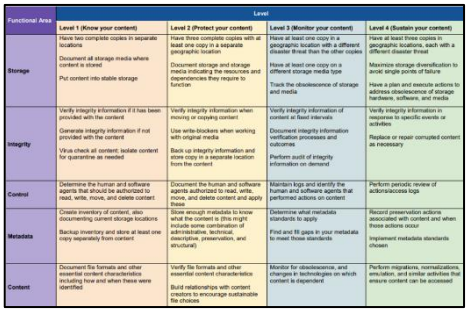
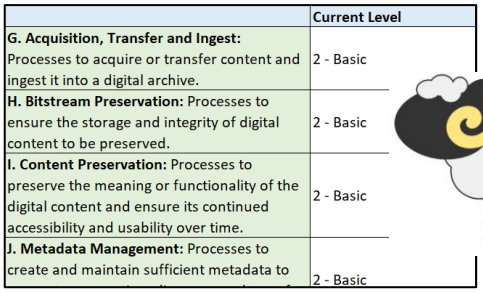
Kickstart Cymru Digital Preservation Bundles



Functional Area	Level			
	Level 1 (Know your content)	Level 2 (Protect your content)	Level 3 (Monitor your content)	Level 4 (Sustain your content)
Storage	Have two complete copies in separate locations Document all storage media where content is stored Put content into stable storage	Have three complete copies with at least one copy in a separate geographic location Document storage and storage media indicating the resources and dependencies they require to function	Have at least one copy in a geographic location with a different disaster threat than the other copies Have at least one copy on a different storage media type Track the obsolescence of storage and media	Have at least three copies in geographic locations, each with a different disaster threat Maximize storage diversification to avoid single points of failure Have a plan and execute actions to address obsolescence of storage hardware, software, and media
Integrity	Verify integrity information if it has been provided with the content Generate integrity information if not provided with the content Virus check all content; isolate content for quarantine as needed	Verify integrity information when moving or copying content Use write-blockers when working with original media Back up integrity information and store copy in a separate location from the content	Verify integrity information of content at fixed intervals Document integrity information verification processes and outcomes Perform audit of integrity information on demand	Verify integrity information in response to specific events or activities Replace or repair corrupted content as necessary
Control	Determine the human and software agents that should be authorized to read, write, move, and delete content	Document the human and software agents authorized to read, write, move, and delete content and apply these	Maintain logs and identify the human and software agents that performed actions on content	Perform periodic review of actions/access logs
Metadata	Create inventory of content, also documenting current storage locations Backup inventory and store at least one copy separately from content	Store enough metadata to know what the content is (this might include some combination of administrative, technical, descriptive, preservation, and structural)	Determine what metadata standards to apply Find and fill gaps in your metadata to meet those standards	Record preservation actions associated with content and when those actions occur Implement metadata standards chosen
Content	Document file formats and other essential content characteristics including how and when these were identified	Verify file formats and other essential content characteristics Build relationships with content creators to encourage sustainable file choices	Monitor for obsolescence, and changes in technologies on which content is dependent	Perform migrations, normalizations, emulation, and similar activities that ensure content can be accessed



	Current Level
G. Acquisition, Transfer and Ingest: Processes to acquire or transfer content and ingest it into a digital archive.	2 - Basic
H. Bitstream Preservation: Processes to ensure the storage and integrity of digital content to be preserved.	2 - Basic
I. Content Preservation: Processes to preserve the meaning or functionality of the digital content and ensure its continued accessibility and usability over time.	2 - Basic
J. Metadata Management: Processes to create and maintain sufficient metadata to	2 - Basic

Mapping to the NDSA Levels of Digital Preservation and Digital Preservation Coalition's Rapid Assessment Model

Version 1.3

October 2023

LLYFRGELL GENEDLAETHOL CYMRU
THE NATIONAL LIBRARY OF WALES

Contents

Overview

Scope

This guidance document provides an indicative mapping of the Kickstart Cymru digital preservation bundles and associated workflow to the self-assessment standards outlined in the NDSA Levels of Digital Preservation (NDSA Levels) and the Digital Preservation Coalition's Rapid Assessment Model (DPC RAM).[1] The guidance aims to provide recipient services of the Kickstart Cymru digital preservation bundles with an introduction to how the hardware and software provided can be used to support digital preservation self-assessment.

Indicative mapping against both the NDSA Levels and DPC RAM is based on the process outlined in the accompanying Kickstart Cymru bundles *Ingest Workflow Guidance*. The suggested levels of attainment are based largely on the functionality that is provided by the bundles and associated workflow; it is assumed that archive services will in practice be able demonstrate a higher level of maturity against both the NDSA Levels and DPC RAM when other policies and procedures are taken into account. As such, the guidance also highlights the areas of digital preservation which the workstations do not cover, and suggests how these gaps can be addressed through additional documentation and procedures.

Why is digital preservation assessment important?

Digital preservation assessments aim to benchmark organisational practices and policies against current professional standards and highlight areas for future improvement. Some assessment frameworks focus on organisational certification as a 'trustworthy' digital repository and these generally require both an advanced level of digital preservation and peer review or assessment by an external body; examples include the CoreTrustSeal certification framework and the ISO 16363 standard.[2] The current guidance document focuses on the NDSA Levels of Digital Preservation and DPC RAM as they both embrace the lighter-touch approach of self-assessment and explicitly include organisations that are just beginning to implement digital preservation. In addition, self-assessment using the NDSA Levels is specifically mandated in the Accredited Archive Service's (2022) Guidance under the criterion relating to Assessment and Management of Digital Preservation Risk.[3] Archive services undertaking accreditation going forwards will therefore be required to undertake self-assessment using the NDSA Levels.

In summary, digital preservation self-assessment can help you to:

- Align with current good practice
- Identify gaps and resources needed for the future

[1] National Digital Stewardship Alliance (NDSA) (2019)

[2] CoreTrustSeal (2022) *CoreTrustSeal Trustworthy Digital Repositories Requirements 2023-2025*. Available at: https://zenodo.org/record/7051012#.Y-OZkXbP2Uk; *ISO 16363:2012: Space data and information transfer systems — Audit and certification of trustworthy digital repositories*, available at: https://www.iso.org/standard/56510.html

[3] Accredited Archives Service (2022) *Archive Service Accreditation: Guidance for developing and completing an application*. Available at: https://cdn.nationalarchives.gov.uk/documents/archives/archive-service-accreditation-guidance-2022.pdf

- Set goals for improvement
- Understand how relationships with other organisational stakeholders are required to support digital preservation
- Establish what documentation you have to evidence existing procedures, and what else may be needed
- Apply for archive service accreditation (NDSA Levels of Digital Preservation)
- Act as a stepping stone towards the achievement of other forms of digital preservation certification, including external certification

Which self-assessment framework should you use?

The framework you select may depend on your reasons for conducting the exercise. If you are required to complete an assessment to support an application for archive service accreditation you will need to use the NDSA Levels. The NDSA levels may also be suited to a narrower assessment of your technical procedures and tools for carrying out specific digital preservation activities. DPC RAM, although relatively quick to complete (hence 'rapid assessment'!), also includes broader organisational resources, strategies and relationships and thus helps to highlight how your wider operational setting can support digital preservation in the long term.

NDSA Levels of Digital Preservation

Overview

The NDSA Levels of Digital Preservation is a resource which organisations can use to evaluate their digital preservation practices. It focuses on the technical aspects of digital preservation and is based around five functional areas comprising:

- Storage
- Integrity
- Control
- Metadata
- Content

In order to benchmark progression across the levels, organisations should assign themselves a score against each of the actions within each functional area. The actions are listed in order of advancement/complexity relative to the level described. For example, within the 'storage' functional area, an organisation is expected to have two complete copies of a record in separate locations for Level 1, whereas three copies in geographical locations with different disaster threats would be needed to achieve Level 4.

Assessment:

For each action an organisation should score themselves either 0, 1 or 2, where:

- 0 represents an action not yet started
- 1 represents a work in progress (i.e. an action which is not yet complete or not consistently implemented)
- 2 represents an action which is achieved (i.e. completed or consistently implemented if it is an ongoing action)

From these scores you can determine which overall level (Level 1 – 4) you achieve for each functional area.

Notes on self-assessment:

- In order to meet a particular level for a given functional area, you need to achieve **all** of the activities listed (i.e. score '2' for each of the activities relevant to that level).
- In order to move on to a higher level for a given functional area, you should already have achieved *all* of the actions listed for the previous Level.

It is important to remember that some of the activities described below cannot be considered 'achieved' or 'in progress' unless you are taking the necessary steps to ensure that they are actually being implemented in practice. Unlike some digital preservation software solutions which can automate processes such as integrity checking, the Kickstart Cymru workstation and the software supplied with it require manual operation to set processes such as file format validation running. This document therefore outlines what is achievable *in principle* through use of the workstation in accordance with the associated Kickstart Cymru workflow guidance; the *practice* of ensuring that the activities are undertaken rests with you(!).

The NDSA Levels matrix and a template which enables you to complete your own self-assessment, together with further resources, are available via:

https://ndsa.org/publications/levels-of-digital-preservation/

What this guidance tells you:

The guidance below maps the functionality provided by the Kickstart Cymru bundles and workflow to the NDSA Levels matrix. It also includes:

- An explanation as to why each level has been selected
- Possibilities for improvement
- Documentation that can serve as helpful evidence of your practices

Achieving Level 1 (Know Your Content) of the NDSA Levels

This document is intended to indicate how the Kickstart Cymru tools and associated guidance can help you to evidence progress against the NDSA Levels of Digital Preservation. It is therefore largely descriptive, rather than prescriptive, and maps *existing* tools and procedures to relevant activities listed in the NDSA levels.

However, this guidance also aims to:

- Show how you can demonstrate the achievement of at least Level 1 (Know Your Content), of the NDSA Levels using the Kickstart Cymru tools and workflow, as supplemented with some additional activities which are not incorporated in the resources you have been provided with in your bundle. This may require some further documentation and procedures to be developed but should be achievable with a minimal level of additional resources and planning.

- Show how you can progress beyond Level 1 to achieve a higher standard using the 'possible improvements' tips which follow each section. In most cases, the improvements outlined do not require significant investment or additional resources but rather focus on steps you can take to make your digital preservation practices as good as possible with the tools you have available.[4]

[4] The possible improvements for the 'Storage' functional area may be an exception as development of an additional storage facility/location may require additional investment if this not freely available, for example, on an accessible network drive or separately located PC/hard drive.

This tool can be used to assist you with determining which aspects of digital preservation you have strength in and which you may need to focus future efforts.									
	Level								
To use this tool, you will enter in a 0, 1, or 2 in each box next to a task. The conditional formatting will color code the tasks.	Functional Area	ENTER 0, 1, 2	Level 1 (Know your content)	ENTER 0, 1, 2	Level 2 (Protect your content)	ENTER 0, 1, 2	Level 3 (Monitor your content)	ENTER 0, 1, 2	Level 4 (Sustain your content)
2 = Achieved	Storage	1	Have two complete copies in separate locations	0	Have three complete copies with at least one copy in a separate geographic location	0	Have at least one copy in a geographic location with a different disaster threat than the other copies	0	Have at least three copies in geographic locations, each with a different disaster threat
		0	Document all storage media where content is stored	0	Document storage and storage media indicating the resources and dependencies they require to function	0	Have at least one copy on a different storage media type	0	Maximize storage diversification to avoid single points of failure
1 = Work in Progress		2	Put content into a stable storage			0	Track the obsolescence of storage and media	0	Have a plan and execute actions to address obsolescence of storage hardware, software, and media
0 = Not started	Integrity	2	Verify integrity information if it has been provided with the content	2	Verify integrity information when moving or copying content	2	Verify integrity information of content at fixed intervals	0	Verify integrity information in response to specific events or activities
		2	Generate integrity information if not provided with the content	2	Use write-blockers when working with original media	2	Document integrity information verification processes and outcomes	0	Replace or repair corrupted content as necessary
		2	Virus check all content; isolate content for quarantine as needed	2	Back up integrity information and store copy in a separate location from the content	1	Perform audit of integrity information on demand		
	Control	0	Determine the human and software agents that should be authorized to read, write, move, and delete content	0	Document the human and software agents authorized to read, write, move, and delete content and apply these	0	Maintain logs and identify the human and software agents that performed actions on content	0	Perform periodic review of actions/access logs
	Metadata	1	Create inventory of content, also documenting current storage locations	1	Store enough metadata to know what the content is (this might include some combination of administrative, technical, descriptive, preservation, and structural)	0	Determine what metadata standards to apply	0	Record preservation actions associated with content and when those actions occur
		2	Backup inventory and store at least one copy separately from content			0	Find and fill gaps in your metadata to meet those standards	0	Implement metadata standards chosen
	Content	2	Document file formats and other essential content characteristics including how and when these were identified	2	Verify file formats and other essential content characistics	0	Monitor for obsolescence, and changes in technologies on which content is dependent	0	Perform migrations, normalizations, emulation, and similar activities that ensure content can be accessed
				0	Build relationships with content creators to encourage sustainable file choices				

Figure 1: Indicative mapping of Kickstart Cymru ingest workstation and workflow to the NDSA Levels of Digital Preservation.

Note: This table displays what is currently achievable through only the use of the Kickstart Cymru ingest workstation and associated workflow. The guidance below provides information about how you can supplement the workflow to achieve at least Level 1 (Know your content) across all five functional areas.

Storage functional area

Overview

Which Storage activities the workstation can help you carry out:

Level 1: Know Your Content

Have two complete copies in separate locations	1 (In progress)
Document all storage media where content is stored	0 (Not started)
Put content into a stable storage:	2 (Achieved)

Not all of the actions required to meet Level 1 of the **Storage** functional area can be achieved using just the Kickstart Cymru workstation.

However, by storing your content in one additional location, and with the addition of documentation which identifies the storage media on which content is stored (e.g. via a Digital Asset Register), your archive service can achieve Level 1 overall.

Suggested Level Achievable for Storage: Level 1

Activities in detail

Have two complete copies in separate locations

Although the RAID External Hard Drive, as configured to RAID 1, allows for mirrored storage across two separate hard disks to enable recovery in the event of a single disk failure, this still only counts as one storage location as they are both part of the same hard drive. Similarly, if you retain a copy of the content on your workstation PC as well as on the external hard disk, the fact that they are not stored in separate physical locations means that this is not an ideal separation of storage.

The working definitions provided by the NDSA suggest that:

- Separate locations means 'two different places' and 'when beginning digital preservation this could simply mean one copy on a server and another on an external hard drive'; 'A desktop and a hard drive in the same workspace is not considered two separate locations'
- In relation to a 'separate geographic locations', this requires that '[i]deally, copies would be stored in a location with a **different disaster threat**' [*emphasis in original*][5]

This activity is therefore not met by virtue of the Kickstart Cymru workstation and associated workflow alone. In order to meet this criterion, your archive service should store a complete copy of the data separately from your RAID hard drive and ingest workstation PC. This could include a network-based drive to which you have access or storage on a PC or additional external hard drive which is not connected to the ingest workstation PC.

Document all storage media where content is stored

This activity is not carried out as part of the Kickstart Cymru ingest workflow, but could be satisfied by inclusion as part of your Digital Asset Register, or by noting the types of storage media used for digital preservation within supporting documentation such as a Digital Preservation Strategy or Collections Care Policy. Remember that where an additional storage medium is used for the purpose of having two copies in separate locations, you should also list the second storge medium in addition to the RAID hard drive.

Put content into stable storage

The working definitions provided by the NDSA suggest that storage means 'dedicated storage designated and managed specifically for digital preservation'. This means a dedicated location where you have decided your preserved digital content should sit, in contrast to records being scattered around in different shared drives, inboxes and other places. Therefore, if utilised consistently and specifically for the storage of your digital records, the RAID External Hard Drive will satisfy this condition.

[5] NDSA (2019) *Working Definitions for the Levels of Digital Preservation Version 2.0*. Available at: https://osf.io/rynmf

Possible Improvements

- Investment in an additional storage solution (e.g. cloud based, network based) which is not physically collocated with the ingest workstation.
- Investment in a digital preservation system that includes storage (e.g. Arkivum, Preservica)

- **Supporting policies and documentation**

- Digital asset register, or other records inventory, which provides evidence of storage media. This could be arranged by format, or at collection level, rather than item level **or** a digital preservation strategy or collections care policy which identifies specific storage media used for the purposes of digital preservation.
- Digital Preservation Policy which outlines that storage is separated across locations and media types to avoid single points of failure.

Integrity functional area

Overview

Integrity activities the workstation can help you carry out:	
Level 1: Know Your Content	
Verify integrity information if it has been provided with the content	2 (Achieved)
Generate integrity information if not provided with the content	2 (Achieved)
Virus check all content; isolate content for quarantine as needed	2 (Achieved)
Level 2: Protect Your Content	
Verify integrity information when moving or copying content	2 (Achieved)
Use write-blockers when working with original media	2 (Achieved)
Back up integrity information and store a copy in separate location from the content	2 (Achieved)
Level 3: Monitor Your Content	
Verify integrity information of content at fixed intervals	2 (Achieved)
Document integrity information verification processes and outcomes	2 (Achieved)
Perform audit of integrity information on demand	1 (In progress)

All of the actions required to meet Levels 1 (Know Your Content) and 2 (Protect Your Content) for the **Integrity** functional area are suggested to be achievable with the Kickstart Cymru workstation, if it is used in accordance with the associated workflow guidance.

The workstation should also enable you to perform some of the actions listed under Level 3 (Monitor Your Content) of the Integrity Functional Area, but not all.

It is important to remember that the workstation does not automate the procedures for verification of integrity information (unlike some digital preservation software suites) and, therefore, in order to achieve Level 2, it is necessary to ensure that you maintain robust integrity verification processes which need to be manually initiated.

Suggested Level Achievable for Integrity: Level 2

Activities in detail

Generating integrity information

The working definitions provided by the NDSA refer to integrity information as: 'Information about digital content that can be used to verify over time that the content is whole and unaltered through loss, tampering, or corruption. At the bit level, this may take the form of checksums'.[6]

Within the Kickstart Cymru workflow, Teracopy and Droid are both capable of generating checksums against which any checksums submitted by the depositor can be verified.

Verifying integrity information

Integrity information (checksums) can be checked at fixed intervals by following the guidance under 'Fixity Checking with Droid' in the Ingest Workflow Guidance, which suggests verifying them against the checksums generated on ingest at intervals of between 6 months and 1 year.

Teracopy and Droid are both used as part of the ingest process to produce integrity information, including checksums. The reports generated by these processes are also saved to provide evidence of the processes undertaken.

Using a USB write blocker

A USB write blocker has been supplied as part of the workstation for use in the transfer of digital records from external media.

Documenting and Backing Up Integrity Information

Integrity information (checksums) are generated through both Teracopy and Droid and copies of reports produced by each tool are stored as metadata both with the content and, as recommended in our workflow, as a copy to be stored separately from the content.

[6] NDSA (2019) Working Definitions for the Levels of Digital Preservation Version 2.0. Available at: https://osf.io/rynmf

Possible Improvements

- Planning for the **verification of integrity information in response to specific events** or activities is not currently included in the Kickstart Cymru workflow which focuses on integrity checking on ingest and at fixed intervals thereafter. The same integrity checking process could in theory be used to verify integrity information in response to specific events, such as perceived risks that might be caused by hardware or disk failure, or in response to a security threat or electrical fault, though this would currently need to be triggered manually. At a basic level, however, this could be achieved by checking content is unchanged after storage media migration (such as when you decide to copy the content to a second storage location).

- Initiation of a process to **review and audit integrity checking processes** and the records generated by them: This would ideally be performed by an external individual. However, in the first instance, you could have an archive staff member run an internal test of the systems. For example, having somebody add and then edit or delete a test file, then subsequently checking to see if the procedures pick up that there has been a change or deletion (i.e. by producing a different checksum).

Supporting policies and documentation

- Digital Preservation Policy which includes a *commitment* to generate, store and check integrity information.
- Digital Preservation Strategy or Collections Care Procedures which provide an overview of *how* integrity information is generated and stored using checksums.
- Droid output reports which are saved as metadata both with the content and separately from it.

Control functional area

Overview

The Kickstart Cymru bundles and accompanying workflow guidance do not cover the requirements for the **Control** functional area of the NDSA Levels. These determinations will be individual to each archive service depending on your staffing, access, and information security controls.

If you do, however, at least make a determination as to which people and software should be authorised to read, write, move and delete content, then this would be sufficient to achieve Level (Know Your Content), and therefore, to achieve Level 1 across overall.

Suggested Level Achievable for Control: Level 1

Possible improvements

Decide which staff members should be responsible for ingesting digital records via the ingest workstation, as well as who should be able to perform fixity checks and access the content for any other reason (e.g to arrange access). Ensure that only these staff members have access to the content and that responsibilities are written down in an information security policy or similar document.

Supporting policies and documentation

- Digital preservation policy which may outline, or refer to, access and information security controls for digital content.
- Information security policy which outlines responsibilities and access controls.

.

Metadata functional area

Overview

Metadata activities the workstation can help you carry out:

Level 1: Know Your Content

Create inventory of content, also documenting current storage locations	2 (Achieved*)
Backup inventory and store at least one copy separately from content	2 (Achieved)

Level 2: Protect Your Content

Store enough metadata to know what the content is	1 (In progress)

*This can be achieved if you take additional action to document current storage locations.

The software included in your Kickstart Cymru workstation, and Droid in particular, are capable of generating most of the metadata required to achieve Level 1 in relation to the **Metadata** functional area. However, a Digital Asset Register or Information Asset Register which details the storage locations for your content, at least at the collection level, will also be required.

Therefore, to achieve the suggested Level 1 (Know your content), some further action and documentation will be required in addition to steps outlined in the Kickstart Cymru workflow.

Suggested Level Achievable for Metadata: Level 1

Activities in detail

Create inventory of content and storage locations

The metadata generated by a Droid output report includes the majority of elements the NDSA lists in the definition of 'inventory', being:

'A broad description or listing of content to be preserved that may include, for example, file listing, file formats, size, media, and location of the content.'

However, a Droid report does not list the media on which, or the location where, the content is stored and this would therefore need to be separately recorded. As a first step, this could be achieved by ensuring that Droid output reports are supplemented by a Digital Asset Register which outlines storage locations at the collection level.

The Kickstart Cymru workflow also recommends storing a copy of any accession or transfer forms in the related content's metadata folder and it is important to remember this as a preliminary step.

Possible Improvements

You should select and consistently apply specific metadata standards, which may include PREMIS for digital records in particular. Specific metadata standards adhered to may be referenced in your digital preservation policy or metadata policy.

Beyond this, automatic recording and reporting of metadata in relation to preservation activities is likely to be possible only through the acquisition of a digital preservation software suite which should both perform preservation-related activities (e.g. normalization, creation of access copies, integrity checking) and produce a record of when these activities are undertaken.

Supporting policies and documentation

- Digital preservation policy which outlines, or links to, your service's metadata policy for digital content. This may include additional elements to the metadata associated with physical records
- Metadata policy for digital records. If not independent, this could be annexed, for example, as part of a wider cataloguing policy or cataloguing procedures document.
- Digital Asset Register which outlines storage locations and storage media (at collection level at least)

Content functional area

Overview

Control activities the workstation can help you carry out:	
Level 1: Know Your Content	
Document file formats and other essential content characteristics	2 (Achieved)
Level 2: Protect Your Content	
Verify file formats and other essential content characteristics	2 (Achieved)
Build relationships with content creators to encourage sustainable file choices	0 (Not started)

The tools supplied as part of your Kickstart Cymru workstation enable you to document the file formats and other essential content characteristics of your digital records, including through the use of Droid and JHove in particular. You can also carry out file format verification using the same tools.

For the **Content** functional area, you can therefore achieve Level 1 (Know your content), and one of the activities for Level 2 (file format verification) using the workstation and software supplied.

If your archive service is able to build relationships with content creators to encourage sustainable file choices then you would also be capable of showing that you reach Level 2 (Protect your content) overall.

Suggested Level Achievable for Content: Level 1

Activities in detail

Documenting and verifying file formats and other essential content characteristics

Droid is used to perform file format identification as part of the Kickstart Cumru ingest workflow, and a copy of the Droid report is recommended to be saved as metadata, both with the content and separately from it, to evidence how and when these formats were identified.

Jhove has also been supplied for use as part of the ingest workflow for format verification and the resulting profile is also recommended to be saved as a metadata report.

Possible Improvements

To achieve Level 2, you could initiative and maintain discussions with content creators surrounding file format sustainability; this may be more feasible for creators with whom your service has a long-standing relationship and those from whom you receive ongoing accruals (e.g. local authority bodies). However, documentation outlining preferred file formats such as depositor guidelines or a collections development policy accessible to your external depositors could also help to ensure you receive files in an optimal format for preservation from non-local authority bodies.

To achieve an even higher level (Level 3: Monitor your content) you could carry out ongoing reviews to assess risks caused by the potential obsolescence of formats, hardware, and software, including the tools and technologies on which the storage or access to the content may be dependent. This may include reviewing your Digital Asset Register at fixed intervals to identify records held on formats which may be at greater risk. The Digital Preservation Coalition's 'BitList' provides a list of 'digitally endangered species' in relation to a wide range of digital content.[7]

Supporting policies and documentation

- Droid and Jhove metadata output reports
- Depositor guidelines or collections development policy which outline preferred file formats
- Digital Preservation Policy which references, in outline, the actions and periodic reviews undertaken to monitor and mitigate against obsolescence (if applicable)
- Digital Asset Register which may be used to support format obsolescence monitoring (if applicable)

[7] DPC (2021) *The BitList 2021: The Global List of Digitally Endangered Species*. Available at: https://www.dpconline.org/docs/miscellaneous/advocacy/wdpd/2521-bitlist2021/file

DPC Rapid Assessment Model

Overview

The DPC's Rapid Assessment Model is designed to allow rapid benchmarking of your digital preservation policies and procedures against a range of organisational and service capabilities, and it is suitable for use by any organisation which preserves digital information over the long term. The DPC RAM maturity model enables you to rate your organisation against a series of levels ranging from beginner ('minimal awareness') to advanced ('optimized') level.

The guidance below outlines how the functionality provided by the Kickstart Cymru bundles, and associated workflow, may be mapped to criteria outlined in the service capabilities section of DPC RAM. It also includes:

- An explanation of which criteria for that service capability are/are not met
- Guidance on how to reach the next level of the maturity model (e.g. progression from 'basic' to 'managed' level)

Resources

The full DPC RAM worksheet is available for download via:

https://www.dpconline.org/digipres/implement-digipres/dpc-ram

Additional guidance on DPC RAM is outlined in the bibliography at the end of this document. The DPC's website also has a wide range of resources, many of which are available to member and non-member organisations alike: https://www.dpconline.org/digipres/implement-digipres/dpc-ram

The DPC have also made available a new series of webpages on 'Levelling Up' with DPC RAM which are also accessible to non-members: https://www.dpconline.org/digipres/implement-digipres/dpc-ram/level-up [8]

DPC RAM organisational capabilities

The DPC RAM self-assessment frameworks is split into two key areas titled organisational capabilities and service capabilities. Unlike the NDSA Levels, which is focussed on specific digital preservation activities, DPC RAM also provides benchmarking criteria to establish how well your wider organisational governance, strategy, legislative compliance, IT capabilities, progress monitoring, and community engagement support digital preservation As these areas are beyond the scope of the ingest workstations and their associated workflows, this part of the self-assessment framework has been excluded from the mapping levels below. However, organisational capabilities are considered as important as service capabilities for understanding how well your wider operating context supports

[8] Available to non-members as of August 2023.

digital preservation and it is recommended that services undertaking self-assessment using DPC RAM complete their self-assessments using both parts of the model.

Mapping to service capabilities of DPC RAM

The table below provides an indication of how the Kickstart Cymru bundles and workflow can be mapped to the services capabilities in DPC RAM. When carrying out your own self-assessment, you should also provide a target level which reflects your organisational aims for improvement of your digital preservation policies and procedures in the future, as well as the steps you will need to take to reach your target level. The guidance on 'how to reach the next level' below may be helpful in filling this part of the matrix out.

SERVICE CAPABILITIES		
	Current Level	**Why did you select this level?**
G. Acquisition, Transfer and Ingest: Processes to acquire or transfer content and ingest it into a digital archive.	1 - Awareness/ 2 - Basic	A documented process for ingest exists; Documentation and metadata is sometimes received or captured as part of the acquisition or transfer process; a working area (physical or virtual) is available for pre-ingest and ingest activities (for example to carry out virus checking and file identification).
H. Bitstream Preservation: Processes to ensure the storage and integrity of digital content to be preserved.	1 - Awareness/ 2 - Basic	Dedicated storage is available to meet current preservation needs; staff know where content is stored; checksums are generated for all content
I. Content Preservation: Processes to preserve the meaning or functionality of the digital content and ensure its continued accessibility and usability over time.	1 - Awareness/ 2 - Basic	File formats are identified; Content is characterized and assessed for preservation and quality issues such as encrypted, broken, or incomplete content and invalid files.
J. Metadata Management: Processes to create and maintain sufficient metadata to support preservation, discovery and use of preserved digital content.	1 - Awareness/ 2 - Basic	Metadata and documentation acquired with content is retained and preserved; basic preservation metadata is captured at item level.
K. Discovery and Access: Processes to enable discovery of digital content and provide access for users.	1 - Awareness/	The ingest workstation and associated workflow currently provides the tools for ingest and preservation of digital records but they do not manage access to the records and a separate workflow for access will need to be devleoped by archive services according to their institutional requirements.

Figure 2: Indicative mapping to DPC Rapid Assessment Model Service Capabilities.

Note: The levels indicated in this document indicate a range between 'Awareness' and 'Basic' for a number of service capabilities, to demonstrate that the Awareness level is achieved, but not all of the Basic criteria are met. In practice, when you are completing your own DPC RAM self-assessment, you should only select **one** level based on your service's actual practices and procedures, and you should only select a specific level (e.g. Basic) if all of the activities for that level are being carried out.

G – Acquisition, Transfer and Ingest

Level Suggested: **1- Awareness to 2 – Basic**

Basic criteria which are met:

- **A documented process for ingest exists**: The ingest workflow guidance represents a documented process for ingest, though each service may wish to adapt this documentation to suit any alternative processes and procedures they use.
- **Documentation and metadata is sometimes received or captured as part of the acquisition or transfer process:** According to the Kickstart Cymru ingest workflow, reports produced by tools such as Droid and Jhove are saved in a metadata folder, copies of which are recommended to be retained both with the content and separately from it.
- **A working area for pre-ingest and ingest activities (for example to carry out virus checking and file identification**): The Kickstart Cymru bundles have been installed in each archive service in a dedicated working space which houses all of the hardware and software required to carry out ingest activities including virus checking and file identification.

Possible improvements for Basic criteria not met

- **Basic guidance for donors, depositors and record creators is available where appropriate:** It is assumed that each archive service will already have guidance and policies for donors and depositors in place, but these are not covered by the Kickstart Cymru workflow.
- **A documented process exists for selecting and capturing digital content where appropriate (e.g. for web archives, email archives, digitized content, records within an EDRMS):** The workstations and associated workflow do not cover the stage of selecting digital content and separate documentation for how this will be managed could be developed (e.g. as part of a digital preservation policy/strategy)
- **Some content is appraised as part of a manual process in line with relevant policies:** Some appraisal actions including file format identification and validation, and checking for duplication, are including in the recommended workflow, but a separate policy on digital content appraisal could be developed or incorporated as part of your existing collections development policy.

How to reach the next level (3 – Managed)

Some of the criteria for 'Managed' level may already be met by the workstation and associated workflows, including:

- **Appraisal is a standard part of the ingest workflow**: our Kickstart Cymru ingest workflow recommends some appraisal steps using the output reports from Droid and Jhove.
- **Successful transfer of content is verified by integrity checking:** This may be achieved through verification of checksums provided by the depositor, by comparison against the checksum generated by Teracopy/Droid; though, in practice many services have reported they are unlikely to receive checksums with digital records at present.

- **Parts of the ingest process are automated:** AVG antivirus should automatically scan content for viruses; Droid and Jhove do automate the processes of file format identification and verification, but this process must be manually initiated so this is not currently a truly 'automated' process.

Services should also need to ensure that they manage their relationships with donors and records creators which, in addition to the above-mentioned guidance documentation, may include ongoing communication and support.

H – Bitstream Preservation

Level Suggested: **1 – Awareness to 2 – Basic**

Basic criteria which are met:

- **Dedicated storage is available to meet the current preservation needs:** The RAID external hard drive supplied as part of the workstation provides 6TB of mirrored storage which is specifically designated for the storage of digital records.
- **Staff know where content is stored:** Archive service staff have been trained on the transfer and storage of digital records to the external hard drive as part of the installation process, and through the associated workflow guidance, though separate documentation to indicate the location of specific digital collections/records would be helpful to evidence this. (E.g. through an inventory or digital asset register).
- **Checksums are generated for all content:** Teracopy and Droid are used to generate checksums for all content that is ingested.

Possible improvements for Basic criteria not met

- **Replication is based on simple backup regimes:** Although content on the external hard drive is mirrored across two drives, and allows for recovery in the event of single disk failure, there is no current process for backing up the content on a separate storage system. A separate backup via storage of a complete copy on a separate medium, in a separate location, would be ideal.
- **There is an understanding of which staff members should be authorised to access the content:** It is assumed that this criteria is likely to already be met in the case of each archive service, but this would need to be separately evidenced/documented through, for example, access and information security policies.

How to reach the next level (3 – Managed)

Some of the criteria for 'Managed' level may already be met by the workstation and associated workflows, including:

- Content is already integrity checked, but may need to be stored/replicated in an additional storage location, as it is currently only suggested to be stored on the RAID external hard drive.

In addition, services would need to:

- Make decisions on the frequency of integrity checking and number of copies held which take into consideration the risks, value of the content and the associated financial and environmental costs. Such discussions, and documented outcomes, may include consideration of whether to invest in software capable of automating integrity checking at fixed intervals (e.g. a standalone tool such as AVG Fixity Pro, or a software suite such as Preservica or Archivematica).
- Repair content failing integrity checks: The present workflow identifies files which fail integrity checks, but does not suggest a tool or process for repairing them
- Enforce and document staff authorisations to access content
- Carry out tests to verify the effectiveness of backups, replication and integrity checking.

I – Content Preservation

Level Suggested: **1 - Awareness to 2 – Basic**

Basic criteria which are met:

- **File formats are identified:** Droid is used to identify file formats in accordance with the PRONOM database and the resulting report is stored as metadata.
- **Content is characterised and assessed for preservation and quality issues such as encrypted, broken, or incomplete content and invalid files:** Jhove is used to perform file format validation and to confirm the status of each file which is transferred. DROID can also help meet this requirement as it will identify content that is encrypted and may flag an error if it finds broken/corrupted content.

Possible improvements for Basic criteria not met

- **There is a basic understanding of current and future users and use cases for the content:** Services will have a strong understanding of the prospective users and uses for the content ingested, but this may be supported through additional documentation such as a digital preservation strategy which outlines the stakeholders involved in digital preservation.

How to reach the next level (3 – Managed)

To reach 'Managed' level services will need to:

- Carry out technology watch activities and identify 'at risk' content. This may include reviewing, and documenting, your Digital Asset Register at fixed intervals to identify records held on formats which may be at greater risk. The Digital Preservation Coalition's 'BitList' provides a list of 'digitally endangered species' in relation to a wide range of digital content.[9]
- Document technical dependencies, which may include the hardware and software required to sustainably store, and access, digital content in the future.
- Ensure the preservation and quality of content through actions such as migration, emulation or modification of creation or capture workflows: this may require consideration of how proprietary file formats can be accessed in the future and whether they require maintenance of specific software or migration to an open-source format;
- Record details (e.g. access logs) of changes to digital content

[9] DPC (2021) *The BitList 2021: The Global List of Digitally Endangered Species*. Available at: https://www.dpconline.org/docs/miscellaneous/advocacy/wdpd/2521-bitlist2021/file

J – Metadata Management

Level Suggested: **1 - Awareness to 2 – Basic**

Basic criteria which are met:

- **Metadata and documentation acquired with content is retained and preserved**: The accompanying ingest workflow guidance suggests that such metadata and documentation are saved in the 'metadata' folder created on ingest of any digital records and is saved both with the content and separately from it.
- **Basic preservation metadata is captured at item level**: Some, but not all, elements of what may be considered basic preservation metadata are created and captured for item-level objects including fixity information (checksums) and file format, and version information which support access to the files in the future. However, other data which might constitute basic preservation metadata, including information about rights/permissions is not automatically captured through the workflow.

Possible improvements for Basic criteria not met

- **Content is described at collection level in a digital asset register:** While it is assumed that archive services are likely to retain digital asset registers which describe content at the collection level, this is not covered by the Kickstart Cymru ingest workflow.
- **An appropriate minimum descriptive metadata requirement exists:** The Kickstart Cymru guidance does not cover descriptive metadata and is agnostic as to the standard that may be adopted by each service.

How to reach the next level (3 – Managed)

To reach 'Managed' level services would need to:

- Identify appropriate metadata standards for your digital content. Some of these may be similar to your physical content, in relation to description for example, but some preservation metadata in particular may be specific to digital records This may include, for example, reference to the PREMIS (Preservation Metadata: Implementation Strategies) standard which supports the long-term preservation of digital materials.[10]
- Maintain guidance and use controlled vocabularies to ensure consistency of metadata entry
- Assign persistent unique identifiers (PIDs) to manage digital content[11]
- Maintain the structural relationships between data and metadata that form a particular digital object

[10] Library of Congress (2023) *PREMIS Data Dictionary for Preservation Metadata, Version 3.0*. Available at: http://www.loc.gov/standards/premis/v3/index.html

[11] See: Digital Preservation Coalition, *Digital Preservation Handbook: Persistent Identifiers*. Available at: https://www.dpconline.org/handbook/technical-solutions-and-tools/persistent-identifiers

K– Discovery and Access

Level Suggested: **1 –Awareness**

Awareness criteria which are met:

[None] The ingest workstation and associated workflow does not cover discovery and access to the content

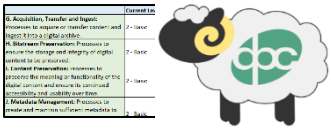
Possible improvements for Awareness criteria not met

- **There is a basic understanding of current and future users and use cases for the content:** Archive services will have a strong understanding of your prospective users and uses for the content ingested, but this may be supported through additional documentation such as a digital preservation policy which outlines the stakeholders involved in digital preservation.

How to reach the next level (2 – Basic)

To reach 'Basic' level services would need to ensure that:

- Basic resource discovery exists for some digital content, which may include the listing of digital content (e.g at collection level) in your archive catalogue
- Users should be able to access digital content and metadata, either remotely or on-site
- Users' access to digital content is recorded through, for example, an access log
- Information on the accessibility of digital content is provided to users, which could take the form of an access policy or access and use information available through your website



DPC RAM

Service capabilities

NDSA Levels

Functional areas

	Switch on	
(J) Metadata management (K) Discovery and access	Create working folders	Metadata
(G) Acquisition, Transfer and Ingest	Update virus definitions	Integrity
(H) Bitstream preservation	Connect via USB write blocker	Integrity
(G) Acquisition, Transfer and Ingest	Virus scan	Integrity
(H) Bitstream preservation	Copy content	Integrity
(I) Content preservation	File format identification	Integrity Content
(I) Content preservation	File format validation	Content
(H) Bitstream preservation	Transfer to hard drive	Storage
(J) Metadata management (K) Discovery and access	Manage desktop files	Metadata

Figure 3: Mapping Kickstart Cymru Workflow Steps to DPC Ram Service Capabilities and NDSA Level Functional Areas

Bibliography and resources

Accredited Archives Service (2022) *Archive Service Accreditation: Guidance for developing and completing an application*. Available at: https://cdn.nationalarchives.gov.uk/documents/archives/archive-service-accreditation-guidance-2022.pdf

NDSA (2020) NDSA Levels of Digital Preservation FAQ: https://docs.google.com/document/d/1jxj7hlV9LNO0grucz88CSyk_WHrc_uAQB-EYJjJvexY/edit#heading=h.mw3frum1afrw

NDSA (2019) Using the Levels of Digital Preservation: an Overview for V2.0: https://osf.io/vnc32

NDSA (2019) Working Definitions for the Levels of Digital Preservation Version 2.0: https://osf.io/rynmf

Digital Preservation Coalition, Rapid Assessment Model, available at: https://dpconline.org/our-work/dpc-ram

Digital Preservation Coalition, Digital Preservation Handbook, available at: http://dpconline.org/handbook

Digital Preservation Coalition, *Assessing readiness for digital preservation*, available at: https://dpconline.org/docs/miscellaneous/training/1677-assessing-readiness-gettingstarted/file

Digital Preservation Coalition, *How to Level Up With DPC RAM* (Updated 2023) available at:

https://www.dpconline.org/digipres/implement-digipres/dpc-ram/level-up