



Digital Preservation in Archaeology

Alicia Wise and Julian Richards

Abstract

Archaeology is in a special position in that much of the creation of its data results from destruction of primary evidence, making access to data all the more critical in order to test, assess, and subsequently reanalyze and reinterpret both the data and the hypotheses arising from them. Over the years, archaeologists have amassed a vast collection of fieldwork and post-excavation data archives, a significant proportion of which remain unpublished. Access to data, even those which are published, is often difficult or inconvenient at best. The Archaeology Data Service (ADS), established October 1st 1996, is working towards provision of digital data over the Internet through an integrated on-line catalogue to its collections and the collections of partner organisations. The work of the ADS thus proceeds along two tracks: digital preservation and facilitating access to digital information.

1 Archaeological data

Initiatives around the globe are exploring access to and preservation of digital information. Archaeologists in many countries are participating in this research largely because computer technology is becoming more and more common for the recording of and the creation of primary data in our discipline, and in related disciplines. Significant international attention is focused on the use of relational and object-oriented databases to manage and retrieve large amounts of archaeological information and the World Wide Web as a vehicle for the wide dissemination of information about archaeology. Though in recent years CD-based multimedia applications have been gaining popularity among archaeologists, the Internet seems poised to subsume most of these projects.

While it is not possible to describe in detail the myriad number of digital resources which exist, not even the most extensive Web gateways can do that, a sample of international projects is offered below.

Information technology is developing rapidly down under. The Australian National University sponsors a WWW gateway called Archaeology World (Australian National University 1997) and sites accessible through this gateway offer on-line access to databases, images, and texts. As digital preservation is a pressing topic at places such as Australia's National Preservation Office (National Preservation Office 1997), the infrastructure to accommodate widespread archiving in archaeology is developing rapidly.

Canadian archaeologists are well served by CHIN (Canadian Heritage Information Network 1997). Information on this extensive website include an inventory of archaeological artifacts and other objects in Canadian museums, and extensive information about a variety of resources of interest to archaeologists and museums professionals.

Europe is producing a number of promising information technology projects, largely through funding programmes of the European Commission and the European Union. One

such project is Aquarelle, a partnership of cultural heritage organisations and software developers in England, France, Greece, and Italy (Aquarelle 1997). The goal of this project is to develop software to allow the international transfer of multimedia folders between museums and cultural resource managers (Dawson, this volume). Another project has resulted in Guder Grave, an impressively designed website containing information about all aspects of the European Bronze Age (Jarl Hansen, this volume and National Museum of Denmark 1997). Currently this resource is available only in Danish, but translations of the web material should make this valuable resource more accessible in countries where Dutch, English, and French are spoken. Europe is also well-served by an extensive web-based gateway, the Archaeological Resource Guide to Europe (Champion and van Leusen 1997).

The Israel Antiquities Authority has developed a centralized database for recording information about all archaeological excavations in the country. Upkeep of this database has been facilitated by the development of a bilingual thesaurus to control terminology entered by excavators, and by a database template distributed to excavators for guiding the construction of recording systems for their projects (Cohen 1996).

In the United Kingdom a variety of resources exist from the relational databases underpinning the National Monuments Records in England, Northern Ireland, Scotland, and Wales to the diverse county-level Sites and Monuments Records. The Council for British Archaeology (CBA) is involved with projects designed to integrate these varied digital resources and others including the British and Irish Archaeological Bibliography (British and Irish Archaeological Bibliography 1997) and an extensive gateway to British Archaeology (Heyworth 1997). CBA projects also include an electronic journal (Internet Archaeology 1997) and a digital archive (Archaeology Data Service 1997).

In the United States there are a wide-variety of good digital resources available over the World Wide Web. One that is frequently used is ArchNet, a virtual library for archaeology maintained by staff at the University of Connecticut's library

(Plunkett and Lizee 1997). ARGOS is a collection of peer-reviewed websites covering classical archaeology which are accessible through a searchable front-end (Beavers and Sonpal 1997). Other projects in the United States include the National Archaeological Database (NADB) which was created by the National Park Service to manage its enormous collection of archaeological survey and excavation data. Recently the Centre for Advanced Spatial Technologies at the University of Arkansas has taken on the challenge of integrating the information in the NADB with a GIS interface (National Park Service and CAST 1997). Additionally, a variety of State Historic Preservation Offices are investing in the creation of large relational databases to maintain their large records of archaeological interventions. Some of these databases have been created with GIS front ends, others are not quite so technically involved, but all are extremely useful resources. Last, but certainly not least, is the Archaeological Data Archive Project at Bryn Mawr University (see Eiteljorg this volume and Eiteljorg 1997).

Worrying is the infrequency with which the long term preservation of these, and other, valuable digital resources is discussed. Two organisations provide notable exceptions. One is the Archaeological Data Archive Project just mentioned. The other is the Archaeology Data Service in the United Kingdom. These two projects share a common interest in *digital archiving and facilitating access to digital information*.

2 The Archaeology Data Service

The idea for a digital archiving service for archaeology in Britain was stimulated in early 1996 by a call for proposals from the Arts and Humanities Data Service (AHDS). The AHDS was then a newly established service of the body responsible for funding all information technology programs in Britain's universities, the Joint Information Systems Committee, and consisted of an Executive at King's College London and one service provider. This first service provider, the History Data Service, is a digital archive for historians residing at the University of Essex Data Archive.

A bid to provide a similar digital archiving service for archaeology was discussed in an informal lunchtime meeting at the Computer Applications in Archaeology UK meeting in York in February 1996 and subsequently developed into a consortium proposal by the Universities of Birmingham, Bradford, Glasgow, Kent at Canterbury, Leicester, Newcastle, Oxford, and York, with the Council for British Archaeology (Richards 1996). An announcement that this bid was successful was made later that spring, the first member of staff was appointed in July, and the Archaeology Data Service began life on October 1st, 1996.

The success of this bid focused on the consortium management approach. Each contributor to the original proposal nominated a representative or representatives to sit on the ADS management committee, in order to guide the service through its development phases. The membership of this management committee has subsequently been broadened to reflect a greater number of perspectives in archaeology. For this reason, a committee member represents the Royal Commissions for Ancient and

Historical Monuments in England, Scotland, and Wales. Another member represents the Association of Local Government Archaeological Officers, the body which advocates needs of archaeologists in England and Wales involved in the curation and protection of archaeology during and after the development control process.

The consortium organisation of the ADS offers other strengths as well. Each member of the management committee brings a different range of archaeological and technological skills a different range of hardware and software. This diversity of resources, uniquely possible in a service primarily based in the university sector, means the ADS is in a position to offer a broader range of services than its 2 full-time and 2 part-time funded posts would otherwise allow.

The ADS is hosted by the Archaeology Department at the University of York, but the consortium structure focuses direct attention on the fact that archaeological data archives are dispersed. From the beginning, it has been assumed that the holdings of the ADS would be dispersed among consortium members and would be linked to digital archives held by other organisations. The Internet, existing initiatives to standardise the contents and descriptions of archaeological data, and new standards for metadata (Miller 1996, Wise and Miller 1997) and international communications standard Z39.50 mean that these data do not need to be brought together in one single physical location in order to make them available to the wider community.

There is also a fractal nature to the ADS. Managed by a consortium of experts, the service is also advised by a diverse 40-member advisory committee and is part of the inter-disciplinary and distributed Arts and Humanities Data Service.

The advisory committee consists of British archaeologists from all sub-fields in the discipline: artifact specialists, computer archaeologists, contracting unit staff, cultural resource managers, environmental archaeologists, government-mandated regulatory bodies, independent or avocational archaeologists, museum curators and archivists, scientific archaeologists, underwater archaeologists, university-based researchers and teachers, and others.

The Arts and Humanities Data Service has developed into a cohesive, integrated, yet distributed whole (Burnard and Short 1994, Greenstein 1996, Greenstein and Trant 1996, Ross 1995). Service providers have been appointed to address the digital archiving needs of a variety of disciplines in addition to archaeology and history. These include ancient and modern texts and linguistic corpora at the Oxford Text Archive; broadcasting, music, and theatre at the University of Glasgow's Performing Arts Data Service; and architecture, art history, museum studies and all other aspects of imaging in the arts and humanities at the Surrey Institute of Art and Design's Visual Arts Data Service.

The aims of the ADS are:

1. to collect, describe, catalogue, preserve, and provide user support for digital resources that are created as a product of archaeological research.

2. to take an AHDS-wide responsibility for promoting standards and guidelines for best practice in the creation, description, preservation and use of spatially referenced information.
3. to take, in collaboration with existing agencies, a major role within the UK archaeological community in promoting standards and guidelines for best practice in the creation, description, preservation, and use of digital archaeological research data.

The brief of the ADS is to provide for data generated in the course of archaeological research by British archaeologists, wherever they are working. In theory, therefore, there are no geographical limits on collections but in practice the focus is expected to be on the archaeology of the British Isles. As digital archives for archaeology develop, like the Archaeological Data Archive Project in the United States, it is intended to develop reciprocal archiving policies.

For those classes of archaeological data where there are existing archival bodies the role of the ADS is to collaborate with the appropriate national and local agencies to promote greater use of existing services. In cases where no existing agency has an obvious responsibility, the ADS will act to accession data directly.

The collections policy of the ADS is therefore developing on two fronts to:

1. archiving and preservation of digital data
2. facilitating access to distributed digital information

3 Digital preservation

Archaeology is in a special position in that much of the creation of its data results from the destruction of primary evidence. Increasingly, the digital record may be the only record of precious research materials. With the ever-increasing pace of change in computer hardware and software, in a few years' time that data may be lost forever.

The ADS is trying to ensure that orphaned data sets are not lost and is developing agreements with those organisations that fund archaeological research to ensure that archaeological data sets are offered for deposit and preservation. A range of funding bodies already recognize the importance of digital preservation and either recommend or require that digital data sets be offered to the ADS or other appropriate service providers within the AHDS:

1. British Academy
2. Carnegie Trust
3. Council for British Archaeology
4. Economic and Social Research Council
5. Leverhulme Trust
6. Natural Environment Research Council
7. Wellcome Trust Institute for the History of Medicine

Data sets offered to the ADS or other service providers in the AHDS will be subject to an evaluation procedure before acceptance. This is because timely deposit of digital

information is not enough to ensure that it can be preserved and re-used in the future. Thorough documentation of the dataset, the purpose for which it was collected, the ways in which it was recorded and transformed, and the ways in which it is structured are essential for digital archiving. This is because data must be migrated to new software, operating systems, and hardware platforms as technology changes. Without thorough documentation the information necessary for accurate data migration is missing.

There is another reason that data sets offered to the ADS will be evaluated before acceptance. Computers have been used to create and record archaeological data for more than three decades. There is therefore an enormous backlog of digital data which could be preserved. More and more data are appearing, however, that can only be archived in digital formats. The ADS must prioritize its collections procedures. It will not be possible to hold all computerised archaeological data: instead those data which have secondary use value and are documented thoroughly to facilitate reuse must come first.

At present there is very little information about either the cost of digital archiving, or the level of demand for re-use from an accessible archaeological archive. To inform its collections policy, the ADS is collecting information about these issues by implementing a number of archiving pilot studies.

3.1 Pilot studies in digital preservation

In order to develop appropriate procedures for handling digital archaeological data, a series of pilot projects are being organised. These pilots project explore two main aspects of digital archiving:

1. Archiving specific data types (e.g. texts, CAD, GIS, databases, geophysics, text)
2. Archiving data derived from a variety of sources (e.g. excavation, buildings recording, museum archives)

Pilot projects in development or currently underway by the ADS include:

CAD Archiving , The RCHME has proposed developing a CAD pilot project with the ADS. We will be looking to the established expertise of the Archaeological Data Archive Project at Bryn Mawr University for collaboration on this project.

Electronic Publishing , Internet Archaeology, a fully electronic journal, and the ADS are currently identifying a project or projects which are ready for both publication and archiving. The aim of this pilot would be to explore the issues related to a synthetic article published electronically and based on a fully functional database archived with the ADS.

Fieldwork Archiving , The ADS is working with English Heritage, a government CRM body, and contracting units, such as the Oxford Archaeological Unit, to identify and develop one or more fieldwork archiving pilots. The objectives of these pilot projects will be to determine the costs of digital archiving; the types of archaeological data which should be digitally archived; the types of digital data

which have commercial value, the types of documentation necessary to inter-connect artifact, paper, and digital archives; and the information/training necessary to ensure archaeologists can reuse the digital archive.

Text Archiving , With the backing of the ADS, the Council for British Archaeology has applied successfully to the Fast Track Digitisation programme for funding to digitise its occasional and research report series. This will involve over 105 volumes of text and images saved in PDF format, and the aim of this pilot is to examine strategies for the delivery of electronic texts.

3.2 Guides to good practice in digital archiving

In its role of promulgating standards the ADS is also developing a series of Guides to Good Practice. These are primarily aimed at the creators and depositors of digital data and will provide advice on appropriate archive file formats as well as covering the generation of metadata. There are four explicitly archaeological guides covering:

1. Excavation and fieldwork archives
2. Geographical Information Systems
3. Geophysics
4. Satellite imagery and aerial photographs

Guides produced by other service providers will cover related areas, such as databases, sound and video images (Wissenburg 1997).

4 Facilitating access to digital information

Widespread access to digital information serves two important functions: it improves communication about archaeology and it aids the preservation of digital data. 'Improving communication' happens in a variety of ways: researchers gain access to primary data too expensive to print conventionally or that were never printed at all, anyone with Internet access gains access to archaeological information even if they have no access to a specialist library, spatially distributed archaeological resources can be discovered without costly trips to the numerous archives, museums, SMRs, and NMRs which have developed, and archaeological synthesis across regional and national boundaries generally becomes easier. 'Preservation of digital data' is aided when digital datasets are downloaded and used in different computing environments thus increasing the number of formats and operating systems datasets are migrated into.

It is necessary to develop national and international strategies for digital archiving in archaeology. The ADS is working closely with national and local agencies in the UK to develop a coherent policy for the preservation of digital fieldwork archives and other datasets. Any digital archiving strategy must recognize that not all digital information can be preserved, and that clear definitions of those categories of digital data which have a secondary usage value, and thus need to be preserved, are available. It is also necessary to distinguish between data which is to be made available for open public access and that which is to be stored and maintained but not yet released. Wherever possible it is

desirable to negotiate research access to such data, whilst recognising that there may be cases (for example, copyright restrictions pending full publication of a project, or developer funding) where the data depositor still has a financial interest in the data set and will wish to maintain a degree of control over it. Finally, developing methods to inter-connect the growing numbers of digital datasets held in different archaeological organisations, and making these as easily accessible as possible, is an important goal for digital archiving strategies.

4.1 Case studies

Two examples clearly illustrate the importance of facilitating access to archaeological information: the first is the experience of an undergraduate archaeologist in Britain and the other is the experience of a professional archaeological consultant.

Gigi Signorelli, a recent graduate of the University of York Archaeology Department, did an undergraduate dissertation entitled "Agricultural Strategies in North East Scotland: Barley Production from the Prehistoric to the Medieval Period". Part of this research required him to assemble a thorough record of all archaeological sites in which evidence for any kind of cereal crop was collected in this region and for this time period. A thorough literature search in the library revealed the few sources as evidence for cereal production were often buried within texts instead of being included as main features of the title or subject keywords (a common frustration in archaeology). His next line of attack was to contact the National Monuments Record of Scotland in Edinburgh. He paid a visit to Edinburgh on his limited student budget, and arrived to discover that the kind of information he needed was not instantly available from the NMRS and that it would be better for him to contact or visit the various SMRs in his region of interest. Out of money, and quickly running out of time, he considered changing his research topic. In consulting with staff in his department, he was told about a computerised database of archaeological plant remains created by Philippa Tomlinson and Allan Hall (Tomlinson and Hall 1996) which he could access over the World Wide Web. In Gigi's own words "Searching through the ABCD gave me a very good starting point for my research and it actually provided me with the main material I needed for the entire project." (Signorelli 1997).

A second case study also highlights the importance of making information in existing archaeological archives more accessible to potential-- important for both the staff and the users of those archives. A professional archaeological consultant, Dr. Catherine Mortimer, is completing a study of cruciform brooches from Anglian contexts in England, and is faced with the challenge of making sure she has identified all known examples of this artifact class. No newcomer to her research topic, Dr. Mortimer worked with several hundred British examples of cruciform brooches for her PhD research and only needs to fill in any gaps in her database. To accomplish this straight-forward task, however, has been anything but easy. She has been forced to write to the National Monuments Record in England and 43 county Sites and Monuments Records. Luckily, most of her queries have met with polite and informed responses from helpful SMR

staff members. Think of the time, though, that has been required to process the same query 44 different times and the effort that will be involved in making sense of responses formatted in 44 different ways.

4.2 Pilot studies

Just as the ADS is developing pilot studies to examine issues of digital preservation, so to is it collaborating with outside organisations to examine issues of connecting distributed digital resources. Again, there are two major areas of research:

1. Building digital connections between existing archives
2. Re-use of digital datasets

Accessing Scotland's Past , Foremost among these pilots is Accessing Scotland's Past (ASP) which aims to build digital connections between existing archives, in this case the National Monuments Record (NMR) of Scotland and the Sites and Monuments Records (SMRs) of the Shetland Isles and the West of Scotland.

Accessing Scotland's Past is a pilot project between the ADS, Royal Commission on the Ancient and Historical Monuments of Scotland (RCAHMS), the Shetland Amenity Trust, and the West of Scotland Archaeology Service. Funding for the pilot comes from the Scottish Cultural Resources Access Network (SCRAN) and the Joint Information Systems Committee (JISC).

The goal is to interconnect data held in a NMR, two SMRs, and the ADS. A key research aim of this pilot is to prototype the ADS metadata cataloguing system. International communication standards such as ISO 23950 (better known as Z39.50) and cataloguing standards such as Dublin Core metadata are the linchpins for this pilot.

Museums Pilot , A possible pilot study demonstrating the integration of archives distributed across museums and the ADS is under development. Partners for this project are currently being sought. Museums in Britain are the official repositories for all artifact and paper-based archives resulting from archaeological fieldwork. Many museums are facing two crises: one a lack of space as the quantity of

recovered artifacts grows and the second a lack of resources for digital archiving as archaeologists begin to deposit computer diskettes.

University Teaching and Learning , A pilot is under development between the ADS and the Association for Environmental Archaeology. The idea is to develop a suite of web-based teaching materials about environmental archaeology which will be used to facilitate access to and reuse of digital datasets mounted at the ADS.

4.3 Workshops

Given that the point of preservation is to enable re-use, the ADS is also developing a range of end user services. The ADS takes responsibility, in collaboration with the AHDS, for the education of researchers and accordingly organizes appropriate training courses for both data providers (e.g. methods of documentation), and data users (e.g. network access tools). The ADS publicises its activities, not only through the central work of the AHDS Executive, but also through its own Web pages, conference presentations, and through a bi-annual hardcopy newsletter circulated widely throughout the archaeological community.

5 Conclusion

In summary, a variety of initiatives are underway at the Archaeology Data Service in collaboration with UK and international partners. The aim of these initiatives is to enable the collection, description, cataloguing preservation and re-use of all forms of digital data in archaeology. It is recognized that without widespread input from all stakeholders within and outside of the archaeological community, it will be impossible to develop workable solutions that address all of our very real needs.

In 1997 one millennium is drawing to a close and the digital age has already begun in archaeology. Now is the time to tackle the difficult issues surrounding preservation of the digital data we are actively creating before such a back-log develops that irreplaceable archaeological information is lost forever.

Bibliography

- Aquarelle, 1997 Aquarelle, <http://aqua.inria.fr>
- Archaeology Data Service, 1997 Archaeology Data Service, <http://ads.ahds.ac.uk/>
- Australian National University, 1997 Archaeology World, <http://artalpha.anu.edu.au/web/arc/arcworld.htm>
- Beavers, A F, and Sonpal, H, 1997. ARGOS, <http://argos.evansville.edu/>
- British and Irish Archaeological Bibliography, 1997 British and Irish Archaeological Bibliography, <http://britac3.britac.ac.uk/cba/bab.html>
- Burnard, L, and Short, H, 1994 An Arts and Humanities Data Service, report of a Feasibility Study commissioned by the Information Systems Sub-Committee of the Joint Information Systems Committee of the Higher Education Funding Councils.
- Canadian Heritage Information Network, 1997 Canadian Heritage Information Network, <http://www.chin.gc.ca/>
- Champion, S, and van Leusen, M, 1997 Archaeological Resource Guide to Europe, <http://www.bham.ac.uk/ARGE/>
- Cohen, R, 1996 Integration of Distributed Databases, *Archeologia E Calcolatori*, 7, 831-835.
- Eiteljorg, H, 1997 Archaeological Data Archive Project, <http://csaws.brynmawr.edu:443/web1/adaparchive.html>
- Greenstein, D, 1996 Humanities Computing in the Digital Age: Present Perspectives and Future Trends, *Times Higher Education Supplement*, April.

- Greenstein, D, and Trant, J, 1996 The Arts and Humanities Data Service: Collecting digital research data; building a framework for digital resource preservation and interchange, *Ariadne*, 4, <http://ukoln.bath.ac.uk/ariadne/issue4/>
- Heyworth, M, 1997 Council for British Archaeology, <http://britac3.britac.ac.uk/cba/>
- Internet Archaeology 1997, <http://intarch.ac.uk/>
- Miller, A P, 1996 Metadata for the Masses, *Ariadne*, 5, <http://www.ukoln.ac.uk/ariadne/issue5/metadata-masses/>
- National Museum of Denmark, 1997 Guder & Grave, <http://www.natmus.min.dk/kulturnet/>
- National Park Service and CAST, 1997 National Archaeology Database, <http://www.cast.uark.edu/other/nps/nadb/>
- National Preservation Office, 1997 Preservation and Digital Information, <http://www.nla.gov.au/dnc/tf2001/padi/padi.html>
- Plunkett, T, and Lizee, J, 1997 ArchNet, <http://spirit.lib.uconn.edu/ArchNet/ArchNet.html>
- Richards, J D, 1996 The Archaeology Data Service, *Archaeological Computing Newsletter*, 46, 19-23.
- Ross, S, 1995 Preserving and maintaining electronic resources in the visual arts for the next century?, *Information Services and Use*, 15,373-384.
- Signorelli, G, 1997 My first approach to and use of the Archaeobotanical Computer Database (ABCD), *Arts and Humanities Data Service Newsletter*, 2, http://www.kcl.ac.uk/projects/ahds/public/newslet/news11_2.htm
- Tomlinson, P, and Hall, A, 1996 A review of the archaeological evidence for food plants from the British Isles: an example of the use of the Archaeobotanical Computer Database (ABCD), *Internet Archaeology*, 1, http://intarch.ac.uk/journal/issue1/tomlinson_index.html
- Wise, A L, and Miller, P, 1997 Why metadata matters in archaeology, *Internet Archaeology*, 2, http://intarch.ac.uk/issue2/wise_index.html
- Wissenburg, A, 1997 Guides to Good Practice of the Arts and Humanities Data Service, <http://www/kcl.ac.uk/projects/ahds/public/guides.htm#create>

Contact details

Alicia Wise and Julian Richards
Archaeology Data Service
Department of Archaeology
University of York
King's Manor
York YO1 2EP
UK
email: info@ads.ahds.ac.uk
<http://ads.ahds.ac.uk/>