


BCC Digital Archiving SOP

SOP Title	BCC Digital Archiving Procedures for Project Owners
SOP Reference	BCC-ISM-001-SOP
Version Number	1.0
Approval Date	17/01/2019
Effective Date	31/01/2019
Review Date	31/01/2019

Author:	Madalyn Hardaker <small>Name</small>	Information Governance Lead <small>Position</small>
Reviewed by:	Zoe Leech	Research Systems Analyst; Digital Archivist
	Michelle Sleeth <small>Name</small>	Clinical Project Manager; Deputy Head of Operations CRUK Prevention Trials Unit <small>Position</small>

Approved by	Jonathan Croft <small>Name</small>	Head of Research Computing <small>Position</small>
	 <small>Signature</small>	17/01/2019 <small>Date</small>

Version	Date	Reason for Change	Person Responsible
1.0	17/01/2019	Initial Version	Madalyn Hardaker

Contents

Contents	2
Glossary	2
1. Objective	2
2. Scope	3
3. Background.....	3
4. Roles and Responsibilities	4
5. Procedures	4
5.1 Preparing files for digital archive.....	4
5.2 Transferring files.....	5
5.3 Accessing files.....	6
5.4 Deleting files from the archive	6
6. Related Documents	6
Appendix 1: Fees	7

Glossary

Payload	The data being archived. The payload is stored in files.
Safeguarding Software	The application used to collect payload files, encrypt them and place them into packages for transfer to the storage locations.
Web UI	The Web User Interface for the Safeguarding Software, accessed in the end-user's browser.
Escrow Tape	Arkivum sends a copy to escrow when there is approximately 2TB of content, enough to fill a tape. The tape is then sent to an offsite escrow facility.
Datapool	Allows a defined group of payload files to be co-located on a set of archive tapes.
Safeguarding Software	The application used to transfer files from the local network to Arkivum

1. Objective

To ensure that research project records are secured for long-term storage at the end of each project to comply with the UK Policy Framework for Health and Social Care Research, the Data Protection Act 2018 and Barts Health Trust and Queen Mary University of London Policy on the Retention and Disposal of Records (based on Department of Health recommendations on records retention).

To aid compliance with Reg 31A of SI 2006/1928 and the Freedom of Information Act 2000 by allowing records and information to be retrieved when needed.

To ensure that research records are managed in the most secure and cost effective way

To retain accessibility and compatibility after original systems and/or software are made obsolete or removed from routine use.

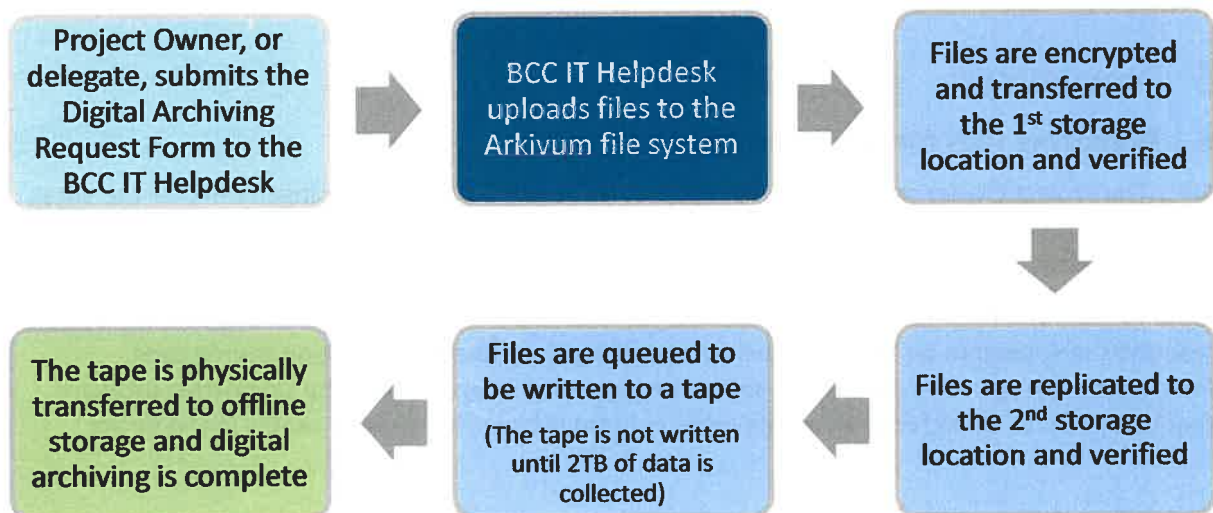
2. Scope

This policy is applicable to all staff who wish to use the digital preservation services offered through the Barts Cancer Centre as well as those who will be responsible for aspects of the internal process, such as the BCC IT helpdesk and the digital archivist.

3. Background

Digital archiving solutions, such as Arkivum, facilitate the long-term secure storage and preservation of critical information. The digital platform ensures the files are accessible when required and that integrity and compatibility can be assured for the full duration of retention. All files and metadata are encrypted and all files that are transferred go through a checksum process to ensure integrity for retrieval purposes. Arkivum guarantees that the files are 100% the same because of the checksum process.

FIGURE 1 SUMMARY OF THE ARCHIVING PROCESS



4. Roles and Responsibilities

Arkivum	Arkivum is a third party vendor selected to provide digital archiving services and will undertake activities as outlined in relevant contracts and agreements. Their primary role is ingesting and maintaining archived documents under the direction of BCC, ensuring content security, integrity and usability.
BCC IT Helpdesk	BCC IT helpdesk facilitates the deletion of files when necessary and conducts three yearly fidelity checks.
Digital Archivist	The Digital Archivist is a member of the BCC IT team, within QMUL, who is the named individual responsible for digital archiving. This individual processes the Digital Archiving Request Forms, facilitates technical support for the compilation of necessary files, and transfers the data from the local environment to the Arkivum Gateway Appliance, ensuring that access is restricted to only authorised individuals.
Project Owner	The Project Owner is generally the Principal Investigator or otherwise the senior person responsible for the project. The project owner may delegate archiving related tasks but is ultimately responsible for ensuring that the content to be archived is original, authentic, complete and accurate; and that relevant guidance has been reviewed and any necessary approvals are obtained including documented approval from the sponsor.
Project Sponsor	Recording the ownership of the archived content including any transfer of responsibility; informing the institution as to when the files no longer need to be retained and when trial records can be destroyed.

5. Procedures

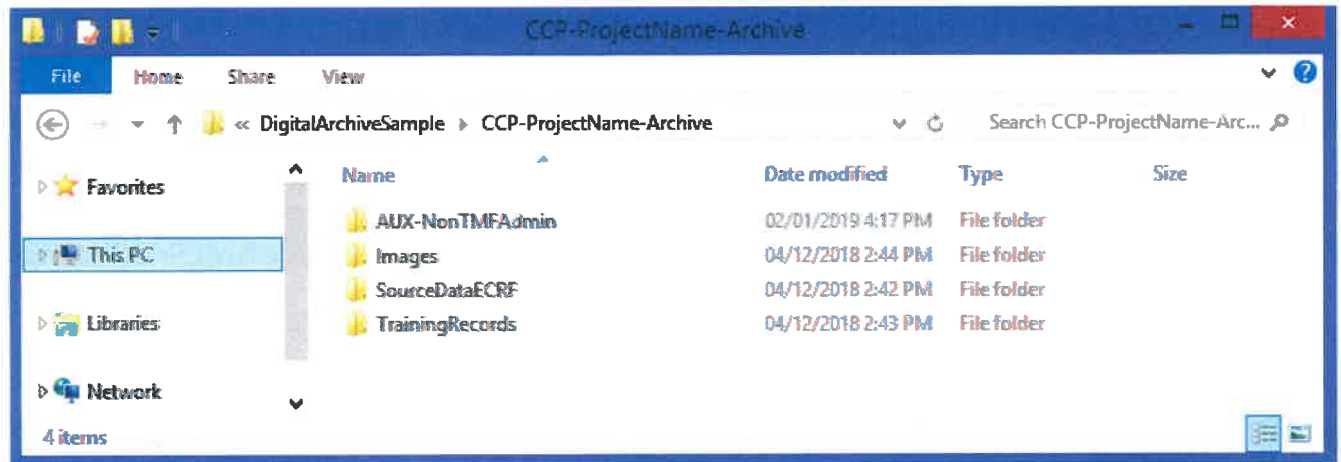
5.1 Preparing files for digital archive

5.1.1 The Project Owner should make the sponsor aware of the storage arrangements for the data, such as the list of contents to be archived and their retention times and these arrangements should be formally agreed and documented.

5.1.2 To maintain future access to certain records and data in their original formats, certain central resources may need to be archived, such as operating systems and application source code. Programmers assigned to the project should be consulted when defining the supportive software that may need to be archived alongside other content in order to render it usable in the future.

5.1.3 The Project Owner or delegate should move a copy of all files to be archived to a new folder named according to the following convention: “Centre-ProjectName-Archive” (NB: no spaces). Files should be logically organised and labelled, such as in the example below.

If files are to be digitally archived but are not subject to any future inspection, indicate “AUX” for auxiliary as a prefix.



5.1.4 The Project Owner or a delegate should complete a [digital archiving request form](#), which will document the types of data in the folder to be archived and highlights items for which technical support for archiving (e.g. operating systems) may be required.

5.1.5 Arkivum can accept all known file formats and will ensure their readability for decades to come, but certain formats may be favourable to others. For example, PDF/A is a standardised version of PDF specifically for archiving. The Digital Archivist may run the files through a program called DROID (Digital Record Object IDentification) to identify what file formats are currently in use and then recommend converting files into preferred file types, which can be done in bulk, when appropriate.

5.2 Transferring files

5.2.1 The digital archivist will review the digital archiving request form and the “Centre-ProjectName-Archive” directory and confirm any additional required information.

5.2.2 The Digital Archivist will transfer these files to Arkivum.

5.2.3 The files will undergo a checksum process upon upload to ensure data integrity during transit. The meta-data is encrypted, the original directory structure is obfuscated, all files are encrypted and a further hash is generated before being placed into an archive package to be transferred to the first storage location.

5.2.4 Each time an archive package is transferred the package will be verified by a checksum process to ensure the files are exactly the same.

5.2.5 The original files and archive package will be retained until the digital archive is moved to tape. The digital archivist will alert the project owner when the transfer to tape has been confirmed and then the archive package can be deleted.

5.3 Accessing files

5.3.1 Data must only be accessed from the archive with the approval of the investigator or institution. The sponsor should not have uncontrolled access to the files

5.3.2 If access to the archived data is required, the project owner or delegate should complete the Digital Archive Retrieval Request form.

5.3.3 All transfers of content to and from the digital archive are recorded.

5.4 Deleting files from the archive

5.4.1 When required, at the end of the defined retention period, submit a Digital Archive Deletion Request Form to have the content permanently removed from the digital archive.

6. Related Documents

BCC Digital Archiving Policy

DG-16 Disposal of Information Policy

JRMO SOP for Transferring Research Project Records to Corporate Records Management (known as Archiving)

QMUL Records Retention Policy

DigitalPreservationCoalition (2019). *Home - Digital Preservation Handbook*. [online] dpconline.org. Available at: <https://www.dpconline.org/handbook> [Accessed 9 Jan. 2019].

Nationalarchives.gov.uk. (2017). *DROID: user guide*. [online] Available at: <http://www.nationalarchives.gov.uk/documents/information-management/droid-user-guide.pdf> [Accessed 9 Jan. 2019].

Appendix 1: Fees

Arkivum Costs

Annual Maintenance Fee	£800
Additional fee per Terabyte	£360

BCC Cost Recovery Strategy

The ongoing maintenance fee of £800 per annum will be divided across the departments using the service. For example, if 4 departments are using the service then each department will pay a base fee of £200 each year.

The additional fee per Terabyte will be calculated according to the storage requirements of each department, rounded to the nearest 500GB (0.5TB).

For example, if Department A is storing 1.5TB of data and is one of four departments using the service, the annual fee will be £740

How much is a Terabyte?

A Terabyte (TB) is 1,000,000 Megabytes (MB) or 1,000 Gigabytes (GB). For reference, this is approximately 75 million pages of text, 17,000 hours of music, or 620,000 photos.

