The Archaeology Data Service: Lifecycle management; day-to-day in a digital archive

Digital Curation MA
Aberystwyth University
17 February 2015

Catherine Hardman
Deputy Director
The ADS: Who we are and what we do

• Founded 1996
• Department of Archaeology
• University of York
• Collections
  o 1,100,000 metadata records
  o 23,000+ unpublished fieldwork reports
  o 700+ rich archives
• Guides to Good Practice
• DPC Decennial Award 2012
• Data Seal of Approval 2010 and 2013
Today’s key points

• Why archive?
• Who’s the audience?
• What do they need?
• How do we provide it? 1. policies
• How do we provide it? 2. technology
• How do we provide it? 3. staffing
• How do we pay for it?
Why is it so important to preserve archaeological data? (preservation by record)

- Archaeology is destructive!
- Archaeological resources are unique
- The documentation becomes the main resource for future interpretation
- Digital data is subject to rapid change to both software and hardware formats
Preservation problems

- IBM 3480
- 3.5" Floppy
- xD Picture Card
- QIC DC600
- DC4_120
- Rectangular Hole Punch Card
- 8mmD-eight
- Cassette tape
- Memory Stick
- 5.25" Optical Disk
- DG90M Tape
- 12" Optical Disk
- MultiMedia Card
- Jaz Disk
- SD Memory Card
- Zip Disk
- 9-Track Reel
- DVD-ROM
- G2000 Tape
- 8" Floppy
- CD-ROM
- 5.25" Floppy
- 5.25" Optical Disk
- 4mm Tape
- Sparq Disk Cartridge
- 8mmD-eight
- Floptical Disk
- Smart Media
- 12" Optical Disk
- Ditto Max
- CompactFlash
- Travan
- CD-ROM
- SD Memory Card
- DLT Tape
Preservation problems: file formats

Images (8)
CAD Vector Graphics (4)
Database (4)
GIS (8)
Movies (3)
Spreadsheets (4)
Statistics (8)
Text (7)
Virtual Reality (4)
Geophysics (3)
Audio (4)

... always under review and growing...
So that was the problem in archaeology

• The policy framework:
  ○ The commercialisation of archaeology (PPG16)
  ○ AHRC and the retention of research data
  ○ Skills and infrastructure shortage in museums
Who are the stakeholders?

- Community archaeologists
- General public
- National agencies
- Commercial archaeologists
- Academics and researchers
- Local Authorities
- Digital curation specialists
Primary reuse of Data

Results from a recent survey (2011) show that 44% of reuse of our data sets is for academic research or teaching and learning.
The ADS supports research, learning and teaching with freely available, high quality and dependable digital resources. It does this by preserving digital data in the long term, and by promoting and disseminating a broad range of data in archaeology. The ADS promotes good practice in the use of digital data in archaeology, it provides technical advice to the research community, and supports the deployment of digital technologies.
How the infrastructure of the ADS looks
How do you know you can trust the way we preserve our data?

- We hold a Data Seal of Approval
- We follow the Open Archival Information System (OAIS) reference model
The sixteen guidelines provide a really good idea of the policies you need!

Towards sustainable and trusted data repositories

There are 16 guidelines that together determine whether your data repository qualifies for the Data Seal of Approval.
Our archive policies

Common to traditional archives:

- A collections policy
- Selection and retention, review and disposal
- Sensitive data DPA/FoI requests
- Human Tissue Act 2004
- Licensing and copyright
- Terms and conditions of use
- Charging policy
Things you never thought you’d have to think about

The Spitalfields Archive

- Ethical issues
- Legal framework
- Personal data
- ‘Open’ data
- Dealing with abuse
This document systematizes an overview of archival practice developed by the ADS since its inception. It does not exist in isolation but as part of a suite of documents guiding good governance and practice by the ADS. Policy and strategy documents include:

- ADS Five Year Plan: April 2008 - March 2013 (strategy document)
- ADS Risk Register
- ADS Preservation Strategy
- ADS Disaster Recovery Plan
- ADS Access Policy

The ADS is further governed by the policy and strategy of its host institution; the University of York. Documents include:

- University of York Records Management Policy 2004
- University of York Information Access and Security Policy
- University of York Legal Statements and linked policy and strategy documents therein

All over seen by a Management Committee
As noted in the Collections Policy the ADS has agreements with a number of funding agencies that support archaeological research, to encourage funding recipients to offer their datasets for deposit

- Arts and Humanities Research Council (AHRC)
- Natural Environment Research Council (NERC), for science-based archaeology

The ADS has Service Level Agreements (SLA) with a number of organisations including

- The UK Data Archive (UKDA) for provision of a remote deep storage facility
- To host and provide an image preservation service to the Parks and Gardens Data Service (PGDS)
- To host and provide a preservation service to the online journal Internet Archaeology

Further, the ADS has Memoranda of Understanding (MoU) with a number of external organisations concerned with preservation and reuse of data

- The Association of Local Government Archaeological Officers (ALGAO)
- The Council for British Archaeology (CBA)
- The Royal Commission on the Ancient and Historical Monuments of Scotland (RCAHMS)
- The Royal Commission on the Ancient and Historical Monuments of Wales (RCAHMW)
- The Royal Commission on the Historical Monuments of England (RCHME now part of English Heritage)
- The mda (formerly the Museums Documentation Association)
- The National Trust
Open Archival Information System (OAIS)
Where are the pinch points?

Depositors!

- End of project
- Wrong format
- No metadata
- No money
- No clue!
To help:

- Guides to Good Practice
- Depositors Guidelines
- ADS easy
- Staff!
The Archaeology Data Service is currently undertaking a collaborative two-year project to revise and develop our series of Guides to Good Practice. The project will encompass important revisions of the existing six ADS Guides as well as the development of entirely new documents covering areas such as marine scanning, laser scanning, GPS, digital audio and digital video. Previous authors have been asked to revise existing content and new authors, from both Europe and the US, will contribute to the development of the guides into new themes and areas.

The project to develop the Guides is predominantly being carried out in support of the Digital Antiquity initiative, a Mellon Foundation funded US-based project with the aim of creating "a collaborative organization devoted to enhancing preservation and access to digital records of archaeological investigations". A major aim of the Guides is to develop the basis for workflows for the creation of digital datasets that can be effectively archived both by Digital Antiquity's DAR repository in the US and by the Archaeology Data Service in the UK. As part of this initiative the development of the Guides will involve close collaboration with teams in the US at both the University of Arkansas and Arizona State University.

Other ADS projects are also planned to feed into the revision and development of the Guides. ADS involvement in the European VENUS project will result in one of the first published guides focussing on marine scanning and photogrammetry. 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, will see a significant contribution to the guides from English Heritage funded projects.

At present the new Guides are currently under development. A marine scanning guide, produced for the VENUS project, has already been produced and is available via the VENUS project page on the ADS website. This guide will form the basis for an expanded marine guide to be released alongside our other new Guides to Good Practice in January 2011.
Workflows and examples
Guidelines for Depositors

Version 1.4, May 2012

Depositing with the ADS
- 1.1. Why Deposit?
- 1.2. How to Deposit

Creating and Documenting your Data
- 2.1. Part 1: Starting the Project
  - 2.1.1. Digital Archive Strategy
  - 2.1.2. The need for Metadata / Documentation
  - 2.1.3. File Naming Strategy
- 2.2. Part 2: Creating and Documenting Your Files
  - 2.2.1. Overview of Preferred Data Formats
  - 2.2.2. Databases and Spreadsheets
  - 2.2.3. Geographical Information Systems
  - 2.2.4. Geophysics and Remote Sensing
  - 2.2.5. CAD and Vector Images
  - 2.2.6. Raster Images
- 2.3. Part 3: Documenting the Project
  - 2.3.1. Creating Metadata Records for Datasets
ADS-easy is a system into which you can upload digital files and associated metadata from archaeological fieldwork and research; on submission these files will be delivered to the ADS for inclusion in our archive.

When should you use ADS-easy?
At least for the time being, ADS-easy is best used for small to medium sized archives, by which we mean archives of around 300 files of a common type. The system works best with projects that contain straightforward file types such as text, images, spreadsheets and CAD files. ADS-easy complements the use of the OASIS system so is especially useful for depositing the digital outputs of small fieldwork projects where you are happy with a simple archive interface i.e. you don’t want an online GIS interface or something similar.

ADS-easy can also accept audio, databases and geophysics files but we currently exclude the upload of 3D laser scanning and larger files due to the limitations of using the web to upload large files.

If you are uncertain about whether you should use ADS-easy please contact us and we can talk to you about the options.

Register / Login to ADS-easy
Click here to register for ADS-easy.
If you have already registered, click on the login link on the menu bar above.
If you have forgotten your login password, please click here.

Latest news about ADS-easy
August 2013: Bosworth Battlefield Project Archive Launched to Coincide with Battle of Bosworth Anniversary Weekend
August 2013: ArchaeologyBritain App Out Now in the App Store
August 2013: New Research Fellowship - Centre for Digital Heritage

Browsers
* Please note that this site works best with the Firefox, Chrome and Safari browsers
Staff and staffing

• Will depend on the style of archive you have
• The personal touch
Staffing levels

- Director
- Deputy Director (Collections)
- Access and Communication Manager
- European Projects Manager
- Lead Applications Developer (systems) + 2
- Digital Archivists (5)
- Administrator
- University of York host
How do we disseminate our data?

All of our holdings are freely available through a web interface.
Open formats; open data

Anglo-Saxon Graves and Grave Goods of the 8th and 7th Centuries AD: A Chronological Framework
John Hines, 2013

Introduction
This archive makes material collected and evaluated under the English Heritage/Lancashire Archaeology Unit project 'Anglo-Saxon England: 500-1050' chronologically accessible to a range of digital files, both to substantiate the printed report of the project 'Anglo-Saxon burial and inter-disciplinary issues of the 6th and 7th Centuries AD' (Lancashire Archaeological Society, 2013) and to enable future researchers to undertake further research using the material.

The archive comprises downloads and an online version of the project database.

Primary contact
Professor John Hines
Department of Archaeology
Cardiff University
John Parkes Building
Colun Drive
Cardiff
CF10 3EU
Wales
Tel: 01222 874929
Send e-mail enquiry

Resource identifiers
ADS Collection 1421
Collection doi: 10.5284/10.18290
How to cite using this DOI

Download
Database files
There is a project database which can be downloaded in Microsoft Access format as compressed files or used online.

<table>
<thead>
<tr>
<th>Database description documentation</th>
<th>PDF</th>
<th>129 Kb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entity relationship diagram</td>
<td>PNG</td>
<td>82 Kb</td>
</tr>
<tr>
<td>References from the Elemental Analyses table</td>
<td>PDF</td>
<td>129 Kb</td>
</tr>
</tbody>
</table>

Database Access features
The Access database has a form called 'Main' which opens automatically when the database is launched. The Main form provides two buttons via which to access the data. Build and Site. Each of these tables is represented by its own tab in the Main window. There is also a Reports button, that provides access to a series of pre-defined reports containing overviews of parts of the data, and a Queries button through which a cross-link of the inventories in the Site can be generated.

The Site button is based upon a gazetteer of later Anglo-Saxon furnished burial sites. It provides an overview of all of the burial sites reviewed from which the sample of burials analysed in detail under this project were drawn, and a subform for the burials that have been recorded from each site in the database. Every site has its own unique Site_ID.

The Build button provides access to all of the burials recorded by the project. Each burial is also designated by a unique ID which is composed of Site_ID (see above) and the grave number. Where necessary, one or more names are inserted before the grave number to allow the burial to be sorted in the correct alphabetical and numerical order.

Downloaded files
These compressed files can be imported into most database packages using the documentation above to link the different tables together.

<table>
<thead>
<tr>
<th>Class</th>
<th>CSV</th>
<th>Kb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>CSV</td>
<td>1 Kb</td>
</tr>
<tr>
<td>Material</td>
<td>CSV</td>
<td>1 Kb</td>
</tr>
<tr>
<td>Otton Age</td>
<td>CSV</td>
<td>1 Kb</td>
</tr>
<tr>
<td>Otton sex</td>
<td>CSV</td>
<td>1 Kb</td>
</tr>
<tr>
<td>Typology 1 Class</td>
<td>CSV</td>
<td>1 Kb</td>
</tr>
<tr>
<td>Typology 2 Type</td>
<td>CSV</td>
<td>1 Kb</td>
</tr>
<tr>
<td>Typology 3 Subtype</td>
<td>CSV</td>
<td>3 Kb</td>
</tr>
<tr>
<td>Typology 4 Variant</td>
<td>CSV</td>
<td>2 Kb</td>
</tr>
<tr>
<td>Typology 5 Subvariant</td>
<td>CSV</td>
<td>5 Kb</td>
</tr>
<tr>
<td>Artifacts</td>
<td>CSV</td>
<td>115 Kb</td>
</tr>
<tr>
<td>Burial</td>
<td>CSV</td>
<td>28 Kb</td>
</tr>
<tr>
<td>CH4</td>
<td>CSV</td>
<td>4 Kb</td>
</tr>
<tr>
<td>County</td>
<td>CSV</td>
<td>1 Kb</td>
</tr>
<tr>
<td>Elemental Analyses</td>
<td>CSV</td>
<td>28 Kb</td>
</tr>
</tbody>
</table>

The University of York

http://archaeologydataservice.ac.uk
Searchable interfaces
Digital Object Identifiers

Types of persistent identifier e.g. Handles, Archival Resource Keys (AR KS) and Persistent URLs (PURLs), all can be resolved to an Internet location. The scheme that is gaining most traction is the Digital Object Identifier (DOI).

http://dx.doi.org/ 10.5284 / 1000389
resolver service prefix suffix
(assigning body) (resource)
Collection level

A Corpus of Early Anglo-Saxon Buckets
Jean Mary Cook, 2003

Introduction
When Jean Cook died in July 2001 the community of Anglo-Saxonists lost one of its eminent members. Her research on grave goods of the pagan period, specifically 'buckets', was a project she began in 1983 and which she diligently pursued through a busy professional life in Museums and University administration. In retirement her interests in teaching adult students and in participating in Oxfordshire historical and archaeological projects, still allowed her to become the first woman Secretary of the Society of Antiquaries of London. When her term of office came to an end she turned almost her whole attention to her research on buckets, intending to bring the work to a conclusion with a database and a monograph. Her copious notes and drawings over years, together with a database planned and operational, were in the files in her study when she died. Four of her friends and colleagues met to discuss how the work, so close to completion, could be brought to publication; they were Dr Diete Brugmann, German archeologist and friend; Helena Hamerow, archeologist and at that time Director of the University of Oxford Institute of Archaeology; Tom Hassall, archaeologist; friend and colleague over many years; and Mary Hodges friend and colleague in many joint projects. Permission was obtained from the family to place the whole archive at the Institute of Archaeology and copyright to any material published was also vested in the Institute. Dr Brugmann prepared a detailed plan showing how the archive could be completed and published, and this plan was submitted to the Society of Antiquaries of London and the Marc Fitch Fund asking for financial support. The support was generously forthcoming and the work began in May 2002. The Institute of Archaeology provided the venue for the archive and for the work itself, including an office and the necessary computer links.

Funders
But how do creators get the maximum benefit?

Internet Archaeology: Data Papers

Sometimes the level of recognition gained from depositing your archaeological research data with an accredited repository is limited, so in an attempt to redress this, Internet Archaeology is working with its sister service ADS (Archaeology Data Service) to establish a series of published, peer-reviewed 'data papers'. You have put a lot of effort into creating your data and an IA data paper allows you to get credit for it, to publicise and share it with the community.

What is a Data Paper?

A data paper is a short, peer-reviewed publication that is designed to raise awareness of your dataset and its re-use potential. A data paper describes the contents of your dataset (already or soon to be deposited with an accredited repository*), the methods used to create that dataset and, most importantly, what further avenues of research are possible.

A data paper is an extension of the 'integrated publication' model we have been developing with ADS since our very earliest issues, integrating data within the article narrative. It differs in that the article is much shorter (and so able to be published more quickly), but it also explicitly credits the referee and makes their comments available to all. A data paper may be used to complement and enhance a related IA publication or help add context to a standalone archive.

The main points that define a data paper are:

- It describes the content and the re-use potential of a particular dataset hosted in a repository (usually by the same authors)
- The data paper is accompanied by an unanonimised referee (peer) statement on the data’s re-use potential and significance
- It is an edited, Open Access publication, disseminated under a CC-BY licence
- It is assigned a CrossRef DOI
- It links back to the original deposited dataset

* We are working in the first instance with the ADS, but other trusted repositories will be added in due course.
Sustainable business models and the ADS Charging Policy

\[ C = A + I + D + R \]

\( C \) (Cost of preservation) =
\( A \) (Management and Administration) +
\( I \) (Ingest costs) +
\( D \) (Dissemination costs) +
\( R \) (Refreshment costs)
Making depositing easier...and cheaper

ADS-easy is a system into which you can upload digital files and associated metadata from archaeological fieldwork and research; on submission these files will be delivered to the ADS for inclusion in our archive.

When should you use ADS-easy?
At least for the time being, ADS-easy is best used for small to medium sized archives, by which we mean archives of around 300 files of a common type. The system works best with projects that contain straightforward file types such as text, images, spreadsheets and CAD files. ADS-easy complements use of the OASIS system so is especially useful for depositing the digital outputs of small fieldwork projects where you are happy with a simple archive interface i.e. you don't want an online GIS interface or something similar.

ADS-easy can also accept audio, databases and geophysics files but we currently exclude the upload of 3D laser scanning and larger files due to the limitations of using the web to upload large files.

If you are uncertain about whether you should use ADS-easy please contact us and we can talk to you about the options.
Costing Calculator

This cost estimation tool allows users to calculate how much deposition through ADS-easy will cost. This will provide potential depositors with guidance on the costs involved in the deposition of their digital data with the ADS and will allow future depositors to more effectively plan for the archiving of a project.

All costs include an initial project start-up fee (currently £150), with further charges on a per file basis.

Find out more about charging and the costing calculator.

Disclaimer: The calculator provides a guide for the potential cost of deposition either via ADS-easy or via a physical media, and should not be treated as a statement of actual costs. This estimate will be valid for the duration of the Financial Year in which it is calculated i.e. an estimate made in January will only be valid until the end of March. If your project requires a definitive costing (especially for AHRC and NERC funding applications) you are encouraged to contact us directly.

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Choose Data Type</th>
<th>File Extension</th>
<th>Quantity (File)</th>
<th>Unit Price</th>
<th>Unit Price (Total)</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database</td>
<td>mdb</td>
<td></td>
<td>4</td>
<td>£3.00</td>
<td>£12.00</td>
<td>Remove</td>
</tr>
<tr>
<td>Text</td>
<td>pdfa</td>
<td></td>
<td>10</td>
<td>£1.00</td>
<td>£10.00</td>
<td>Remove</td>
</tr>
<tr>
<td>GIS</td>
<td>gmi</td>
<td></td>
<td>3</td>
<td>£1.00</td>
<td>£3.00</td>
<td>Remove</td>
</tr>
<tr>
<td>Video</td>
<td>mpeg</td>
<td></td>
<td>5</td>
<td>£3.00</td>
<td>£15.00</td>
<td>Remove</td>
</tr>
</tbody>
</table>

Subtotal: £40.00

TOTAL (incl. £150.00 project start-up fee): £190.00

If you choose to deposit by traditional methods, the estimated price is considerably higher at: £280.00
The Value and Impact of the ADS

Neil Beagrie,
Charles Beagrie Ltd.

John Houghton,
Centre for Strategic Economic Studies, Victoria University
Economic Impact of the ADS

£8.30 return for every £1 spent

Return over 30 years?
Increase in returns on investment in data and related infrastructure arising from additional use facilitated by ADS
Keep in touch

Follow us on Twitter: @ADS_Update

Friend us on Facebook: http://www.facebook.com/archaeology.data.service

catherine.hardman@york.ac.uk