guides into new themes and areas.

The project has been undertaken in collaboration with the Digital Antiquity initiative, a US-based project with the aim of enhancing the preservation of and access to digital records of archaeological investigations. A major aim of the Guides is to provide the basis for archaeological project workflows that will create digital datasets that can be archived and shared effectively by Digital Antiquity's iDAR archive and repository in the US and by the Archaeology Data Service in the UK. The development of the Guides involves close collaboration with teams in the US at both the University of Arkansas and Arizona State University.

Other ADS projects have also fed into the revision and development of the Guides. ADS involvement in the European VENUS project has formed the basis of a guide focussed on marine survey. In addition, the incorporation of findings from the ADS Big Data project, together with the revision of the existing guide on aerial photography and remote sensing data, has seen a significant contribution to the guides from English Heritage funded projects.

Previous versions of the ADS/AHDS Guides to Good Practice have been archived and are still available on the old Guides to Good Practice page. View the full new Guides to Good Practice Table of Contents

ARCHAEOLOGY | DATA SERVICE |
 The Digital Archaeological Record
A SERVICE OF DIGITAL ANTIQUITY

The Andrew W. Mellon Foundation | ENGLISH HERITAGE |

Fig. 16.2 Guides to Good Practice

As with the previous series of Guides, the new updated series aims to specify standards and ‘good practice’ for a variety of techniques used within archaeological projects and covers the data formats that they produce, the suitability of these formats for long-term data preservation and, importantly, the metadata and documentation (at a number of levels) required to archive, understand and reuse these datasets in the long term. Building on the work of the VENUS project, the new Guides also examine 3D techniques and data types such as laser scanning and photogrammetry within the larger project lifecycle of data acquisition, processing, reprocessing and the creation of various types of derived data such as CAD models, still images and video.

16.5 Current Projects

A number of recent and current projects are also contributing to the way in which the ADS approaches the archiving and dissemination of 3D datasets.

The deposit of data from the Virtual Amarna Project provided the first opportunity for the ADS to ingest and disseminate a large collection of 3D PDF files (Kemp 2011). A number of artefacts held in the site museum at Tell el-Amarna in Egypt had been scanned by a team from the University of Arkansas, working with Barry Kemp, the excavation director. As deposit of the archive coincided with the production of the laser scanning Guide to Good Practice, this provided the opportunity to create an excellent exemplar archive and highlight the various incarnations of data created within a laser scanning project together with the metadata required to understand and reuse these data (Limp et al. 2011; Payne 2011) (Fig. 16.3).

The ADS is also currently involved in the European CARARE project as a content provider mapping and adding—amongst other data—3D datasets to the Europeana portal. The use within the project of the 3D PDF format as the primary means of object dissemination has highlighted this format’s potential for easy and flexible dissemination of complex 3D models as well as allowing ADS to work with similar European organisations that are creating or disseminating data in what is a relatively underused format.

In addition to 3D-specific projects, the ADS has also undertaken work to ensure that the data it stores, regardless of type, is both secure and reusable. In 2011, the ADS was awarded the Data Seal of Approval verifying that it meets the standards of a trusted digital repository (Mitcham and Hardman 2011). In addition, in 2012, the ADS was also accredited as an official Data Archive Centre (DAC) for the UK Marine Environmental Data and Information Network (MEDIN). The ADS has also implemented the ISO standard Digital Object Identifier (DOI) system for its collections via DataCite (<http://datacite.org/>). From an ADS perspective, the availability of DOIs for its digital collections not only ensures that data citations remain stable and resolvable but also that they are quantifiable in terms of reuse and impact via metrics from the DataCite DOI resolver.

The screenshot shows the website for the Virtual Amarna Project. At the top is the 'ads' logo and 'ARCHAEOLOGY DATA SERVICE' with a search bar. Below is a navigation menu with links for HOME, ARCHSEARCH, ARCHIVES, LEARNING, ADVICE, OUR RESEARCH, BLOG, ABOUT US, and LOGIN. A secondary menu includes ALL, JOURNALS AND SERIES, GREY LITERATURE, PROJECT ARCHIVES, BIBLIOGRAPHIES, and THESES. The main heading is 'The Virtual Amarna Project' by Barry Kemp, 2011. On the left, there are links for Introduction, Overview, and Downloads, along with copyright information for Egypt Exploration Society and Amarna Trust. A 'Primary contact' section lists Prof Barry Kemp at the McDonald Institute for Archaeological Research. A 'Send e-mail enquiry' button is present. Below that are 'Resource identifiers' including ADS Collection: 1077, doi:10.5284/1011330, and a link to cite using the DOI. The main content area has an 'Introduction' section with text about the Virtual Amarna Museum and the city of Tell el-Amarna. To the right of this text is a photograph of an ancient Egyptian mask. Further text describes the project's methodology using 3D laser scanning and mentions archaeologist Anna Stevens. A final note states that the small selection cannot fully illustrate the theme and provides a link to the project website.

Fig. 16.3 The Virtual Amarna project archive

16.6 Conclusion

Archaeological data requires active management throughout the project lifecycle to ensure that it will be ‘fit for purpose’ for future preservation and access. A number of factors are encouraging researchers to think about providing open access to their data and to plan for its long-term preservation. These include policy recommendations from research councils and governments, as well as an increasing desire from researchers themselves to share data. A number of countries are now developing digital research infrastructures and data archives and there have been attempts to promote more integrated access.

3D data faces the same preservation issues as other archaeological datasets although, in some cases, certain problems are somewhat heightened, including issues of data storage, ensuring adequate metadata, documenting processing techniques and dealing with proprietary software. While, for example, the use, and even preference, for proprietary or compressed 3D data formats within the community is a well-recognised preservation issue, new complex problems in terms of access (due to the sheer size or number of files) is a particular characteristic of 3D data in archaeology.

This increasing volume of data produced by 3D projects has another aspect in that, when the larger workflow is viewed and multiple incarnations of the ‘same’ data are viewed as holding value, storing and providing access to these multiple sets of files becomes problematic. While storage capabilities continue to increase alongside decreasing hardware costs, indicating a possible solution to such data storage issues, it is worth noting that many technologies which capture 3D data continue to generate larger and larger volumes of data countering the perceived savings implied by lower hardware costs. It should also be noted that the most significant cost component of any digital repository is actually the labour required to ingest, migrate and subsequently manage the data. This cost is contingent on many factors, the least of which may be data volume (Richards et al. 2010).

The continuing growth and use of 3D data and the varied technical systems used to generate it will also hopefully continue to see a parallel development of both data format and metadata standards. Ongoing collaborative research at both a national and international level, as demonstrated by the projects discussed in this chapter, will be an important factor in developing systems that ensure that the data produced remain secure and usable for future generations.

References

- Alcala, F., Alcocer, A., Alves, F., Bale, K., Bateman, J., Caiti, A., et al. (2008). VENUS (Virtual Exploration of Underwater Sites) two years of interdisciplinary collaboration. In: *Proceedings of VSMM'08 Conference on Virtual Systems and Multimedia Dedicated to Digital Heritage*, (pp. 250–258) Cyprus, Limassol: VSMM
- Austin, T. & Mitcham, J. (2007). Preservation and management strategies for exceptionally large data formats: ‘Big Data’, Archaeology data service. Retrieved 14 Nov 2012 from http://archaeologydataservice.ac.uk/attach/bigData/bigdata_final_report_1.3.pdf
- Austin, T., Mitcham, J. & Richards, J.D. (2008). From questions to answers; outcomes from the ‘Big Data project’ in Layers of Perception. In: A. Posluschny, K. Lambers, I. Herzog (Eds.), *Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA)* Berlin, Germany (pp. 194–199), April 2–6, 2007, 2008, Kolloquien zur Vor- und Frühgeschichte Band 10, Bonn. Retrieved 14 Nov 2012 from http://archiv.ub.uni-heidelberg.de/propylaeumdok/volltexte/2010/497pdf/06_04_austin_et_al_bigdata.pdf
- Austin, T., Bateman, J., Jeffrey, S., Mitcham, J. & Niven, K. (2009). Marine remote sensing and photogrammetry: A guide to good practice. In K. Niven (Ed.), *Archaeology data service*. Retrieved 14 Nov 2012 from http://guides.archaeologydataservice.ac.uk/g2gp/VENUS_Toc
- Bewley, R., Donoghue, D., Gaffney, V., van Leusen, M., & Wise, A. (1998). *Archiving aerial photography and remote sensing data: A guide to good practice. AHDS guides to good practice*. Oxford: Oxbow Books. Retrieved 14 Nov 2012 from <http://ads.ahds.ac.uk/project/goodguides/apandrs/>
- Binding, C., May, K., & Tudhope, D. (2008). Semantic interoperability in archaeological datasets: Data mapping and extraction via the CIDOC CRM. *Lecture Notes in Computer Science*, 5173, 280–290.
- Binding, C. (2010). Implementing archaeological time periods using CIDOC CRM and SKOS. *Lecture Notes in Computer Science*, 6088, 273–287.

- Buckland, P. I., Eriksson, E., Linderholm, J., Viklund, K., Engelmark, R., Palm, F., et al. (2011). Integrating human dimensions of Arctic palaeoenvironmental science: SEAD – the strategic environmental archaeology database. *Journal of Archaeological Science*, 38(2), 345–351.
- CCSDS (2012). Consultative committee for space data systems (CCSDS) reference model for an open archival information system (OAIS). CCSDS 650.0-M-2, Magenta Book, Issue 2, June 2012. Retrieved 29 Nov 2012 from <http://public.ccsds.org/publications/archive/650x0b1.pdf>
- Drap, P. (2009). VENUS: Virtual exploration of underwater sites: Port-miou C wreck, Marseille. *Archaeology Data Service*. Retrieved 14 Nov 2012, doi:10.5284/1000004
- Eiteljorg, III, H. (1994). The archaeological data archive project, center for the study of architecture. Retrieved 14 Nov 2012 from <http://csanet.org/archive/adap/adaplond.html>
- Eiteljorg, III, H., Fernie, K., Huggett, J. & Robinson, D. (2002). *CAD: A guide to good practice. AHDS guides to good practice*. Oxford: Oxbow books. Retrieved 14 Nov 2012 from <http://ads.ahds.ac.uk/project/goodguides/cad/>
- European Commission (2010). Digital agenda for Europe: Communication from the commission. Retrieved 14 Nov 2012 from [http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52010DC0245R\(01\):EN:NOT](http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52010DC0245R(01):EN:NOT)
- Fernie, K. & Richards, J.D. (2002). *Creating and using virtual reality: A guide for the arts and humanities. AHDS guides to good practice*. Oxford: Oxbow Books. Retrieved 14 Nov 2012 from http://guides.archaeologydataservice.ac.uk/g2gp/Vr_Toc
- Gillings, M. & Wise, A. (Eds.). (1998). *GIS guide to good practice. AHDS guides to good practice*. Oxford: Oxbow Books. Retrieved 14 Nov 2012 from <http://ads.ahds.ac.uk/project/goodguides/gis/>
- Higgins, S. (2008). The DCC curation lifecycle model. *The International Journal of Digital Curation*, 3(1), 134–140.
- ISO (2012). ISO 16363:2012. Space data and information transfer systems-Audit and certification of trustworthy digital repositories. *International Standards Organisation*. Retrieved 30 Nov 2012 from http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=56510
- Jones, S. (2011). *How to develop a data Management and sharing plan. DCC how-to guides*. Edinburgh: Digital Curation Centre. Retrieved 14 Nov 2012 from <http://www.dcc.ac.uk/resources/how-guides>
- Kansa, E., & Whitcher, Kansa, S. (2009). Open context: Developing common solutions for data sharing. *CSA Newsletter*, 21(3). Retrieved 14 Nov 2012 from <http://www.csanet.org/newsletter/winter09/nlw0902.html>
- Kansa, E., & Whitcher, Kansa, S. (2010). Publishing data in open context: Methods and perspectives. *CSA Newsletter* 23(2). Retrieved 14 Nov 2012 from <http://csanet.org/newsletter/fall10/nlf1001.html>
- Kemp, B. (2011). The virtual amarna project. *Archaeology Data Service*. Retrieved 14 Nov 2012. doi:10.5284/1011330
- Kenny, J. & Richards, J.D. (2005). Pathways to a shared European information infrastructure for cultural heritage. *Internet Archaeology* 18, Retrieved 14 Nov 2012. http://intarch.ac.uk/journal/issue18/kenny_index.html
- Kintigh, K. W. (2006). The promise and challenge of archaeological data set integration. *American Antiquity*, 71(3), 567–578.
- Kintigh, K. W., McManamon, F. P. & Spielmann, K. A. (2010). *Synthesis and cyberinfrastructure for SBE research. White paper for SBE 2020: Future research in the social, behavioral and economic sciences*. Phoenix: Arizona State University. Retrieved 14 Nov 2012 from <http://www.tdar.org/wp-uploads/2010/12/201010115-Cyberinfrastructures-for-SBE-Research-final-w-abstract.docx>
- Limp, F., Payne, A., Simon, K., Winters, S. & Cothren, J. (2011). Developing a 3-D digital heritage ecosystem: From object to representation and the role of a virtual museum in the 21st century. *Internet Archaeology* 30. Retrieved 14 Nov 2012 from http://intarch.ac.uk/journal/issue30/limp_index.html
- McManamon, F. P., Kintigh, K. W. & Brin, A. (2010). Digital antiquity and the digital archaeological record (tDAR): Broadening access and ensuring long-term preservation for

- digital archaeological data. *CSA Newsletter* 23(2). Retrieved 14 Nov 2012 from <http://csanet.org/newsletter/fall10/nlf1002.html>
- Mitcham, J. & Hardman, C. (2011). ADS and the data seal of approval—a case study for the DCC. Retrieved 14 Nov 2012 from <http://www.dcc.ac.uk/resources/case-studies/ads-dsa>
- Mitcham, J., Niven, K. & Richards, J. D. (2010). Archiving archaeology: Introducing the guides to good practice. In: A. Rauber, M. Kaiser, R. Guenther, P. Constantopoulos (Eds.), *Proceedings of the 7th International Conference on Preservation of Digital Objects. iPRES 2010*. Vienna, Austria, September 19–24, 2010. Vienna: Austrian Computer Society
- Payne, A. (2011). Laser scanning for archaeology: A guide to good practice. In: K. Niven (Ed.), *Archaeology data service/digital antiquity guides to good practice*. Retrieved 14 Nov 2012 from http://guides.archaeologydataservice.ac.uk/g2gp/LaserScan_Toc
- Richards, J. D. (2006). Archaeology, e-publication and the semantic web. *Antiquity*, 80, 970–979.
- Richards, J. D. (2008). Managing digital preservation and access: The archaeology data service, in managing archaeological resources: Global context, national programs, local actions. In: F. P. McManamon, A. Stout, J. A. Barnes (Eds.), *One World Archaeology* 58, (pp. 173–194). Walnut Creek: Left Coast Press
- Richards, J.D. & Robinson, D. (Eds.). (2000). *Digital archives from excavation and fieldwork: Guide to good practice* (2nd ed.). *AHDS Guides to Good Practice*. Oxford: Oxbow Books. Retrieved 14 Nov 2012 from <http://ads.ahds.ac.uk/project/goodguides/excavation>
- Richards, J. D., Austin, A. F., & Hardman, C. (2010). Covering the costs of digital curation. *Heritage Management*, 3(2), 255–263.
- RIN (2011). Data centres: Their use, value and impact. Research Information Network & Joint Information Systems Committee. Retrieved 14 Nov 2012 from <http://www.jisc.ac.uk/media/documents/publications/general/2011/datacentres.pdf>
- Schmidt, A. (2001). *Geophysical data in archaeology: A guide to good practice. AHDS guides to good practice*. Oxford: Oxbow Books. Retrieved 14 Nov 2012 from <http://ads.ahds.ac.uk/project/goodguides/geophys/>
- Scopigno, R., Callieri, M., Cignoni, P., Corsini, M., Dellepiane, M., Ponchio, F., et al. (2011). 3D models for cultural heritage: Beyond plain visualization. *IEEE Computer Graphics and Applications*, 44(7), 48–55.
- Snow, D. R., Gahegan, M., Lee Giles, C., Hirth, K. G., Milner, G. R., Mitra, P., et al. (2006). Cybertools and archaeology. *Science*, 311, 958–959.
- Swan, A. (2011). *Sharing knowledge: Open access and preservation in Europe*, (pp. 25–26). Conclusions of a strategic workshop, Brussels, November 2010. Luxembourg: European Commission. Retrieved 14 Nov 2012. doi:10.2777/63410. http://ec.europa.eu/research/science-society/document_library/pdf_06/oa-preservation-2011_en.pdf
- TDR (2010). Memorandum of understanding to create a European framework for audit and certification of digital repositories. Trusted Digital Repository.eu. <http://trusteddigitalrepository.eu/Site/Trusted%20Digital%20Repository.html>. Accessed 30 Nov 2012.
- Tudhope, D., May, K., Binding, C. & Vlachidis, A. (2011). Connecting archaeological data and grey literature via semantic cross search. *Internet Archaeology* 30. Retrieved 14 Nov 2012 from http://intarch.ac.uk/journal/issue30/tudhope_index.html/