Digital Archiving and Commercial Archaeology

Louisa Matthews
Collections Development Manager
Introduction

- The ADS: Trusted Digital Repository
- Why Archive?
- How digital archiving can help solve some practical problems
- Tools and Implementation
- Epilogue - HERALD
The ADS: Who we are and what we do

- Founded 1996
- Department of Archaeology, University of York
- Collections
  - 1,100,000 metadata records
  - 34,535+ unpublished fieldwork reports
  - 700+ rich archives
- Guides to Good Practice
- DPC Decennial Award 2012
Our original philosophy

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.
Why were we worried about digital archives?

“The rate of change in computing technologies is such that information can be rendered inaccessible within a decade...”

- Joint Information Systems Committee: Why Digital Preservation?
Problems

Hardware failure / obsolescence

“...erm, does anyone have a floppy disk drive...?”

Format failure / obsolescence
Problems

Lack of Documentation

“Have we got the box with the picture on?”

Access Failure

“Dear former employer / supervisor,
I need my project from a few years ago as my laptop met with an accident, can I have a copy please...?”
How we enact the vision

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
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

PLUS:
- ADS Preservation Policy
- ADS Repository Operations
- ADS Ingest Manual
- Copyright Infringement Policy
- Reference Model for an Open Archival Information System
If it’s good enough for NASA, it’s good enough for us!
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 IDAR 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.
Dissemination

How do we disseminate our data?

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

The Prehistoric Stones of Greece: A Resource from Field Survey

Introduction
Overview
Site Search
Survey Search
Usage Statistics

Primary contact
Dr Gilbert Marshall
Department of Geography
Tel: 01794 443209

Send e-mail enquiry

Resource identifiers
ADS Collection: 1620

The University Of

http://archaeologydataservice.ac.uk

15/05/2019

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

Types of persistent identifier e.g Handles, Archival Resource Keys (ARKs) 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)
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 archaeologist and friend; Helena Hamerow, archaeologist 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
Can digital archiving help solve some practical problems?

As well as being good practice?

- Big datasets (not Big Data)
- Museum Archives
- Analogue obsolescence
- Data Discovery, Accessibility and Reuse

Currently developing relationships with:
- Southampton City
- Derbyshire CC
- Devon + Plymouth
- Worcestershire
- Northamptonshire
- Northumberland (images)
- Leicestershire CC HE Team and Museum Service
Laser Scanning and other 3D

<table>
<thead>
<tr>
<th>Dataset Description</th>
<th>Preservation Format</th>
<th>Metadata</th>
<th>Required?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original raw scan files</td>
<td>ASCII TXT and Native</td>
<td>Project Metadata, Scan Metadata</td>
<td>Yes</td>
</tr>
<tr>
<td>Registered point cloud Transformation Matrices for each scan Control Point File if georeferenced</td>
<td>ASCII TXT</td>
<td>Registration Metadata</td>
<td>Yes</td>
</tr>
<tr>
<td>Pre-mesh point cloud</td>
<td>ASCII TXT</td>
<td>Pre-Meshing Metadata</td>
<td>Strongly Advised if archiving a derived Polygonal mesh</td>
</tr>
<tr>
<td>Polygonal mesh</td>
<td>OBJ + MTL (ASCII)</td>
<td>Polygonal Mesh Metadata</td>
<td>Required if final dataset and strongly advised if used to create a derived 2D/3D CAD Dataset</td>
</tr>
<tr>
<td>Decimated Mesh</td>
<td>OBJ + MTL (ASCII)</td>
<td>Decimated Mesh Metadata</td>
<td>Required if final dataset</td>
</tr>
<tr>
<td>2D CAD Dataset</td>
<td>DXF/DWG</td>
<td>2D CAD Model Metadata</td>
<td>Required if final dataset</td>
</tr>
<tr>
<td>3D CAD Dataset</td>
<td>DXF/DWG</td>
<td>3D CAD Model Metadata</td>
<td>Required if final dataset</td>
</tr>
<tr>
<td>DEM</td>
<td>SDTS</td>
<td>DEM Creation Metadata</td>
<td>Required if final dataset</td>
</tr>
<tr>
<td>Video</td>
<td>Suitable MPG format</td>
<td>Video Creation Metadata</td>
<td>No</td>
</tr>
</tbody>
</table>

‘CD inna box’: Southampton Museums

Southampton’s Designated Archaeology Collections

Southampton Arts and Heritage is the final recipient of all archives generated by archaeological fieldwork within the Southampton city boundary. Its role is to store and care for the objects and the paper, photographic and digital records which provide the permanent record of any excavation, watching brief or building recording.

Making these objects and records available is also an essential part of the service, and archives can be viewed by appointment. A database of itemised finds, individual objects made of metal, worked bone, worked stone, wood and other materials, is available on the Southampton City Council website.

The increasing amount of digital data in archives from archaeological fieldwork presents particular problems for long term archiving. The need to refresh and migrate data presents particular problems with increasing pressure on resources. Working with ADS, who can provide long term archiving, ensures that digital data from fieldwork can be made available to the widest possible audience.

Southampton’s archaeology collections are designated as of national importance. The strength of the collections lies in the programme of systematic excavations undertaken since the 1950s. These have produced important evidence of everyday life in the medieval town and its Saxon predecessor, Hamwic. There is also increasing evidence of prehistoric activity in the area, and of life in the Roman town of Clausentum, the first urban settlement within the modern city.
‘Negative’ paperwork

- Use of digital archiving for negative investigations – i.e. where there is no physical archive, is a digital paper substitute sufficient?
Analogue Obsolescence

• Chemical Photography is becoming extremely precarious
• Traditional archive repositories are under pressure, in terms of space and resources
• Solutions are
  1. Print everything out and trust that people have used the right ink, paper, etc. and you can store it in the right conditions.
  2. Use a sustainable digital archiving solution
Northumberland Images

![Commemorative stone set in the eastern elevation of the building](image1.jpg)

**Northumberland Images**

Commemorative stone set in the eastern elevation of the building (OASIS ID: vindomor1-197083)

**Downloads**

<table>
<thead>
<tr>
<th>Reports</th>
<th>Images</th>
</tr>
</thead>
</table>

**Images**

| Image Metadata | CSV | 7 Kb |

**Records 1 - 40 of 40**

| **General view of the school** (BOS14001.jpg) | JPG | 7.22 Mb |
| **Commemorative stone set in the eastern elevation of the building** (BOS14002.jpg) | JPG | 7.58 Mb |
| **The outbuilding** (BOS14003.jpg) | JPG | 4.93 Mb |
| **The southern elevation of the school, showing the original phase on the right** (BOS14004.jpg) | JPG | 4.64 Mb |

---

Genesis Way, Consett
County Durham
DH8 5XP

*Send e-mail enquiry*

**Resource identifiers**

- **ADS Collection**: 1897

---

http://archaeologydataservice.ac.uk

15/05/2019
Implementation

• Brief
• WSI
• Museum Deposition
• Regional statement of good practice

What constitutes a ‘good’ digital archive?
<table>
<thead>
<tr>
<th>What is the agreed archive for the project?</th>
<th>What mechanism should be used to archive it?</th>
<th>How do I estimate the cost?</th>
<th>Notes and requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>A grey literature report of an historic environment event</td>
<td>OASIS</td>
<td>Free</td>
<td>The receiving HER must be content that the ADS will archive the report only as a pdf and will NOT archive the images separately</td>
</tr>
<tr>
<td>A grey literature report of an historic environment event for which up to 3 images are included within the report file</td>
<td>OASIS</td>
<td>Free</td>
<td>A zipped package of images may be uploaded to OASIS. This MUST include the ADS images metadata spreadsheet.</td>
</tr>
<tr>
<td>Report and up to 50 images of the site/building recording</td>
<td>OASIS</td>
<td>£40 + VAT</td>
<td>A zipped package of images may be uploaded to OASIS. This MUST include the ADS images metadata spreadsheet.</td>
</tr>
<tr>
<td>Raw + processed geophysics data and associated files for a single project up to a combined area of 10ha.</td>
<td>ads easy</td>
<td>Use the <a href="#">ADS-easy costing calculator</a></td>
<td>Up to 10ha of geophysics data can be submitted using ADS-easy. The semi-automated ingest process means that the cost is lower than that of the standard ADS archiving procedure.</td>
</tr>
<tr>
<td>Raw + processed geophysics data and associated files for a single project over a combined area of 10ha.</td>
<td>ads</td>
<td>Contact ADS <a href="#">Collections Development Manager</a> for a quote</td>
<td>Costs are calculated per 10ha unit (e.g. 8.5ha would be one unit, whereas a 12.7ha will be 2 units).</td>
</tr>
<tr>
<td>Archives of up to 300 files including both images and other data types</td>
<td>ads easy</td>
<td>Use the <a href="#">ADS-easy costing calculator</a></td>
<td>The semi-automated ingest process means that the cost is lower than that of the standard ADS archiving procedure.</td>
</tr>
<tr>
<td>Archives of up to 300 files including images and other data types not covered by ADS-easy (e.g. Laser scanning and Photogrammetry)</td>
<td>ads</td>
<td>Contact ADS <a href="#">Collections Development Manager</a> for a quote</td>
<td>Please contact ADS <a href="#">Collections Development Manager</a> for a quote</td>
</tr>
<tr>
<td>Larger archives or archives requiring specially designed interfaces</td>
<td>ads</td>
<td>Contact ADS <a href="#">Collections Development Manager</a> for a quote</td>
<td>Please contact ADS <a href="#">Collections Development Manager</a> for a quote</td>
</tr>
</tbody>
</table>
Would you like to know what the new OASIS system will do?

October 23, 2015   OASIS   Jo Gilham

We’re in the process of writing the project design for Stage 2 of the HERALD project – the redevelopment of the form. The actual redevelopment will start next year (if funded by Historic England and Historic Environment Scotland). The draft project design was shown to the OASIS Management Board yesterday (we meet twice a year) and it was agreed to circulate it more widely to see if there was comment on how we have tried to represent the opinions and information gathered during the user needs survey and interviews which were Stage 1 of the project. The requirements identified in Stage 1 were many and varied and so we have tried to make the new OASIS system more flexible so it will meet more people’s needs.

http://archaeologydataservice.ac.uk/blog/oasis/
Collecting and Validating Data

The HERALD Project
Where are we now?

• HERALD Stage 1 Report – March 2015
• BIAB Stage 1 Report – September 2015
• DES-OASIS Project Grant Application – Sept 2015 (from April 2016 if successful)
• HERALD Stage 2 Project Design inc. BIAB – November 2015
• OASIS Redevelopment 2016 onwards
How did we envision OASIS helping?

Fieldwork input → Print out → Post → Backlog → Local Archive govt

National agency ← Backlog ← Post ← Print out
What information does OASIS collect?

- Site details: Project name, Description, Project dates, Event type, Monument and artefact, Associated Identifiers, Other event info.
- Location: Site name, County, District, Parish, Grid reference, Postcode.
- People involved: Project brief originator, Project design originator, Project director or manager, Project supervisor, Funder, Organisation.
- Archive: Archive recipient, Archive identifier, Archive contents, Media available, Archive notes.
- Report details: Title, Author(s), Date, Issuer or publisher, Place of issue or publication, Description, URL, Report number.
Where does information go?

- Web service
- OAI-PMH
- Research projects
- Grey literature library
- Geophysical Survey database
- Radiocarbon Dating
- Recording Form
- DES Form

Acronyms:
- FAME: Federation of Archaeological Managers & Employers
- MEDIN: Marine Environmental Data & Information Network
- DATA.GOV.UK: Opening up Government
- ADS: Archaeology Data Service
- OASIS: Open Access Service for Interdisciplinary Science
OASIS is like a transit lounge...

(... and currently about this old too!)
Unfortunately we have to deal with some problems too...

There’s a problem of flow in and out of the transit lounge

And in some places it isn’t used enough!
What the survey tells us:

- Ongoing communication
- Flexible workflows
- Different levels of interaction
- Archiving and dissemination of reports
- Inclusion of Museum curators
- Record specialist data when appropriate
- Move focus from archaeology to historic environment
- Include simple to use import and export systems
Different for almost everyone:

- some want to start and complete the record
- some want units to do it
- some want a mini record
- some want a very rich record
- some want access to richer records but may not use them
- some want the museums involved....

...so no one size fits all!!
Records remain ‘OPEN’ with alerts when core fields are completed

- Environmental
- VAG
- Geophysics
- Other modules: BIG ANCHORS
- Pottery

OASIS LITE

OASIS STANDARD

OASIS PLUS

Bibliographic record & Report uploaded by HERs or contractors

Constitutes a basic HER record

Creates a rich metadata record for specialist / project specific recording
Will OASIS allow automated report upload?

Options:

1. Continue status quo with proxy sign-off from Historic England
2. Automated release (recommended option)

In option 2, the following are proposed:

1. Non-reviewing HERs - reports are released directly into BIAB (marked as un-reviewed)
2. HER which is reviewing records/reports - reports are released into BIAB once reviewed. If the reports are not reviewed within a period of time (i.e. one month) then they are released into BIAB (but marked as un-reviewed).

If so, what should the period be? How should these reports be ‘flagged’?
Currently 34538 reports in the library and over 18000 of these come from OASIS.

They are linked to by HERs and Units.

Every report has a DOI
How might HERALD help?

Library of Unpublished Fieldwork Reports

Introduction
The aim of this resource is to make available unpublished fieldwork reports in an easily retrievable fashion. There are currently 30005 reports available and this number is increasing steadily through the OASIS project in England and Scotland.

Staff time saved by no longer needing to:
• check copyright / permission to disseminate on a case-by-case basis
• copy / scan then email / post reports
• help visitors find paper reports

Also:
• Report retrieval time via the HER is almost instantaneous for files held by the ADS
• Future maintenance of this system will be minimal as each report has a DOI
Digital Archiving and Commercial Archaeology

Louisa Matthews
Collections Development Manager