

Digital Data and Archaeology: Management, Preservation and Publishing

ads ARCHAEOLOGY DATA SERVICE

Digital Preservation

Archaeology Data Service
Department of Archaeology
The Henry Moore
University of York
York
YO1 7EP

Dr Katie Green
katie.green@york.ac.uk

06/04/2016 <http://archaeologydataservice.ac.uk> 1

ads ARCHAEOLOGY DATA SERVICE

Digital Preservation

What happens to data once its deposited?



Image © www.digitalbevaring.dk

06/04/2016 <http://archaeologydataservice.ac.uk> 2

ads ARCHAEOLOGY DATA SERVICE

Digital Preservation

Digital Preservation

- Complex process
- Lots of time and effort required
- Standards and guidelines to follow

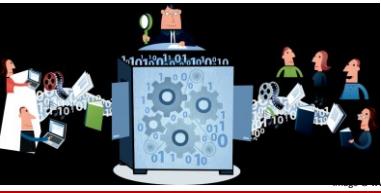


Image © www.digitalbevaring.dk

06/04/2016 <http://archaeologydataservice.ac.uk> 3

ads ARCHAEOLOGY DATA SERVICE

Digital Preservation



Digital Preservation

3 Methods

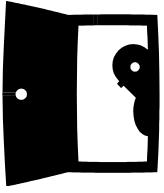
- The Hardware Museum
- The Software Emulator /Virtualisation
- Migration

NB much more intervention is needed than conventional archives

06/04/2016 <http://archaeologydataservice.ac.uk> 4

ads ARCHAEOLOGY DATA SERVICE

Digital Preservation



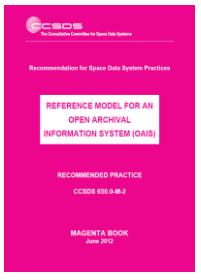
Behind the Scenes at the ADS

- Ensure the multiple and regular backups and the renewal of storage media
 - 23 Virtual Machines
 - Tape backup at University of York
 - Deep Store – UK Data Archive
 - Five year hardware rotation strategy
- Use data migration strategies
- Follow the Open Archival Information System (OAIS) reference model

06/04/2016 <http://archaeologydataservice.ac.uk> 5

ads ARCHAEOLOGY DATA SERVICE

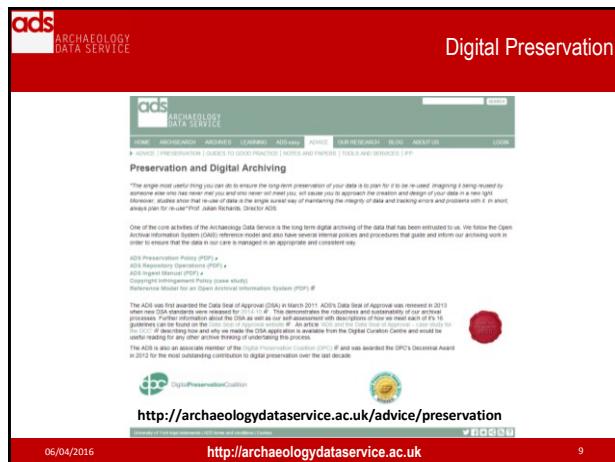
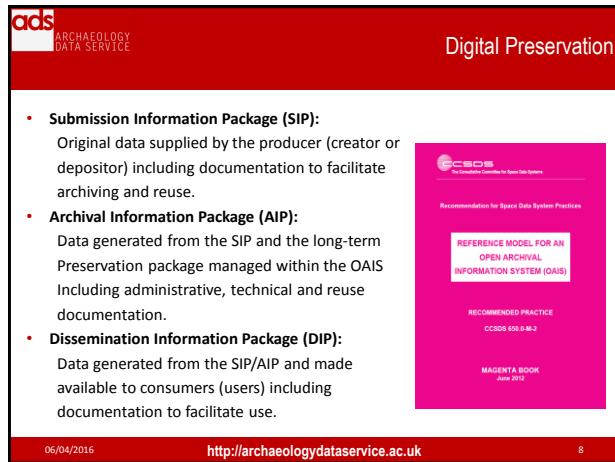
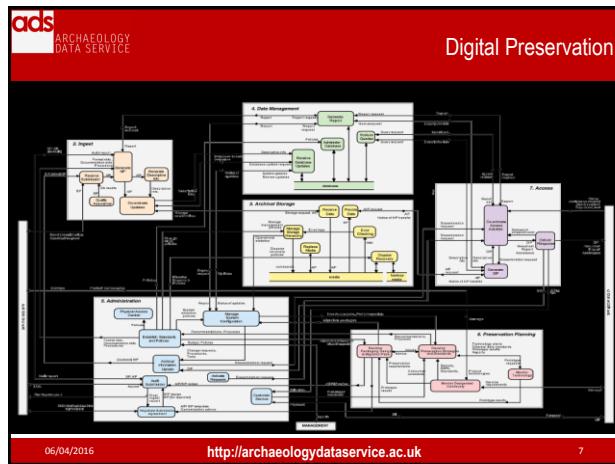
Digital Preservation



Open Archival Information System (OAIS) reference model

- Established by the Consultative Committee for Space Data Systems (CCSDS)
- International ISO standard 14721

06/04/2016 <http://archaeologydataservice.ac.uk> 6



Digital Preservation

Open Archival Information System (OAIS) reference model

```

graph TD
    SIP[SIIP] --> INGEST[INGEST]
    INGEST --> AIP1[AIP]
    AIP1 --> DATA[DATA MANAGEMENT]
    DATA --> ARCHIVAL[ARCHIVAL STORAGE]
    ARCHIVAL --> ACCESS[ACCESS]
    ACCESS --> CONSERVATION[CONSERVATION]
    CONSERVATION --> DIP[DIP]
    AIP1 --> PRESERVE[PRESERVATION PLANNING]
    PRESERVE --> AIP2[AIP]
    AIP2 --> MANAGEMENT[MANAGEMENT]
    MANAGEMENT --> SIP
    
```

06/04/2016 <http://archaeologydataservice.ac.uk> 10

Digital Preservation

Open Archival Information System (OAIS) reference model

06/04/2016 <http://archaeologydataservice.ac.uk> 11

Digital Preservation Standards

Deposit Evaluation

- Intellectual content & potential interest in their re-use
- Viability of data management, preservation, and distribution
- Other suitable archives?
- Authority to deposit the data
- Material is 'complete'
- Digital form in preferred file format – consult repository websites
- Sufficient project documentation and file Metadata

06/04/2016 <http://archaeologydataservice.ac.uk> 12

Digital Preservation

[HOME](#) | [PRESENTATION](#) | [GUIDES TO GOOD PRACTICE](#) | [NOTES AND PAPERS](#) | [TOOLS AND SERVICES](#) | [PP](#) | [LOGIN](#)

ADVICE

Guidelines for Depositors
Version 2.0 September 2014.

CONTENTS

- [Introduction to the Guidelines](#)
- [Why Deposit Data?](#)
- [Depositing with the ADS](#)
- [What to deposit](#)
- [How to deposit](#)
- [Costs](#)
- [Preparing Collections for Deposit](#)
- [Data Management Plans](#)
- [File Management \(Format, Structure, Naming, Metadata\)](#)
- [Selection and Retention](#)
- [File Format Requirements](#)
- [Documents](#)
- [Databases, Spreadsheets and Statistics](#)
- [Raster Images](#)
- [Geographics and Remote Sensing](#)
- [CAD and Vector Images](#)
- [Geographical Information Systems](#)
- [Video and Audio](#)
- [Virtual Reality](#)
- [Collection-level Metadata](#)
- [Deposit Agreement List](#)
- [Downloads](#)
- [Acknowledgements](#)

Example of completed metadata sheet for Raster Images files deposited with the ADS

File name	Format	Version	Creator	Number of Rows	Number of Columns	Dimensions	File size	Date of file creation	Date of last modification	Number of bytes
test.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test2.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test3.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test4.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test5.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test6.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test7.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test8.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test9.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test10.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test11.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test12.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test13.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test14.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test15.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test16.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test17.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test18.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test19.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000
test20.tif	TIFF	1.0	John Doe	1000	1000	1000x1000	10000000	2014-09-01	2014-09-01	10000000

http://archaeologydataservice.ac.uk 13

Digital Preservation Standards

Open Archival Information System (OAIS) reference model

The diagram illustrates the OAIS reference model as a flow process. It starts with 'INGEST' leading to 'PRESERVATION PLANNING'. This is followed by 'DATA MANAGEMENT' which includes 'ARCHIVAL STORAGE' and 'ADMINISTRATION'. 'ACCESS' leads to 'DISSEMINATION' (DIP). A feedback loop labeled 'MANAGEMENT' connects back to the 'PRESERVATION PLANNING' and 'DATA MANAGEMENT' stages. The process is monitored by 'DESCRITIVE INFO' (SIIP) and 'CONSERVATION'.

http://archaeologydataservice.ac.uk 14

Digital Preservation

Submission Information Package (SIP)

- **Virus check**
- **Media and file readability check**
- **Data resource integrity check**
- **Check file formats suitable for deposit**
- **Documentation completeness check**
- **Data validation and consistency checks**
- **Web interface text check**
- **Copy to data server**

© Digital Preservation Business Case Toolkit

http://archaeologydataservice.ac.uk 15

Digital Preservation

Submission Information Package (SIP)

- Authenticate original version
- Replace spaces with underscores
- Log details of SIP in **Collections Management System**
- Create **checksums**
- Run **Droid** to generate file level metadata
- Store **licence** in AIP directory
- Scan paper documentation
- Acknowledge receipt of data
- Store original media

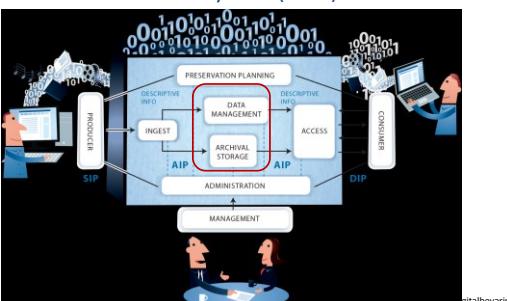


© Digital Preservation Business Case Toolkit

06/04/2016 <http://archaeologydataservice.ac.uk> 16

Digital Preservation

Open Archival Information System (OAIS) reference model



06/04/2016 <http://archaeologydataservice.ac.uk> 17

Digital Preservation

Archival Information Package (AIP)

- Check **licence**
- Check copyright and confidentiality clearance
- Consistency** checks
- Selecting **preservation** and **dissemination** file formats
- Develop a **conversion** plan
- Convert the files
- Validate file conversion
- Metadata update

- Create and store **checksums** for the AIP
- Submit AIP for checking



© Digital Preservation Business Case Toolkit

06/04/2016 <http://archaeologydataservice.ac.uk> 18

Digital Preservation

Internal ADS guidance for all data types

Delivery	Presentation	Description	Notes
Uncompressed Baseline TIFF v.6 .TIF	Portable Network Graphics (.png) or Joint Photographic Expert Group (.jpg)	Any EXIF & IPTC metadata will also need preserving	***
Portable Network Graphics (.png)	Portable Network Graphics (.png)		***
Joint Photographic Expert Group (.jpg / .JPG)	Joint Photographic Expert Group (.jpg / .JPG)	***	***
Graphic Interchange Format (CompuServe).gif	Portable Network Graphics (.png)	***	***
Bit-Mapped Graphics Format (Microsoft).bmp	Portable Network Graphics (.png)	***	***
PhotoCD .pcd	Portable Network Graphics (.png)	***	***
Photoshop (Adobe).psd	Portable Network Graphics (.png)	***	***
CorelPaint .cpt	Portable Network Graphics (.png)	***	***
Adobe Digital Negative .dng	Adobe Digital Negative (.dng) and Joint Photographic Expert Group (.jpg)	***	***
JPEG2000 (.jp2 / .jpx)	JPEG2000 (.jp2 / .jpx)	***	***

06/04/2016

<http://archaeologydataservice.ac.uk>

19

Digital Preservation

Migration

Requires active and continuous management

The Antikythera Survey Project (Collection Id: 1001115)
See this archive on ADS

Update Collection | Go to Tracking Project (190320) | Go to DOI (10.5394/ADS2014)

General | Coverage | Relationships | Accessions | Files | Processes | Web Admin | Archsearch | Export

Processes:

- Migration - Preservation (id: 26995, Microsoft Word for Windows Document 97-2003 - DOC to Microsoft Word for Windows Document 2007 - DOCX)
- Migration - Preservation (id: 26996, ESRI Shapefile - SHP to Geography Markup Language - GML)
- Migration - Preservation (id: 26997, ASCII Text File - CSV to Plain Text File - TXT)
- Editing - Corrective (id: 26998, Plain Text File - TXT to Plain Text File - TXT)
- Migration - Deserialization (id: 26999, ESRI Shapefile - SHP to Zipped Archives - ZIP)
- Migration - Deserialization (id: 26996, ASCII Text File - CSV to Plain Text File - TXT)

06/04/2016

<http://archaeologydataservice.ac.uk>

20

Digital Preservation

Migration

Requires active and continuous management

Migration - Preservation (id: 26995, Microsoft Word for Windows Document 97-2003 - DOC to Microsoft Word for Windows Document 2007 - DOCX)

Date: 13-Jan-2012 to 13-Jan-2012
Description: Conversion of doc file to presentation.
Results:
Input: 1 also file from ADS_preservationArch-1115-10genat23200br
Output:
Hardware: PC
Software: Microsoft Office Word 2007 SP2
Operating System: Microsoft Windows XP Professional 2002
Comments:
Agent: Meera Ray
Accession #: 100330
Added: 13-Jan-2012
Added By: rher103

Migration - Preservation (id: 26996, ESRI Shapefile - SHP to Geography Markup Language - GML)

Date: 13-Jan-2012 to 13-Jan-2012
Description: Conversion of shapefile to gml for preservation, a process that also included the .shp/.dbf/.shx elements.
Results:
Input: 8 shp/dbf/shx files from ADS_preservationArch-1115-10genat23200br
Output:
Hardware: PC
Software: PWTech 2.4.6
Operating System: Microsoft Windows XP Professional 2002
Comments: NB: .prj and .xml files were also added to the preservation versions, along with textual documentation supplied by the depositor.
Agent: Meera Ray
Accession #: 100330
Added: 13-Jan-2012
Added By: rher103

Migration - Preservation (id: 26997, ASCII Text File - CSV to Plain Text File - TXT)

06/04/2016

<http://archaeologydataservice.ac.uk>

21

Digital Preservation

Open Archival Information System (OAIS) reference model

The diagram illustrates the OAIS reference model. It starts with a SIP (Storage Information Package) on the left, which undergoes PRESERVATION PLANNING, INGEST, and DATA MANAGEMENT to become an AIP (Archival Information Package). The AIP then moves through ARCHIVAL STORAGE and ACCESS to become a DIP (Dissemination Information Package). Finally, the DIP is disseminated via DISSEMINATION. The process also includes ADMINISTRATION and MANAGEMENT.

06/04/2016 <http://archaeologydataservice.ac.uk> 22

Digital Preservation

Dissemination Information Package (DIP)

- Convert the files
- Validate file conversion
- Create web interface
- Allocate permanent urls / DOIs
- Pre-Release interface
- Make any depositor changes
- Release Archive
- Publicise Archive

Image © www.digitalbevaring.dk

06/04/2016 <http://archaeologydataservice.ac.uk> 23

Digital Preservation

The Sutton Hoo Research Project 1983-2001

Sutton Hoo

University-led research projects

Major infrastructure programmes

Chamber Tunnel Rail Link Section 1

High Speed

06/04/2016 <http://archaeologydataservice.ac.uk> 24

Digital Preservation

DOIs

A Corpus of Early Anglo-Saxon Buckets
Joan Cook, 2002

Introduction
When Joan Cook died in 2002, the community of Anglo-Saxonists lost one of its greatest intellects. Her extensive notes on the paper plates (especially buckets) she began in 1958 and which she diligently passed through a busy professional life in Museums and University, archaeological units and the like, were a treasure trove of information. Her notes on buckets, which she attended to her research on buckets, standing in being the work to a conclusion with a database planned and a monograph. Her copious notes and drawings over years, together with a database planned and research notes, were left to the University of Oxford. The Director of Archaeology at the University, Dr Tim French, archaeologist, friend and colleague over many years, was asked by the family to take charge of the notes and drawings. He invited me to discuss how the work, as close to completion, could be brought to publication; they were to be given to the Society of Antiquaries of London, the body that had been instrumental in the development of the corpus. I accepted the invitation, and the family agreed to place the whole archive at the Institute of Archaeology and copy it to the University of Oxford. The notes and drawings were copied to the University of Oxford, and the work began in May 2012. The Institute of Archaeology provided the venue for the archive and for the work to be done, decisions on style and the necessary computer work.

Funders
Mark Fitch Fund, The Society of Antiquaries of London, Institute of Archaeology, Oxford

Source Identifiers
<http://dx.doi.org/10.5284/1000389>
doi:10.5284/1000389
How to cite using the DOI

University of Oxford

06/04/2016 <http://archaeologydataservice.ac.uk> 25

Digital Preservation

Persistent Identifiers

Types of persistent identifiers include, 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](http://dx.doi.org/10.5284/1000389)
resolver service prefix suffix
 (assigning body) (resource)

06/04/2016 <http://archaeologydataservice.ac.uk> 26

Digital Preservation

Archive of files and downloads of data

Open formats; open data

06/04/2016 <http://archaeologydataservice.ac.uk> 27

ads
ARCHAEOLOGY
DATA SERVICE

Digital Preservation

Searchable interfaces

06/04/2016 <http://archaeologydataservice.ac.uk> 28

ads
ARCHAEOLOGY
DATA SERVICE

Digital Preservation

Online Publication and Digital Archive
Silchester

06/04/2016 <http://archaeologydataservice.ac.uk> 29

ads
ARCHAEOLOGY
DATA SERVICE

Digital Preservation

Online Publication and Digital Archive

06/04/2016 <http://archaeologydataservice.ac.uk> 30