



## **Digital Preservation Policy**

### **1. Purpose and Scope**

The purpose of this policy is to state and communicate the principles that guide the digital preservation activities of Northumberland Archives at Woodhorn and Berwick-upon-Tweed Record Office in fulfilling all preservation aspects of their statutory duties.

Digital preservation is critical to many aspects of the work of Northumberland Archives. Northumberland Archives is committed to managing the long-term preservation and access to born digital records and digital surrogates of the historical resources in its care.

The bulk of Northumberland Archives holdings have so far been received in traditional formats, such as paper or parchment. However we expect more items to be received into the archive as born digital records, comprising digital images, sound recordings, moving images, word processed documents, financial spreadsheets, archaeological surveys prepared using CAD (Computer Aided Drawing) and correspondence in electronic mail formats is also expected.

The amount of born digital records has increased dramatically over the last 5 years. With the continued dominance of the desktop personal computer in the workplace and educational establishments, huge increases in processor power and continuous developments in storage capacity and media, the amount of information created will continue to rise exponentially.

### **2. Strategic Aims**

This policy forms the basis for the preservation of all aspects of born digital records and digital surrogates. This also includes surrogate copies of material loaned to the archives.

To address the risk of losing digital records, Northumberland Archives has developed a Digital Preservation Policy and Strategy. The policy outlines the approach to digital preservation, whilst the aim of the strategy is to describe this approach in more detail, including financial requirements, specifications, procedures and staffing needs where appropriate.

The policy should also be used in conjunction with other Woodhorn policies for example 'Northumberland Archives: Preservation Policy' and 'Northumberland Archives: Archive Collection Policy Statement'.

The strategic aims of the service are that:

- Born digital records, including surrogate copies, remain accessible in the long term.
- The data integrity and authenticity of born digital records be maintained throughout the lifecycle by following an active digital preservation programme.

- Copyright and intellectual property rights in born digital records be managed and negotiated with producers ensuring public accessibility.
- Digital preservation methods follow current and international standards.

### 3. Standards and Codes of Practice

Northumberland Archives will aim to be compliant with:

**PD5454:2012** *Guide for the storage and exhibition of archival materials*

**BS4971:2002** *Repair and allied processes for the conservation of documents.*

**BS4783:1988-94** *Storage, transportation and maintenance of media for use in data processing and information storage.*

**BS17799:2000** *Code of practice for information security management.*

**BS6266:2002** *Code of practice for fire protection for electronic equipment installations.*

**BS7083:1996** *Guide to the accommodation and operating environment for Information Technology (IT) equipment.*

### 4. Appraisal, Selection & Acquisition

Appraisal and selection of born digital records will follow the guidelines set out in the Archive Collection Policy. The physical format on which the information resides will not normally determine its selection, however with born digital records the digital format will assume a greater importance since access to view the record is paramount.

Thought must be given to long-term access and preservation concerns at the appraisal and selection stage, as once cataloguing begins, access must be made and migrated versions may need to be prepared. A decision will need to be made as to whether the data being appraised warrants the level of resources required for its effective preservation and continued access.

#### File Formats

At the born digital record creation stage, the file format should be carefully considered, not only for the immediate requirements of the creator but also for its long-term sustainability.

Well established formats tend to have broader and longer-lasting support from software suppliers than those which only have a niche market. Popular formats, which are supported by as wide a range of software suppliers as possible, are therefore to be preferred.

There are 3 main types of Format Standards.

◆ **Proprietary Standards:** e.g. Microsoft Word (DOC). The specification is owned by a company and is not generally made available to others.

◆ **Available Formats:** e.g. Adobe Portable Document Format (PDF). Although still a proprietary standard, the specification is made available to other software producers.

◆ **Open Formats** – e.g. Tagged Image File Format (TIFF). The specification is freely available. Access to data held in these formats does not require any specific viewing software.

## 5. Storage

It is important to be able to retrieve a born digital record in a timely manner. It may be required to answer a Freedom of Information enquiry or a historical enquiry, and an archive repository should be able to retrieve the information required to meet its legal requirements and as part of its customer service operations.

Long-term storage methods include online (network), nearline (shared folder) and offline (CD-R, DVD-R/DVD+R, external hard drive).

A Digital Preservation Strategy should ensure that at least three copies of a born-digital record exist. The original record, a preservation copy and one access copy. For digital surrogates two copies should exist. One preservation TIFF / WAV / AVI saved to DVD+R and one access JPEG / MP3 / MPEG2 saved to the Reprographics folder located on server [\\w2knas1\](#).

The original born digital record will be stored on a server in line with current National Archives guidelines. This will be facilitated via an access restricted folder called DIGITAL\_COLLECTIONS and located on the Archives Shared Folder on server [\\w2k8woodhorndc1\](#).

The preservation copy of a born digital record is created during the back-up procedure as part of the Service Level Agreement with Northumberland County Councils Information Services.

The access copy of a born digital record or digital surrogate will be available off line via dedicated public access PC's at Woodhorn and Berwick, or on external hard drives / CD-R / DVD+R as appropriate.

All digital media, for example the physical case and disc should be stored in the Woodhorn Digital Media Store or the Berwick equivalent.

## 6. Preservation and Data Integrity

No electronic storage medium can be considered archival, irrespective of its physical longevity – technological obsolescence is inevitable. The need to refresh born digital records onto new servers is unavoidable for the foreseeable future. A cost implication will manifest itself here.

Digital surrogate preservation copies are saved to DVD+R. This type of disc has a life expectancy of between 5 and 10 years if stored in optimum conditions. These discs will be checked at regular intervals, typically every 5 years, for signs of physical deterioration. If signs of deterioration are detected the contents of the discs will be refreshed onto new discs. A cost implication will manifest itself here.

A process for checking the data integrity of born digital and digital surrogates using MD5 checksum algorithms, as well as migrating data to open standards will need to be established and included in Northumberland Archives Digital Preservation Strategy.

## **7. Access and Use**

### **Digital still images (surrogates)**

National Archive guidelines recommend that a preservation copy is saved as TIFF during digitisation. Access copies should be created as JPEG and saved to external hard drive / DVD+R as appropriate with access provided via a dedicated public access PC at Woodhorn and Berwick. In addition, access copies can be provided via Northumberland Archives online catalogue but only if the item has been catalogued.

### **Digital still images (born digital)**

National Archive guidelines recommend that a TIFF is created. Access copies can be saved as JPEG or migrated to PDF and saved to external hard drive / DVD+R as appropriate with access provided via dedicated public access PC's at Woodhorn and Berwick. In addition JPEG copies can be accessed via Northumberland Archives online catalogue but only if the item has been catalogued.

### **Digital sound recordings (born digital & surrogate)**

National Archive guidelines recommend that an uncompressed WAV is created with a compressed MP3 version generated for access. The compressed version will be saved to external hard drive / DVD+R as appropriate with access provided via dedicated public access PC's at Woodhorn and Berwick. In addition, edited copies can be accessed via Calm if the item has been catalogued.

### **Digital moving image**

Born digital is generally encrypted as movie DVD. These discs, including any packaging are physically stored in the digital store at Woodhorn or Berwick equivalent. A migrated AVI version is prepared by archives staff and saved to DVD+R as a preservation copy, with an access MPEG2 copy provided via dedicated public access PC's at Woodhorn and Berwick. Digital moving image surrogates are prepared from Cine film or video cassette by external contractor or archives staff with the resultant AVI processed in the same way as born digital.

### **Other documents**

Microsoft formats, such as Excel and Powerpoint, are currently supported by Northumberland Council and are therefore recommended for use by digital producers. The original record is saved to server along with migrated PDF versions. Access copies are saved to external hard drive / DVD+R as appropriate with access provided via dedicated public access PC's at Woodhorn and Berwick.

### **Portable Document Formats (PDF)**

This is the preferred format for accessioning into the archives. Migration to PDF is also recommended for other documents which may not be supported over the long term. Access copies are saved to external hard drive / DVD+R as appropriate with access provided via dedicated public access PC's at Woodhorn and Berwick.

### **Specialist Formats**

Formats which may be unsuitable for long-term preservation are specialist proprietary software e.g. family history programs and computer aided drawings (CAD). This is because the resources required to access the original record and to migrate it to PDF may not warrant the expenditure. Careful consideration should be made as to whether to accession such formats at the deposit phase.

### **Online Accessibility**

It is possible to provide online access via CALM for images of documents of born-digital accessions but only where a unique catalogue records exists. However, on the advice of its suppliers Axiel, CALM is unable to provide access to non-image based documents, e.g. PowerPoint, Excel or CAD. Migrating to PDF offers a solution since this format can be accessed via CALM and requires further investigation.

## **8. Woodhorn Operations**

The creation of digital records in order to carry out the operational activities of Woodhorn Trust departments should be undertaken with due thought and diligence. The use of appropriate and supported software must be followed at all times. A file plan should be used to ensure effective saving of, and access to, any records prepared.

The use of the Woodhorn Shared Drive / Archives Network or Reprographics Network must be used when creating born digital material. Storage of items on individual PC's should be avoided since no backup procedures are in place.

Migration of born digital records which were created using older software packages to currently supported software should be undertaken as a priority.

The use of standalone PC's should be discouraged since no anti-virus, or back-up procedures will be in place for the records stored on them.

## **9. Roles & Responsibilities**

The cooperation of all with a role in the generation, use, management, distribution and preservation of digital data is essential.

### **Director of Woodhorn**

Should ensure that individual employees follow the guidelines set out in both the Digital Preservation Policy and Strategy to ensure effective preservation of born digital and digital surrogate data.

By including digital creation/curation and preservation as part of their induction process.

Including reference to the digital policies in the Woodhorn Trust Manual.

**Director of Woodhorn****Role**

- ◆ To ensure all staff at Woodhorn adhere to the digital policies.

**Responsibilities**

- ◆ To have the policies communicated to the Assistant Director of Woodhorn and the Divisional and Executive Directors of Northumberland Council.
- ◆ To have the policies passed at Management Board and Woodhorn Board of Trustees.
- ◆ To ensure sufficient financial support for the preservation technologies needed to comply with the digital policies.

**Head of Collections**

The Head of Collections will ensure that standards and policies are prepared and updated to maintain currency.

By reviewing standards & policies.

By consulting TNA where necessary.

**Head of Collections****Role**

- ◆ To ensure effective standards and policies are in place to preserve born digital and digital surrogates arriving in the Collections Service.

**Responsibilities**

- ◆ Review and revise standards and policies with the Berwick Borough Archivist.
- ◆ Ensure all policy documents contain version control.
- ◆ Assign appropriate staff to implement the digital curation & preservation policies.
- ◆ Ensure staff are trained in digital preservation techniques.
- ◆ Ensure environmental conditions are maintained for the storage of digital materials.
- ◆ Ensure digital equipment is fit for purpose.

## Archivists

The Archivists will advise Woodhorn Trust staff regarding which material should be digitised and preserved.

They will also advise depositors regarding digital preservation of their collections.

### **Archivists**

#### Role

- ◆ To ensure that digital records are preserved in the most appropriate format and no information is lost.

#### Responsibilities

- ◆ To train staff in digital preservation & storage techniques, in consultation with the Digitisation Officer.
- ◆ To assist depositors with the digitisation of material.
- ◆ To assist depositors with the transfer of born digital materials for permanent preservation.
- ◆ To advise depositors and Northumberland Council staff in maintaining currency of digital preservation.
- ◆ To check and maintain the currency of preserved digital material.
- ◆ Apply cataloguing/listing techniques that will allow digital material to be identified when upgrading and/or migration is necessary.

## Berwick Archivist

The Berwick Borough Archivist will identify appropriate material that is to be digitised and preserved. The Archivist will also ensure safe preservation of appropriate born-digital data which is deposited with the Berwick-upon-Tweed Record Office.

The Archivist will advise depositors regarding digital preservation of their collections.

### **Berwick Borough Archivist**

#### Role

- ◆ To ensure that digital material deposited at the Berwick-upon-Tweed Record Office is preserved in the most appropriate format and no information is lost.

#### Responsibilities

- ◆ To assist depositors with the digitisation of material.
- ◆ To assist depositors with the transfer of born-digital material for permanent preservation.
- ◆ To advise depositors and Berwick-upon-Tweed Council staff in maintaining currency of preservation.
- ◆ To check and maintain the currency of preserved digital material at the Berwick-upon-Tweed Record Office.
- ◆ Apply cataloguing/listing techniques that will allow digital material to be identified when upgrading and/or migration is necessary.
- ◆ To work with the staff at Woodhorn in relation to any digital issues.

## Digitisation Officer

The Digitisation Officer will offer technical and practical advice on the preservation of born digital and digitised records.

Including help and advice to members of the public.

Help and advice to Woodhorn and Berwick-upon-Tweed Record Office.

Help and advice to Northumberland Council staff.

### Digitisation Officer

#### Role

- ◆ To manage digital and digitised services at Woodhorn and Berwick.
- ◆ Offer assistance to archivists regarding technical aspects of preserving born digital and digitised material.
- ◆ To act as Woodhorn's representative for digitisation and digital preservation at internal and external meetings.

#### Responsibilities

- ◆ Maintain technical equipment.
- ◆ Advise Head of Collections of new equipment on the market.
- ◆ Manage paid-for digitisation service.
- ◆ Assist archivists in training staff in digital preservation techniques.
- ◆ To report to Head of Collections any deviation from PD5454 & BS4783 (environmental storage conditions of the Digital Media Store.)
- ◆ To assist with the preservation and migration of more complex born digital accessions.

## All employees of Northumberland Archives

All employees must ensure that any born digital or digitised data is created, stored, maintained and preserved according to the terms of the digital policies.

They must use the corporate classification system and file plan set up by the Records Management Service.

#### Role

- ◆ To adhere to the digital policies in operation.

#### Responsibilities

- ◆ To create only those records that are necessary for business needs.
- ◆ To identify the master copy of each document.
- ◆ To store documents according to the corporate classification scheme and retention policies.
- ◆ To transfer identified documents to Northumberland Archives at the end of their administrative life.
- ◆ To store documents in open accessible formats.
- ◆ To seek advice on the storage formats for long term preservation.
- ◆ To seek training in IT technologies to ensure compliance with the digital policies.



## All employees of Woodhorn Trust

All employees must ensure that any born digital or digitised data is created, stored, maintained and preserved according to the terms of the digital policies.

They must use the corporate classification system and file plan set up by the Records Management Service.

### Role

- ◆ To adhere to the digital policies in operation.

### Responsibilities

- ◆ To create only those records that are necessary for business needs.
- ◆ To identify the master copy of each document.
- ◆ To store documents according to the corporate classification scheme and retention policies.
- ◆ To transfer identified documents to Northumberland Archives at the end of their administrative life.
- ◆ To store documents in open accessible formats.
- ◆ To seek advice on the storage formats for long term preservation.
- ◆ To seek training in IT technologies to ensure compliance with the digital policies.

## 10. Training and development of staff

All staff at Woodhorn and Berwick-upon-Tweed should receive training in digital management and preservation techniques and these skills will require updating periodically to maintain currency.

Professional archives, library and museum staff are encouraged and supported to undertake continuing professional development in digital management and preservation, using external and in-house providers.

In order to deliver many aspects of this policy extensive training by the Digitisation Officer will need to be delivered. Many of the processes are new and are open to change following revised guidelines issued by the National Archives and the Digital Preservation Coalition. The very nature of digital preservation requires a fluid approach to training needs and whilst this policy outlines the commitment of Northumberland Archives in delivering digital preservation, resource for continued training and development must be provisioned.

## 11. Disaster Plan

Woodhorn and Berwick-upon-Tweed Record Office will maintain a Disaster Plan and Disaster Box which will be regularly checked and updated.

Selected staff will receive Disaster Plan training which will be cascaded to other staff.

A Disaster Recovery Plan should be agreed with Computer Services at Northumberland Council.

Vital records should be identified and additional back-ups should be created and stored in more than one format.

## **12. Preservation of Northumberland Council & District Council Records**

Northumberland Archives will ensure that Northumberland Council's or District Council born digital and digital surrogate records which have been selected for permanent preservation are stored at Woodhorn or Berwick, and will be maintained in a manner suitable for long-term preservation.

Northumberland Archives will advise Council Directorates on the appropriate methods for the creation and preservation of their records in non-traditional formats, including Internet, Intranet, digital media, audio and video tapes.

We will also advise on the legal admissibility of scanning and storage of data in digital format.

## **13. Advice and Preservation of records from external customers**

Northumberland Archives will encourage good practice and actively support owners of privately deposited items at Woodhorn in the management of their digital records from their creation through to eventual deletion or permanent preservation.

We will act as a centre of excellence and provide advice and guidance to all record creating bodies, both pro-actively and reactively.

The Service has provided records management and archive advice to both internal and external bodies for a number of years and this will now be extended to electronic sources through digital preservation.

We will act as a signpost to good records management and archive practice and to standards and documentation produced by others at the forefront of digital preservation, e.g. The National Archives, the Digital Preservation Coalition, The British Library and the UK Data Archive.

## **14. Review**

This policy will be reviewed every 3 years.

## **15. Glossary of Terms**

AVI	A specific file format which is used in the creation of moving images.
Born Digital	Documents that are created electronically and remain in electronic format.
CALM	A software application used by the Archives Sector in order to manage and catalogue the collections held in its care.

Computer Aided Drawing (CAD)	The result of using a software application to prepare 2D and 3D drawings. Primarily used in architectural design and archaeological excavations.
Data Integrity	Ensuring that the data stream is intact and has not suffered from any type of data loss.
Digital Media	The physical carrier of the digital record, for example a Compact Disc.
Digital Preservation	The management of digital information over time.
Digital Surrogates	The collective name for the preparation of digital versions using scanning or digital photography of analogue sources, for example a scan of a postcard.
Digitise	To transform analogue information into a digital format. E.g. scanning.
DOC	The file extension associated with text documents created using Microsoft Word Processing software.
Electronic Storage Medium	The physical hardware used to hold digital records. For example a server.
Excel	A software package owned by Microsoft which can be used to create financial spreadsheets.
JPEG	Joint Photographic Expert Group. An image file format which stores the digital information in a compressed state allowing the file size to be greatly reduced.
MD5 Checksum Algorithm	A checksum is a fixed-size single piece of data computed from an arbitrary block of digital data for the purpose of detecting accidental errors that may have been introduced during its copying or transfer to storage.
Migrated Versions	To move electronic information into another file format, possibly one that is more open source software independent. For example Microsoft Word saved to Adobe PDF.
Movie DVD	This type of disc is prepared in such a way that the copying of it is prohibited. The physical structure of the disc is robust and life expectancy is far greater than DVD+R.
MP3	A compressed file format used in audio production.
MPEG2	A compressed file format used in Video / audio production.
Near Line	Data which is stored near line is held in second level storage which is cheaper. Files take slightly longer to

	retrieve. E.g. Shared folders.
Offline	Data which is stored offline is not immediately accessible and may take some time to load, possibly from external media. E.g. data memory stick.
Online	Data which is stored online is immediately accessible. It is usually stored on a network or the hard drive of a PC.
PDF	The Portable Document Format is recommended for long term storage with guaranteed access with open source software over the long term.
Physical Format	The physical carrier of the digital record, for example a Compact Disc.
PowerPoint	A software package owned by Microsoft which can be used to create presentations which utilise, text, image, sound and moving images.
Record	Information produced or received in the initiation, conduct or completion of an institutional or individual activity and that comprises content, context and structure sufficient to provide evidence of the activity.
Server	A large computer which can be accessed by multiple users at the same time. Servers generally have large storage capacities and are the recommended storage medium for preservation copies of born digital and digital surrogate records.
TIFF	Tagged Image File Format is the archival standard used to save images in a non-compressed way. TIFF images are generally much larger in file size and the format is recommended by the National Archives.
WAV	Waveform Audio File Format or more commonly known as WAV due to its filename extension is a Microsoft and IBM audio file format standard for storing an audio bit-stream on a PC. It is an uncompressed audio file format and is the recognized archival standard for long term preservation.

Date: February 2016  
Review date: February 2019