



ENGLISH HERITAGE

Centre for Archaeology



Preservation Management Procedures Manual

Adrian Brown

Information Management & Collections

DIGITAL ARCHIVING PROGRAMME

PRESERVATION MANAGEMENT PROCEDURES MANUAL

Status: Approved
Access: Open
Version: 1.0
Author: Adrian Brown
Location: S:\Projects\Pr652\Strategy\DAP Manual 1-0.doc

©ENGLISH HERITAGE 2002. This document contains ENGLISH HERITAGE proprietary information and must not be copied reprinted or reproduced in any material form either wholly or in part without the written consent of ENGLISH HERITAGE.

CHANGE RECORD

Version	Author	Date	Status
1.0	Adrian Brown	October 2002	Approved dissemination version

TABLE OF CONTENTS

CHANGE RECORD.....	ii
TABLE OF CONTENTS.....	iii
LIST OF FIGURES	iii
1 INTRODUCTION	1-1
1.1 ABOUT THIS DOCUMENT	1-1
1.2 PRESERVATION MANAGEMENT PROCEDURES.....	1-1
2 EVALUATION.....	2-1
2.1 PREPARATION FOR EVALUATION	2-1
2.2 EVALUATION.....	2-3
2.3 TRANSFER.....	2-4
2.4 DESTRUCTION.....	2-4
3 INGEST	3-1
3.1 INTRODUCTION	3-1
3.2 PRE-ACCESSION.....	3-1
3.3 ACCESSION	3-2
3.3.1 METADATA GENERATION	3-2
3.3.2 PHYSICAL ACCESSION.....	3-6
4 MIGRATION.....	4-1
5 REFRESHMENT.....	5-1
BIBLIOGRAPHY	I
APPENDIX A: DAP EVALUATION RECORD.....	III

LIST OF FIGURES

Figure 1: Preservation Functions	1-2
Figure 2: The DAPPROG Spreadsheet.....	2-2
Figure 3: Evaluation Decision Tree	2-3
Figure 4: Collections and Objects Record	3-3
Figure 5: Migration Pathway Record.....	3-4
Figure 6: Manifestations and Files Record.....	3-6

1 INTRODUCTION

1.1 ABOUT THIS DOCUMENT

This document describes the detailed procedures for evaluating and accessioning data into the Centre for Archaeology (CfA) digital archive system, and for the long-term preservation management of that data. It should be used in conjunction with the CfA Digital Archiving Strategy¹ (DAS), which describes the overall strategy underpinning the CfA Digital Archiving Programme.

The Head of Archaeological Archives is responsible for the maintenance of this document. Comments and proposed amendments should be submitted to the Head of Archaeological Archives, and all changes of substance will be circulated to the Digital Archiving Programme Working Party for comment, prior to implementation.

Paper copies of this strategy are held by the Head of Archaeological Archives, the Computing Manager, and the CfA library.

1.2 PRESERVATION MANAGEMENT PROCEDURES

Detailed procedures have currently been developed for the following areas of operation:

- ☒ Evaluation: These cover the collation of data and its preparation for evaluation, and the actual evaluation process. These procedures will be triggered by the completion of a project, or by the need to evaluate a periodic ‘snapshot’ of dynamic resource.
- ☒ Ingest: These cover the initial ingest of data into the archiving system, and are triggered upon the completion of evaluation.
- ☒ Migration: These cover the periodic migration of archived resources from one file format to another, and are triggered either as part of the ingest process, or by the technology watch process.
- ☒ Refreshment: These cover the periodic refreshment of data from one storage media instance to another, and are triggered either as part of the ingest process, or by the technology watch process.

The functional relationships between these areas are illustrated in Figure 1 (overleaf).

¹ : Brown (2002).

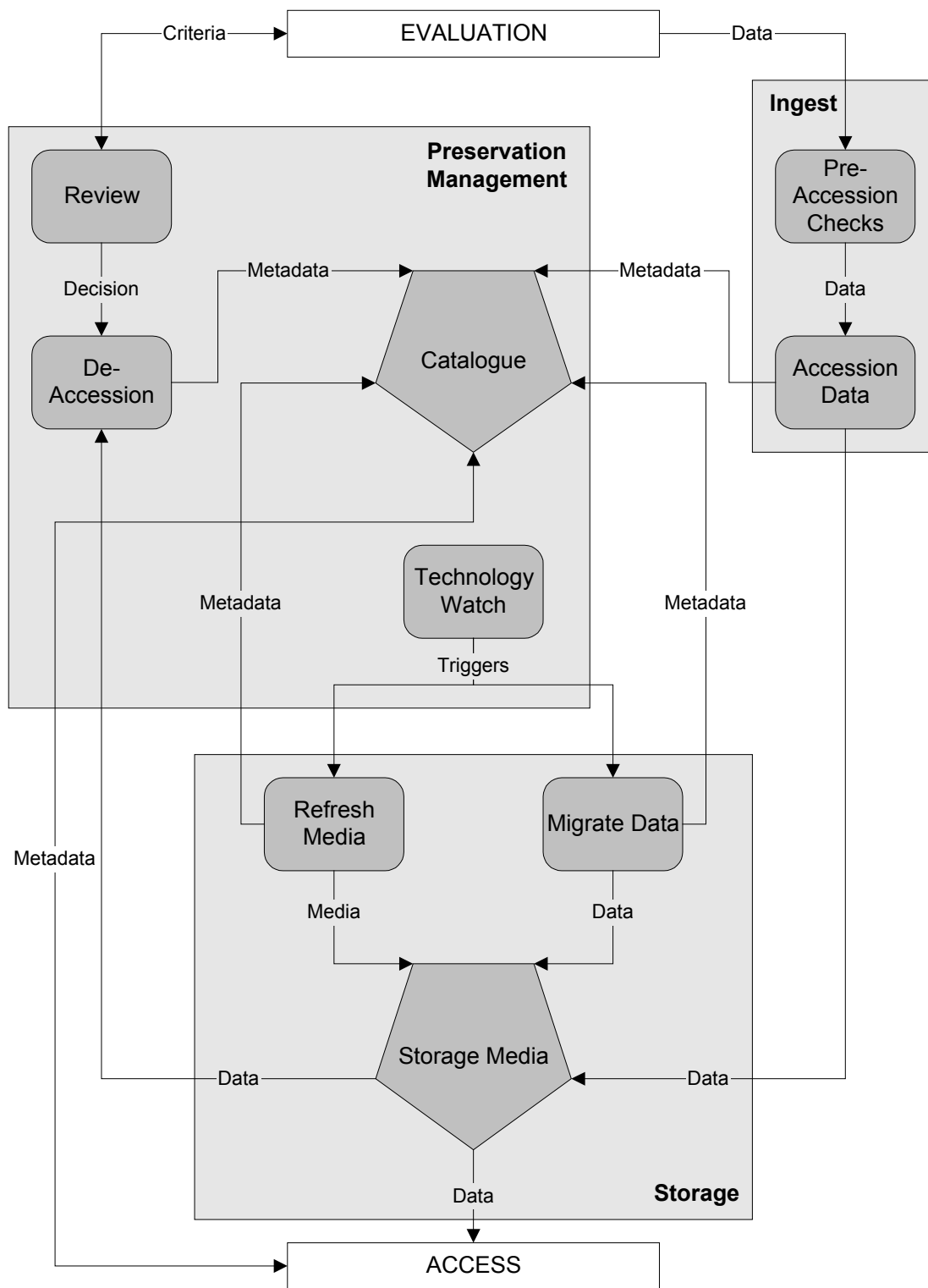


Figure 1: Preservation Functions

2 EVALUATION

2.1 PREPARATION FOR EVALUATION

These procedures should be read in conjunction with DAS Section 2.5.

Prior to evaluation, all data relating to a collection should be collated in one place by the Head of Archaeological Archives. It is essential to ensure that all data is located, and the following sources must therefore be checked:

- ☐ Network Project folders (Drive S:\Projects).
- ☐ Network Head of Archaeological Archives project folders (Drive H:\DAP\Projects)
- ☐ Network personal folders of project team members and data creators (Drive H:\).
- ☐ Network Delilah site folders (Drive O:\Delilah).
- ☐ Existing archive disks.
- ☐ Other disks.
- ☐ Project PhotoCDs.

Once identified, all data should be copied onto the Head of Archaeological Archives' H:\DAP\Collate folder. Project data should be copied into a subfolder named in the form *PrCfA Project No.*, and non-project collections should be copied into a subfolder named after the collection.

A record of this collation process should be maintained by the Head of Archaeological Archives, using the DAPPROG¹ spreadsheet (see Figure 2):

¹ : Currently located at S:\Projects\Pr652\Implement\DAPPROG.XLS

Project No	Project Name	Project	IMC	Team	Dellah	Archive	Other	PhotoCD	Documentation	Comments
1	Winklebury	N	N	N	N	Y	N	N	CJB 2	Collated
2	Kings Langley	N	N	N	N	N	N	N	CJB 2	No digital archive extant
3	Wilderspool	N	N	N	N	Y	N	N	CJB 2	Collated
4	Carlisle, Tanaby Lane	N	N	N	N	Y	N	N	CJB 1,2	Collated
5	Woolley Barrows	N	N	N	N	Y	N	N	CJB 2	Collated
6	Ashley	N	N	N	N	Y	N	N	CJB 2	Collated
7	Burgh-by-Sands, Field 0120	N	N	N	N	N	N	N		No digital archive extant
8	Little Somborne	N	N	N	N	Y	N	N	CJB 1,2	Collated
9	Taunton Priory	N	N	N	N	Y	N	N	CJB 2	No digital archive extant
10	Shaugh Moor, Cairns	N	N	N	N	Y	N	N	CJB 1,2	Collated
11	Carlisle, Fisher Street	N	N	N	N	Y	N	N	CJB 1,2	Collated
12	Hampton Court Palace								CJB 2	Project incomplete
13	Brancaster	N	N	N	N	Y	N	N	CJB 1,2,4	Collated
14	Ospringe, Maison Dieu	N	N	N	N	Y	N	N	CJB 1,2,4	Collated
15	Shaugh Moor, Enclosure	N	N	N	N	Y	N	N	CJB 1,2,8	Collated
16	Dorchester, Wollaston House								CJB 1,2,4	Project incomplete
17	Magiovinium	N	N	N	N	Y	N	N	CJB 1,3,5	Collated
18	Magiovinium	N	N	N	N	Y	N	N	CJB 1,3	Collated
19	Hampton Court Palace								CJB 1,3	Project incomplete
20	Ospringe, Roman Cemetery	N	N	N	N	Y	N	N	CJB 7	Collated
21	Brancaster	N	N	N	N	Y	N	N	CJB 1	Collated
22	Marisco Castle, Lundy	N	N	N	N	Y	N	N	CJB 1,3	Collated
23	Mucking North Ring	N	N	N	N	Y	N	N	CJB 1,3,5,6	Collated
24	Launceston Castle								CJB 1,3,6	Project incomplete
25	Brancaster	N	N	N	N	N	N	N	CJB 1	No digital archive extant
26	Verulamium	N	N	N	N	Y	N	N	CJB 1,3	Collated
27	Layton, DMV								CJB 1,3	Project incomplete
28	Lower Ashmoor Farm	N	N	N	N	Y	N	N	CJB 7	Collated
29	Ardleigh	N	N	N	N	Y	N	N	CJB 3,5,6,8	Collated
30	Stonehenge, Car Park	N	N	N	N	N	N	N		No digital archive extant
31	Bar Point, Isles of Scilly	N	N	N	N	Y	N	N	CJB 3	Collated
32	Carnoon Bank	N	N	N	N	Y	N	N	CJB 1,3,5	Collated
33	Poldowan	N	N	N	N	Y	N	N	CJB 5	Collated
34	Little Bay, Isles of Scilly								CJB 7	Project incomplete
35	Nuneaton Priory	N	N	N	N	N	N	N		No digital archive extant
36	Balksbury Camp	N	N	N	N	Y	N	N	CJB 3,8	Collated
37	Kirkham Priory	N	N	N	N	Y	N	N	CJB 5	Collated

Figure 2: The DAPPROG Spreadsheet

The existence of any documentation relating to the collection, such as Computing Job Books, should also be noted. Finally, the current collation stage should be recorded (unchecked, project incomplete, no digital archive extant, or collated).

2.2 EVALUATION

These procedures should be read in conjunction with DAS Section 2.6.

Evaluation of all resources will be undertaken by the Head of Archaeological Archives, in consultation with the data depositor. The evaluation procedure can be illustrated as a decision tree (see Figure 3):

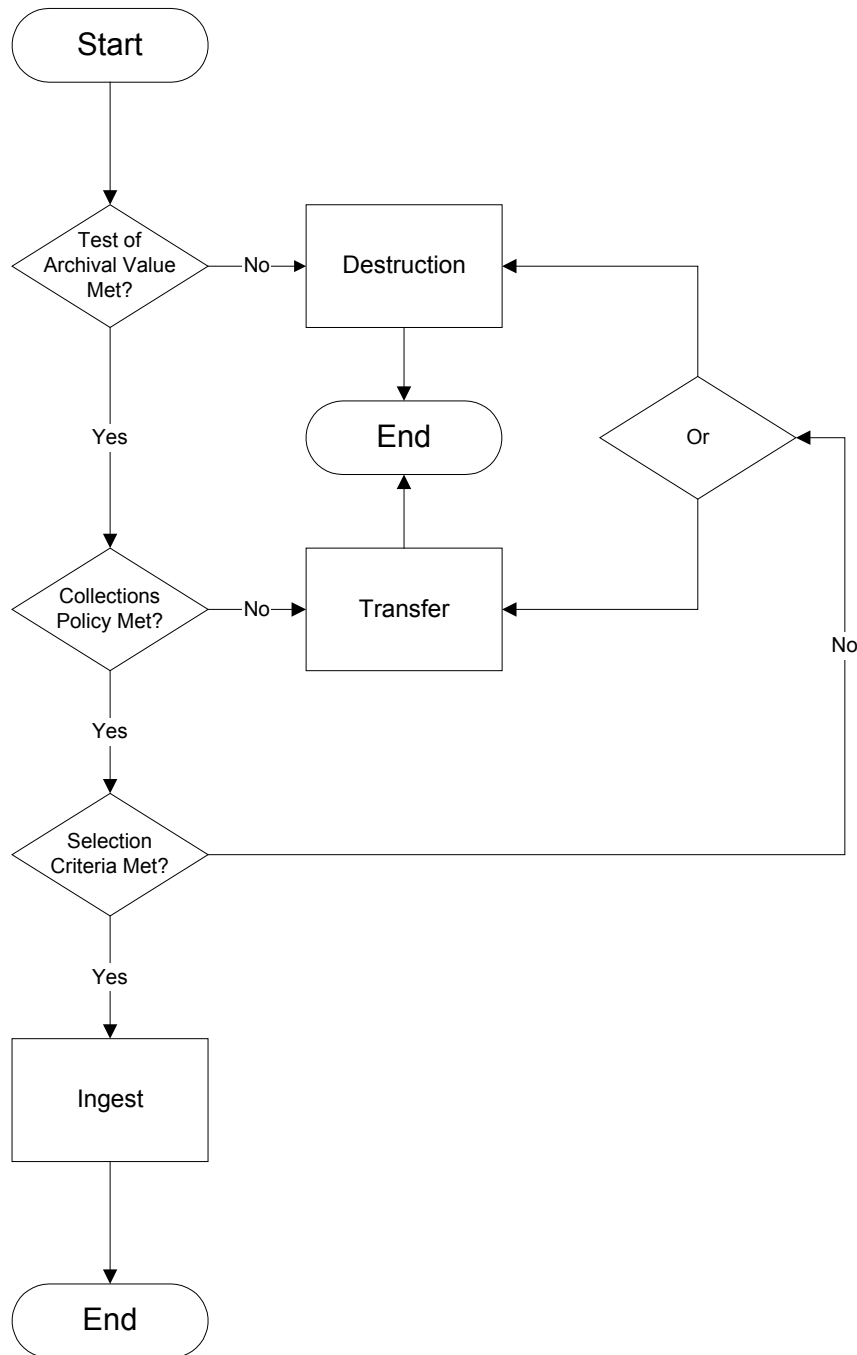


Figure 3: Evaluation Decision Tree

This comprises three sets of evaluation criteria: the **Test of Archival Value** determines whether or not a particular resource is worthy of preservation; the **Collections Policy** defines the types of material which the CfA will preserve; and the **Selection Criteria** describe the essential characteristics which a digital resource must possess to be suitable for preservation and access. These criteria are defined more fully in DAS Section 2.6.

The evaluation process must be fully documented, as a permanent record of the decisions taken. The DAP Evaluation Record form (see Appendix A) should be used for this purpose. For each file evaluated, the following information should be recorded:

- The original file name (e.g. C5.DAT).
- The original file format (e.g. CEU Version 1 Multi-Line Record file).
- A description of the file content, together with any comments (e.g. Contexts 1-105).
- A description of any actions required to ensure that the file meets the Selection Criteria (e.g. ASCII SUB (1A) codes removed at EOF).
- A record of whether the file passes the Test of Archival Value, Collections Policy, and Selection Criteria (Yes/No for each).
- A record of the evaluation decision taken (Archive/Transfer/Delete).

This information forms the basis of the relevant Collection and Manifestation level metadata records on the CfA Metadata System (*CAMS*). Completed Evaluation Record Forms are retained by the AA team as an archive.

2.3 TRANSFER

Procedures for the transfer of data to another body will vary, depending on the requirements of the receiving organisation. However, it is likely that these may include the conversion of data to different file formats and storage media, and the recasting of documentation. It is the responsibility of the Head of Archaeological Archives to liaise with the receiving organisation, in order to ensure that appropriate standards are agreed and observed.

2.4 DESTRUCTION

The decision to destroy any data which has undergone the evaluation process must be taken jointly by the Head of Archaeological Archives and the relevant project manager or Section Head. A record of this decision, together with a list of the deleted data, must be placed on the appropriate registered file.

Ordinarily, data which has been identified for destruction can be deleted in the normal manner within Windows. However, it is important to be aware that deletion does not constitute destruction, and that deleted files may still be retrievable unless either the storage media has been reformatted, or sufficient time has elapsed for the space to be rewritten with new data. It may therefore be necessary to use more rigorous methods to ensure the thorough destruction of confidential information. These may include reformatting of discs, or physical destruction of the media. It is also important to ensure that all extant copies of the data are destroyed. It is the responsibility of the data holder, in liaison with the Head of Archaeological Archives, to ensure that the destruction process is correctly implemented.

3 INGEST

3.1 INTRODUCTION

This section describes the procedures which must be followed to ingest a new digital resource into the archive, and should be read in conjunction with DAS Section 3.2. It applies only to data which has been identified for archive through the evaluation process. There are two discrete stages to the ingest process: pre-accession and accession.

3.2 PRE-ACCESSION

Pre-accession procedures comprise those tasks identified during evaluation as necessary to prepare data for accession into the archive.

- ☐ **Readability:** The readability of each file must be checked. This will usually be accomplished by opening and viewing the file with the appropriate software.
- ☐ **Processing:** Any processing of the data which is required should be carried out. This may include resorting of data, the removal of spurious data (e.g. junk data introduced at the end of a Wordstar non-document file by opening it in document mode), or the aggregation of files. Great care must be taken to ensure that processing does not affect the authenticity of the record. Wherever possible, either the original software or a text or hex editor should be used for this purpose
- ☐ **Renaming:** Where necessary, the file should be renamed in accordance with the filenaming conventions described in the DAS Section 2.4.2. The three-letter file extension should also be checked to ensure that it is the correct one for the file format.
- ☐ **Migration:** If the file is not in an approved archive file format, as specified in DAS Section 3.4.1, then it must be migrated into an approved format, following the migration procedures described in Section 4. The original version must also be retained for accession.
- ☐ **Documentation:** Any accompanying documentation, as specified in DAS Section 2.3.4, must be checked for accuracy and completeness
- ☐ **Virus Scanning:** Each file must be scanned for computer viruses. Although the network anti-virus software automatically scans the servers and workstations on a regular basis, every file should be specifically scanned with the latest version of *Norton AntiVirus* prior to accession.

Each file should then be transferred to the H:\DAP\Archive folder, ready for accession. Project data should be copied into a subfolder named in the form *PrCfA Project No.*, and non-project collections should be copied into a subfolder named after the collection.

3.3 ACCESSION

These procedures cover the actual accession of data in to the digital archiving system, and comprise two stages: metadata generation and physical accession.

3.3.1 METADATA GENERATION

Preservation management metadata for CfA archives is held in the CfA Metadata System database (*CAMS*), which follows the preservation metadata scheme described in DAS Section 3.3.1. Metadata should be created as follows:

Collections

If data is being accessioned which forms part of a collection not previously recorded on *CAMS*, then a new Collection record must be created, as follows:

- 1: From the main switchboard, select “View Collections”.
- 2: Move to a new record.
- 3: Enter the following information: collection name, CAS project no. (if appropriate), CfA project no. (if appropriate), any other reference no. which applies to the collection (if appropriate), a textual description of any associated documentation which exists for the collection (e.g. the relevant Computing Job Book), and a brief textual description of the collection.

Objects

A new object record must be created for each object being accessioned, as follows:

- 1: From the main switchboard, select “View Collections”.
- 2: Move to the appropriate Collection record.
- 3: Click on the Objects subform and move to a new record.
- 4: Enter the following information: object name, object type (select from list), name of object creator, date of creation, and CfA Archive Index No. (select from list, the associated Index Name will be filled in automatically).
- 5: Click on the Description tab and enter a textual description of the object.
- 6: Click on the Rights and Access tab and enter the following information: rights holder, license details, copyright statement (select from list), security level (select from list), access conditions (select from list).
- 7: Click on the Ingest and De-Accession tab and enter the following information: ingest date and ingest notes (record details of any pre-accession processing or other relevant information).

8: Click on the Related Objects tab. If there are any related objects in the archive, move to a new record, enter the object name (select from list) and any relevant notes on the nature of the relationship. In some cases, this information may need to be entered retrospectively. A related object record can be viewed by clicking on the button to the right of the object name.

Figure 4: Collections and Objects Record

Migration Pathways

If an object manifestation has been created using a migration pathway not previously documented on CAMS, then a new Migration Pathway record must be created prior to creating the Object Manifestation record, as follows:

- 1: From the main switchboard, select “View Migration Pathways”.
- 2: Move to a new record.
- 3: Enter the following information: migration pathway name (in the form *source file format name – target file format name*), any general comments on the pathway, the name of the person certifying the pathway for use within the archive, and the date on which the pathway was certified for use. The Pathway No. field will be automatically filled in by the database.

4: For the source and target file formats, enter the following information: format name, a short description of any existing documentation, a short description of any encoding scheme used, a short description of any compression type used, any other notes on the format, and the three letter file extension used.

5: For the test file used, enter the following information: filename (select from list), the source of the file, and any notes about the file, such as the software used to create it.

6: For the conversion method used, enter the following information: the full name and version of the software used, a note of any conversion errors encountered, and a note of any other relevant factors. For multi-stage migration pathways, a separate conversion method record should be created for each stage, in order.

The screenshot shows the 'Migration Pathways' application window. At the top, the title bar reads 'Migration Pathways'. Below the title bar, the main window has a header 'Migration Pathways'. The interface is divided into several sections:

- Pathway Information:** Pathway No.: 2, Name: Windows Bitmap - TIFF 6.0
- Source Format:** Name: Windows Bitmap, Documentation: Microsoft/IBM, Encoding: (empty), Compression Type: None, Notes: (empty), Extension: BMP
- Target Format:** Name: Tagged Image File Form, Documentation: Adobe 1992, Encoding: (empty), Compression Type: None, Notes: (empty), Extension: TIF
- Test File:** Name: Image1.bmp, Source: Created by CfA 2001, Notes: Created using Windows 95 Photoeditor
- Conversion Methods:** Software: Equilibrium DeBabelizer Pro 4.5, Errors: None noted, Notes: Batch conversion method
- Comments:** (empty text area)
- Certification:** Certified By: Adrian Brown, Date: 22 March 2002

At the bottom of the window, there are navigation controls for records. The first record is shown as 'Record: 1 of 1' and the second as 'Record: 2 of 2'.

Figure 5: Migration Pathway Record

Manifestations

A new Object Manifestation record must be created for each manifestation of an object which is being accessioned. This will comprise a record for the original manifestation, and optional records for any other current manifestations created through migration as part of the pre-accession process. The record should be created as follows:

1: From the main switchboard, select “View Collections”.

- 2: Move to the appropriate Collection record.
- 3: Click on the Objects subform and move to the appropriate record.
- 4: Click on the View Current Manifestations button. This will open the Object Manifestations record in a new window.
- 5: If this is the first manifestation record being created for an object, the window will open on a new record, otherwise click on the Manifestations record and move to a new record.
- 6: When creating an original manifestation, enter the following information: manifestation name (this must be unique to the database), a textual description of any technical dependencies, and set the Manifestation status to “Original” (select from list). The Manifestation ID will be generated automatically by the database.
- 7: When creating a subsequent manifestation, enter the following information: manifestation name (this must be unique to the database), the Manifestation ID of the previous manifestation, a textual description of any technical dependencies, and the manifestation status (select from list). The Manifestation ID will be generated automatically by the database. In the Migrations subform, enter the name of the migration pathway used to create the current manifestation (select from list), the migration date, and any relevant notes on the migration process. The selected Migration Pathway record can be viewed by clicking the button to the right of the migration pathway name.

Files

The generation of file level metadata is largely automated on *CAMS*. Although the generation process is designed to manage any number of nested subfolders, it is not recommended that metadata be generated for more than one collection at a time, to avoid confusion. File metadata is created as follows:

- 1: From the main switchboard, select “Import File Metadata”.
- 2: Enter the full pathname of the folder containing the files to be accessioned. This will normally be the relevant collection folder in H:\DAP\Archive. All subfolders will automatically be read as well.
- 3: Click OK.
- 4: *CAMS* will create and display a temporary table (“file_temp”) containing the pathname, filename, file size (in KB) and file date of each file in the named folder(s).
- 5: For each file, enter the relevant manifestation name and file format (select from lists).
- 6: Click OK.
- 7: *CAMS* will ask if you wish to append this data to the database.

8: After checking that the data is all correct, click OK.

9: CAMS will report that metadata import has been completed successfully.

10: Click OK.

The file level metadata will now appear in the database, associated with the relevant manifestation record.

The screenshot shows the 'Object Manifestations' application window. The top section displays 'Manifestation ID: 1' and 'Manifestation Name: Winkpot1'. Below this, 'Status' is set to 'Original' and 'Previous Manifestation' is empty. The 'Migrations' section includes fields for 'Pathway', 'Date', and 'Notes'. The 'Technical Dependencies' section contains the text: 'Created using CEU Version 1 software suite, comprising data file and associated RDF. Requires text editor to view.' The 'Files' section is active, showing details for a file named 'P1C.RDF' with a size of 0.4 KB, dated 02 April 1991, and an MD5 checksum of 98478b17d2c116b9822581f967fd6fee. There are checkboxes for 'Preservation Master', 'Security Master', and 'Working Copy'. The interface includes navigation controls for records, with 'Record: 1 of 1' for the migration section and 'Record: 1 of 2' for the files section, and a total of 98 records at the bottom.

Figure 6: Manifestations and Files Record

3.3.2 PHYSICAL ACCESSION

Once all relevant metadata has been created, the physical accession of the data can take place.

1: A fixity value for authentication purposes must be generated for each file using the *MD5Summer* freeware utility¹. One .md5 file, containing the relevant values, should

¹ : Available to download from <http://homepages.ihug.co.nz/~floydian/md5/about.html>.

be generated for each collection being accessioned, and stored in the same folder as the files to which it refers (within H:\DAP\Archive).

2: The relevant .md5 file should be opened using a text editor (e.g. *Notepad*). Each checksum value should be copied and pasted from here into the Fixity Value field of the relevant *CAMS* File record. The Fixity Method field should be set to “MD5 Checksum” (select from list).

3: The data should be written to archival storage media in accordance with the refreshment procedures described in Section 5, to create a Preservation Master, Security Master, and Working Copy.

4: If a new storage media instance is being used, then a new Storage Media record must be created on *CAMS*, as follows:

4a: From the main switchboard, select “View Storage Media”.

4b: Move to a new record.

4c: Enter the following information: media type (select from list), manufacturer’s type name (select from list), file structure (select from list), status (select from list), location (select from list), and any other notes on the media instance. The Media ID will be automatically assigned by the database.

5: For each file being accessioned, in the relevant File record, enter the Media ID of the Preservation Master, Security Master, and Working Copy. The full storage media details can be viewed by clicking on the button to the right of each field.

6: The storage media should then be transferred to archival storage as follows:

Preservation Master: Guardhouse datasafe.

Security Copy: Laying-Out Shed datasafe.

Working Copy: IMC Office (House 4) or Archaeometry Office (House 3) for geophysical survey data.

4 MIGRATION

This section should be read in conjunction with DAS Section 3.4.6. The migration process is triggered under the following circumstances:

- ☐ When the Technology Watch (see DAS Section 3.3.2) determines that the current archive file formats are about to become obsolescent, or that more appropriate formats are now available.
- ☐ As part of the pre-accession processing of new data.

Migration is the responsibility of the Head of Archaeological Archives. Detailed migration procedures will differ for each migration pathway, and will be developed as part of the certification of each pathway. However, the generic migration process comprises the following stages:

1: Source and target folders should be created – migrated data should always be saved in a separate location to the source data.

2: Data should be migrated using the process defined in the migration pathway for that file type.

3: A random 5% sample of the files generated should be validated against the originals for authenticity.

4: Create a new Manifestation record on *CAMS* for each migrated object (see Section 3.3.1).

5: Generate new File records on *CAMS* for each migrated file (see Section 3.3.1).

6: An updated fixity value must be generated for each file, in accordance with the procedures described in Section 3.3.2, procedures 1 and 2. Since one .md5 file is stored for each archival collection, this procedure must be carried out for all files in the collection. The relevant Fixity values must then be updated on *CAMS* (See Section 3.3.2).

5 REFRESHMENT

This section should be read in conjunction with DAS Sections 3.4.5. The refreshment process is triggered under the following circumstances:

- ☐ When the Technology Watch (see DAS Section 3.3.2) determines that the life expectancy of the current storage media is about to be reached, or that more appropriate media are now available.
- ☐ If Technology Watch triggers the migration cycle (see DAS Section 3.4.6).
- ☐ Upon accession of new data.

Refreshment is the responsibility of the Head of Archaeological Archives and comprises the following stages:

1: Prior to use, all new media should be inspected for damage. The media should be acclimatised for at least 24 hours in the operational area before use, in order to allow the effects of temperature and relative humidity to stabilise. Before use, DLT cartridges should be subjected to a full forward and rewind cycle to equalise tape tension.

2: DLT cartridges should be written on the Digital Archiving Workstation, using the Windows 2000 Backup applet (Start – Programs – Accessories – System Tools – Backup). The “Verify data after backup completed” option should be checked. After writing a DLT cartridge, the tape should be retensioned, fully rewound, and the write-protect switch set.

3: CD-Rs should be written on the Digital Archiving Workstation, using Roxio Easy CD Creator 5. Files should be written using the ISO 9660 file structure with Joliet extensions¹, and the “Test and Create” option should be selected. The integrity of the data must then be verified, by performing a bit-level comparison between the source and copy versions of each file. This should be undertaken using the DOS File Compare utility (in MS_DOS mode type “`fc /b source pathname*. * target pathname*. * >fclog.txt`”, then check the “fclog.txt” file for errors using a text editor such as *Notepad*). All media must be scanned with the latest version of *Norton Antivirus* after writing.

4: Once the new Preservation Master, Security Master and Working copies have been produced, the *CAMS* database must be updated in accordance with Section 3.3.2, procedures 4 and 5.

¹ : This is also defined in the equivalent British Standard, BS 7061:1989.

BIBLIOGRAPHY

Brown, A, 2002, *Centre for Archaeology Digital Archiving Strategy (Version 2.0)*, English Heritage

BS 7061: 1989, *Volume and file structure for compact read only optical discs (CD-ROM) for information interchange*

APPENDIX A: DAP EVALUATION RECORD

