DiAGRAM

Digital Archiving Graphical Risk Assessment Model

THE NATIONAL ARCHIVES

QUESTION 1: DIGITAL OBJECT

Definition: The proportion of your archive made up of born-digital, digitised and surrogate files.
Explanation: Different types of digital material hold different risks, for example some file formats may be easier to preserve than others. The amount of metadata held about the material and their conditions of use will also differ. All these aspects contribute to their digital preservation risk. If your archive does not distinguish between surrogates and digitised records put the percentage for surrogates only.

What proportion of your digital archive are the following?

- Born Digital - Records were created in a digital format.
- Digitised - Records have been created as a result of converting analogue originals, but you do not hold those originals.
- Surrogate - Digital images have been created as a result of converting analogue originals, and you also hold the originals.

QUESTION 2: STORAGE MEDIUM

Definition: The type of media on which your digital material is stored such as USB hard drives, CDs or the Cloud.
Explanation: Storage medium makes a big difference to the longevity of your digital material. Some are designed to be very robust (some degree of error detection and correction built in or less susceptible to many errors by design) and others are not intended for long-term storage.
What proportion of your records are stored on the following media types?

- A Less stable - Expected lifespan below 10 years or unknown, highly susceptible to physical damage, requires specific environmental conditions and very sensitive to changes, does not support error-detection methods, supporting technology is novel, proprietary and limited. Examples include USB flash drives (memory sticks), floppy disks, SD drives and CD-R discs.

- B More stable - A proven lifespan of at least 10 years, low susceptibility to physical damage, tolerant of a wide range of environmental conditions without data loss, supports robust error-detection methods, supporting technology is well established and widely available. Examples include LTO tapes, Blu-ray discs, enterprise/corporate managed hard drives and CD-ROM discs.

- C Outsourced Data Storage - An external company is responsible for our digital storage. Examples include Amazon Simple Storage Service, Microsoft Azure Archive Storage and Google Cloud Storage.

QUESTION 3.1: REP AND REFRESH

Definition: Your archive's policies on making copies and regularly moving digital material on to newer versions of the storage media.

Explanation: It is good practice to keep more than one copy of your digital material. If you do not do this for any material answer 0%.

What percentage of your material do you have at least one additional copy of?

QUESTION 3.2: REP AND REFRESH

Definition: Your archive's policies on making copies and regularly moving digital material on to newer versions of the storage media.

Explanation: As well as keeping copies of your digital material, it is good practice to refresh your storage media (i.e. move the material on to newer versions of the LTO tapes or hard drives for example). This reduces the risk that the material is not corrupted as storage media ages and requires replacement.

For those files with an additional copy, do you ensure you always have at least 2 independent copies?

QUESTION 4.1: OP ENVIRONMENT

Definition: Your archive's policy on the storage location of your digital material.

Explanation: It is good practice to keep a copy of your digital material offsite in case of a disaster at your primary location.

What percentage of your digital material has a copy kept offsite?

QUESTION 4.2: OP ENVIRONMENT

Definition: Your archive's policy on the storage location of your digital material.

Explanation: In the current version of the tool, the operating environment variable is only affected by the physical disaster of flood (not for example by fire or earthquake) so we are only asking if you have protection in place for flood. Other types of physical disasters may be added to later versions of the tool.

If all of your digital material is in one location, is there adequate protection against damage from a flood?

- Yes
- No
- Not Applicable - we have copies offsite

QUESTION 5: PHYSICAL DISASTER

Definition: The risk of a flood at your archive's primary storage location.

Explanation: Where the risk level varies for different risk types (i.e. flood risks from rivers or the sea, flood risks from surface water, flood risks from reservoirs), please answer based on the highest result. Note that for this first version of DiAGRAM, flood is the only physical disaster included, as this is the most likely physical disaster in the UK.

Based on the Government's long term flood risk assessment, how likely is it that your safest digital storage location will experience a flood? Click here to check your flood risk: https://flood-warning-information.service.gov.uk/long-term-flood-risk/postcode

- Very Low
- Low
- Medium
- High

QUESTION 6: CHECKSUM

Definition: A unique numerical signature dreived from a file that can be used to compare copies Definition from the DPC handbook. A checksum is needed to ensure integrity of the digital object.

Explanation: A checksum is needed to ensure the integrity of the digital object. Some depositors are unable to include checksums for the material they deposit in the archive. In these cases, archivists may decide to generate checksums for the material when they receive it to enable them to check it has not changed while in their custody. If your digital material does not have checksums or the checksums are generated after it is accessioned, choose 100% for C.

For what proportion of files do you have a checksum from following sources?

- TRUE - The depositor
- Archivist-generated - The archivist, generated on receipt of the record but prior to accessioning
- FALSE - You don’t have checksums at all, or they were generated sometime after initial receipt

QUESTION 7.1: SYSTEM SECURITY

Definition: A secure system can protect data from deletion or modification from any unauthorized party, and it ensures that when an authorized person makes a change that should not have been made, the damage can be reversed. Definition from Forcepoint, https://www.forcepoint.com/cyber-edu/cia-triad.

Explanation: Many archives do not have direct control of the security of their archival systems. If this is your situation, you can answer for the security managed at a corporate level.

Does your organisation hold a recognised security accreditation such as Cyber Essentials or ISO 27001 (or has it carried out equivalent assessments)?

- FALSE
- Cyber Essentials
- Cyber Essentials Plus
- ISO 27001

QUESTION 7.2: SYSTEM SECURITY

Definition: A secure system can protect data from deletion or modification from any unauthorized party, and it ensures that when an authorized person makes a change that should not have been made, the damage can be reversed. Definition from Forcepoint, https://www.forcepoint.com/cyber-edu/cia-triad.

Explanation: If you don't know what a penetration (or "Pen") test is then choose "No test".

Have your archival systems had a penetration test? If yes, are any issues outstanding?

- No test
- Critical issues outstanding
- Severe issues outstanding
- None, or only minor issues outstanding

QUESTION 7.3: SYSTEM SECURITY

Definition: A secure system can protect data from deletion or modification from any unauthorized party, and it ensures that when an authorized person makes a change that should not have been made, the damage can be reversed. Definition from Forcepoint, https://www.forcepoint.com/cyber-edu/cia-triad.

Explanation: By asking you to assess your skills against a standard, the tool will be able to give a more objective score for this question.

Referring to the NDSA Levels of Preservation, what level is your archive for the Control functional area?

- Not achieved
- Level 1
- Level 2
- Level 3
- Level 4

QUESTION 7.4: SYSTEM SECURITY

Definition: A secure system can protect data from deletion or modification from any unauthorized party, and it ensures that when an authorized person makes a change that should not have been made, the damage can be reversed. Definition from Forcepoint, https://www.forcepoint.com/cyber-edu/cia-triad.

Explanation: If the result is not recorded choose No as you cannot prove that the file was virus-free when it was received.

Is all of your digital material virus checked and the result recorded?

- FALSE
- TRUE

QUESTION 8.1: INFO MANAGEMENT

Definition: Internal systems and support for coherent information management and documentation of preservation actions. This is needed to ensure integrity and provenance of the digital object.

Explanation: For this tool, information management systems refers to the recording of digital preservation activities such as emulation, copying and fixity checking. It does not refer to your online catalogue or the broader information management capabilities of your wider organisation. By using existing standards we will get more consistent answers to the questions, so answers will be more comparable between different occasions that you use the tool

Referring to the NDSA Levels of Preservation, what level is your archive for the Metadata functional area?

- Not achieved
- Level 1
- Level 2
- Level 3
- Level 4

QUESTION 8.2: INFO MANAGEMENT

Definition: Internal systems and support for coherent information management and documentation of preservation actions. This is needed to ensure integrity and provenance of the digital object.

Explanation: For this tool, information management refers to the recording of digital preservation activities such as emulation, copying and fixity checking. It does not refer to the information management capabilities of your wider organisation such as a library, university or business.

Referring to the NDSA Levels of Preservation, what level is your archive for the Content functional area?

- Not achieved
- Level 1
- Level 2
- Level 3
- Level 4

QUESTION 8.3: INFO MANAGEMENT

Definition: Internal systems and support for coherent information management and documentation of preservation actions. This is needed to ensure integrity and provenance of the digital object.

Explanation: By asking you to assess your skills against a standard, the tool will be able to give a more objective score for this question.

Referring to DPC RAM, what level is your archive for Service capability, I - Content preservation
Referring to DPC RAM, what level is your archive for Service capability, J - Metadata management

- Minimal awareness
- Awareness
- Basic
- Managed
- Optimized

QUESTION 9: TECHNICAL SKILLS

Definition: Bespoke digital preservation skills such as awareness of technological trends, detailed knowledge of storage media, hardware and software, skills to perform file format migration, skills to find emulating software etc?
Explanation: By asking you to assess your skills against a standard, the tool will be able to give a more objective score for this question.

KIA 1.9 Apply appropriate technological solutions
KIA 1.12 Digital preservation standards
KIA 1.15 Information technology definitions and skills
KIA 1.16 Select and apply digital curation and preservation techniques
KIA 3.4 Continuously monitor and evaluate digital curation technologies
KIA 5.1 Data structures and types
KIA 5.2 File types, applications and systems
KIA 5.3 Database types and structures
KIA 5.4 Execute analysis of and forensic procedures in digital curation
PQ 3.9 Translate current digital curation knowledge into new services and tools

- None
- Basic
- Intermediate
- Advanced