

TAHO

Tasmanian Archive + Heritage Office

State Records Guideline No 19

Digital Preservation Formats

Table of Contents

1	Purpose.....	4
1.1	Scope.....	4
1.2	Authority.....	4
2	Preserving Digital Records.....	5
3	Characteristics of file formats for preservation.....	6
3.1	Open and accessible formats.....	6
3.2	Ubiquity.....	6
3.3	Metadata support.....	6
3.4	Feature set.....	6
3.5	Interoperability.....	6
3.6	Viability.....	7
4	Destroy digital records when appropriate.....	7
5	Monitor for Obsolescence.....	7
6	Back-up Systems.....	8
6.1	A backup system is not a recordkeeping system.....	8
7	Definitions.....	8
	Further Advice.....	10
	Acknowledgements.....	10
	Appendix 1 – Table of recommended file formats.....	12
	Appendix 2 – Guidelines for creating word documents, spreadsheets and emails.....	14

Information Security Classification

This document has been security classified using the Tasmanian Government Information Security classification standard as PUBLIC and will be managed according to the requirements of the Tasmanian Government Information Security Policy.

Document Development History

Build Status

Version	Date	Author	Reason	Sections
2.0	04-06-2015	Christine Woods	Template	All
1.0	23-04-2012	Ross Latham	Initial Release	All

Amendments in this Release

Section Title	Section Number	Amendment Summary
All	All	Document imported into new template

Issued: April 2012

Ross Latham
State Archivist

UNDER REVIEW

I Purpose

This guideline identifies digital file formats that the Tasmanian Archive and Heritage Office recommend as being suitable for permanent and long term preservation. It also includes guidance on creating, storing and managing digital records.

Preserving digital information is complicated and involves the active commitment of organisations, the development of appropriate policies and plans, and the implementation of sound practices. It requires all organisations with an interest in preserving digital information to share expertise, advice and best practices.¹

The *Archives Act 1983* stipulates that 'the relevant authority is to cause all [State] records to be preserved and accessible until they are dealt with in accordance with this Act'.²

As file formats encode information into a form which can only be processed and rendered comprehensible by specific combinations of hardware and software, the accessibility of that information is highly vulnerable in the rapidly evolving technological environment.³

The approach to digital preservation must ensure:

- 1) The content of the record remains unchanged despite preservation processes;
- 2) An appropriate audit trail of preservation activities is recorded; and
- 3) Compliance with digital recordkeeping standards of authenticity.

The following legislation must be considered in the selection of an approach to digital records preservation. The *Archives Act 1983* relating to the making and keeping of records by agencies and the management of State archives; the *Evidence Act 2001* for the provision of copies of records from electronic systems; and the *Electronic Transactions Act 2000* for the equivalence of records of business conducted electronically or in physical form.⁴

I.1 Scope

The focus of this guideline is on media-independent content, ie. digital content that is managed as file types and not inextricably linked to a physical storage medium (in contrast to videotape which is dependent both on the physical carrier and the playback equipment). It does not address recommendations for physical preservation media.

I.2 Authority

This guideline is issued under the provisions of Section 10A of the *Archives Act 1983*. Guidelines issued by the State Archivist under this Section set standards, policy, and procedures relating to the making and keeping of

¹ Library and Archives Canada, Home>About Us>Digital Initiatives>Digital Policies, Guidelines and Tools>Local Digital Format Registry (LDFR) File Format Guidelines for Preservation and Long-term Access Version 1.0 at LAC 22-10-2010 <http://www.collectionscanada.gc.ca/digital-initiatives/012018-2220.01-e.html#n> Accessed May 2011.

² s 1b The *Archives Act 1983* (Tas)

³ The National Archives (UK), 'Digital Preservation Guidance Note 1 - Selecting File Formats for Long-Term Preservation', 19 June 2003, Accessed Jan 2011 http://www.nationalarchives.gov.uk/documents/selecting_file_formats.pdf, p. 4.

⁴ State Records New South Wales, 'Digital Records Preservation Requirements', <http://www.records.nsw.gov.au/recordkeeping/topics/digitalrecordkeeping/digital-records-preservation-discussion-paper/digital-records-preservation-requirements> , Accessed Feb 2011.

State records. This section also requires all relevant authorities to take all reasonable steps to comply with these guidelines, and put them into effect.

Keyword	Interpretation
MUST	The item is mandatory.
MUST NOT	Non-use of the item is mandatory.
SHOULD	Valid reasons to deviate from the item may exist in particular circumstances, but the full implications need to be considered before choosing this course.
SHOULD NOT	Valid reasons to implement the item may exist in particular circumstances, but the full implications need to be considered before choosing this course.
RECOMMENDS RECOMMENDED	The item is encouraged or suggested.

'MUST' and 'MUST NOT' statements are highlighted in capitals throughout the Guideline. Agencies deviating from these MUST advise TAHO of the decision to waive particular requirements.

Agencies deviating from a 'SHOULD' or 'SHOULD NOT' statement MUST record:

- the reasons for the deviation,
- an assessment of the residual risk resulting from the deviation,
- the date at which the decision will be reviewed, and
- whether the deviation has management approval.

Agencies deviating from a 'RECOMMENDS' or 'RECOMMENDED' requirement are encouraged to document the reasons for doing so.

2 Preserving Digital Records

There are hundreds of digital file formats currently used by Tasmanian government agencies. In most cases the issue of limiting file formats cannot be avoided, unless an archives is willing either to accept all formats and commit to preserving them all (which is unlikely to be successful) or, to accept records for which there are no fully-developed preservation plans and rely on the development of tools to migrate and render them in the future. Documentation available on the subject of file formats confirms that limiting the number of file formats for accession is often necessary.⁵

Generally, open formats should be used whenever possible, available formats should be considered next and proprietary formats should be considered as a last resort.⁶

⁵ InterPARES 2 Project: International Research on Permanent Authentic Records in Electronic Systems, 'General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation', Dec 2006
[http://www.interpares.org/display_file.cfm?doc=ip2_file_formats\(complete\).pdf](http://www.interpares.org/display_file.cfm?doc=ip2_file_formats(complete).pdf). Accessed May 2011

⁶ Library and Archives Canada, Home>About Us>Digital Initiatives>Digital Policies, Guidelines and Tools>Local Digital Format Registry (LDFR) File Format Guidelines for Preservation and Long-term Access Version 1.0 at LAC 22-10-2010 <http://www.collectionscanada.gc.ca/digital-initiatives/012018-2220.01-e.html#n> Accessed May 2011.

The file format suitable for long term retention will depend on the kind of information preserved, the uses to which the objects may be put in the future, the expectations of current and future users, and how far into the future the objects are intended to remain useful.⁷

3 Characteristics of file formats for preservation

The following criteria should be considered when selecting file formats for long term preservation:

3.1 Open and accessible formats

Using open formats can avoid loss of control over government owned information as a result of changes in commercial arrangements.⁸ Nonproprietary, open standards are usually more fully documented and more likely to be supported by tools for validation than proprietary formats.⁹

3.2 Ubiquity

Formats that are in widespread use are more likely to have ongoing and extensive support from software suppliers and user communities,¹⁰ and tools for migration and emulation are more likely to emerge from industry.¹¹

3.3 Metadata support.

Formats that support the inclusion of metadata are highly recommended as the metadata provides vital information on the provenance and technical characteristics of the information. The Tasmanian Archive and Heritage office endorse the *Australian Government Recordkeeping Metadata Standard*.

3.4 Feature set

Although formats should be selected which support the full range of features and functionality required for business processes, it is equally important to avoid choosing over-specified formats. In general, the more complex the format, the more costly it will be to manage and preserve.

3.5 Interoperability

The ability to exchange electronic records with other users and IT systems is an important consideration. Formats which are supported by a wide range of software or are platform-independent are highly desirable.

⁷ Hodge, G and N Anderson, 'Formats for digital Preservation: A review of alternatives and issues, 2007, <http://iospress.metapress.com/content/9v61373118049755/fulltext.pdf> Accessed Feb 2011.

⁸ State Records New South Wales, 'Policy on digital records preservation', Accessed Jan 2011 <http://www.records.nsw.gov.au/recordkeeping/governmentrecordkeeping-manual/documents/recordkeeping-policies/Policy%20on%20digital%20records%20preservation%20Final.pdf> p.5.

⁹ Arms, C and Carl Fleischhauer, compilers, *Sustainability of Digital Formats: Planning for Library of Congress Collections*. Washington, D.C.: Library of Congress, updated March 6, 2006. <http://www.digitalpreservation.gov/formats/sustain/sustain.shtml> Accessed May 2011.

¹⁰ The National Archives, 'Selecting File Formats for Long-Term Preservation', Accessed 20/1/2011 <http://www.nationalarchives.gov.uk/documents/selectingfile-formats.pdf>

¹¹ Arms, C and Carl Fleischhauer, compilers, *Sustainability of Digital Formats: Planning for Library of Congress Collections*. Washington, D.C.: Library of Congress, updated March 6, 2006. <http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>

3.6 Viability

Formats that provide error-detection facilities to allow detection of file corruption which may have occurred during transmission are recommended.¹² One such example is PNG (Portable Network Graphic), an image format which includes multiple methods for checking integrity built into the file itself.

4 Destroy digital records when appropriate

Maintaining accessibility to digital records over time is an expensive and complex undertaking and should only be considered for records that are required to be kept for long term retention or as State archives.

Government agencies should actively dispose of their records using appropriate disposal schedules. All authorised disposal schedules are listed on the GISU website. All agencies are encouraged to develop a disposal schedule.

5 Monitor for Obsolescence

All organisations with digital records need to monitor their condition, the ongoing viability of the systems they are contained in and the support for the format they are stored in. This will require knowledge of the systems and formats used within the organisation and an awareness of changes by vendors, forthcoming obsolescence, impending withdrawal of vendor support or other factors that may affect organisational systems and digital resources. There are a number of tools and resources that can assist with monitoring digital records for obsolescence and to therefore facilitate migration.¹³

There should also be an ongoing program for performing fixity or integrity checks, such as checksums, and taking restoration actions identified as necessary by such monitoring; this is important for monitoring the stability of bitstreams over time.¹⁴

Properly maintaining and refreshing the hardware systems is also important, so that the records have the best chance of surviving over time.

¹² State Records NSW Guideline 'Managing Digital Records', Accessed 20/1/2011 <<http://www.records.nsw.gov.au/recordkeeping/governmentrecordkeeping-manual/guidance/documents/recordkeeping-digital/Managing%20digital%20records.pdf>> For further information on total destruction of digital records see State Records New South Wales, 'Recordkeeping in Brief 51 - Destroying digital records: When pressing 'delete' is not enough', at<http://www.records.nsw.gov.au/recordkeeping/rib_51_destroying_digital_reco_15307.asp>.

¹³ For further information on these tools see pp.11-12 State Records NSW Guideline 22 'Managing Digital Records', Accessed 20/1/2011 <<http://www.records.nsw.gov.au/recordkeeping/government-recordkeeping-manual/guidance/documents/recordkeepingdigital/Managing%20digital%20records.pdf>>

¹⁴ For further information see Paradigm, 'Digital Preservation Strategies: Bitstream Preservation', Jan 2008 <http://www.paradigm.ac.uk/workbook/preservation-strategies/degree-bitstream.html> Accessed Jan 2011.

6 Back-up Systems

6.1 A backup system is not a recordkeeping system

Information technology professionals often use the term 'archiving' to describe back-up regimes. For recordkeeping purposes conducting back-ups does not constitute an archiving or preservation strategy, it is a business continuity or disaster recovery precaution.¹⁵

Backup systems capture a complete image of a file server at a particular point in time. They serve disaster management purposes, allowing data to be restored as it was at that particular point in time. Backup systems are not recordkeeping systems; they are difficult to search and cannot be relied upon as the means to maintain your organisational records.¹⁶

7 Definitions

Born digital records - 'Born digital' records (those that began their life in digital form) such as emails, web pages or database records.¹⁷ Other common types include documents, spreadsheets, presentations and online transactions. Born digital records are differentiated from digital materials that have been created as a result of converting non-digital source material, and non-digital materials which might have originated from a digital source but have been printed to paper or otherwise converted into analogue form.¹⁸

Formats

- **Available formats** - Proprietary formats with freely available specifications.¹⁹
- **File format** - The specific, pre-established structure for arranging data within a file.²⁰ This is done as a stream of bits (ones and zeros) and is unique to each file format. Storing data in a specific file format facilitates the storage, retrieval, processing, presentation, and/or transmission of information by software which supports it.²¹ TIFF, WAVE, AVI and other common image and audiovisual formats are not actually file "formats" as such, but rather file "container formats" (also known as "wrappers"). These are designed to combine multiple bitstreams into a single file.²²

¹⁵ Archives New Zealand: Government Recordkeeping Group, 'Digitisation Standard' 2007, <http://archives.govt.nz/sites/default/files/s6.pdf> Accessed April 2011

¹⁶ State Records NSW Guideline 'Managing Digital Records', Accessed Feb 2011 <<http://www.records.nsw.gov.au/recordkeeping/governmentrecordkeeping-manual/guidance/documents/recordkeeping-digital/Managing%20digital%20records.pdf>>

¹⁷ State Records New South Wales, 'Policy on digital records preservation', Accessed Jan 2011 <http://www.records.nsw.gov.au/recordkeeping/governmentrecordkeeping-manual/documents/recordkeeping-policies/Policy%20on%20digital%20records%20preservation%20Final.pdf> p.5

¹⁸ International Organization for Standardization: Technical Report, 'Information and documentation – Implementation guidelines for digitisation of records', First edition 12-1-2010 ISO/TR13028, Accessed 17 March 2011.

¹⁹ UK Data Archive, *UK Data Archive Preservation Policy*. Colchester: University of Essex, version 2.0, September, 2005, p. 19. <http://www.dataarchive.ac.uk/news/publications/UKDAPreservationPolicy0905.pdf> as cited in InterPARES 2 Project 'General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation', 2006 < [http://www.interpares.org/display_file.cfm?doc=ip2_file_formats\(complete\).pdf](http://www.interpares.org/display_file.cfm?doc=ip2_file_formats(complete).pdf)> Accessed Jan 2011.

²⁰ Preservation Metadata: Implementation Strategies (PREMIS), *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group*. United States: On-Line Computer Library Centre and Research Libraries Group, May 2005, p. 237. <http://www.oclc.org/research/projects/pmwg/premisfinal.Pdf>

²¹ The InterPARES 2 Project Dictionary, InterPARES 2 Web site, "File format," http://www.interpares.org/ip2/display_file.cfm?doc=ip2_dictionary.pdf

²² InterPARES 2 Project: International Research on Permanent Authentic Records in Electronic Systems, 'General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation', Dec 2006 [http://www.interpares.org/display_file.cfm?doc=ip2_file_formats\(complete\).pdf](http://www.interpares.org/display_file.cfm?doc=ip2_file_formats(complete).pdf). Accessed Jan 2011.

- **Proprietary/closed formats** - the specification on how the file format is arranged is not publicly available. It is generally controlled by a corporate entity.
- **Open formats** - file formats for which the specification is made freely available, often created and maintained by a standards group such as International Standards Organisation (ISO).²³ These formats are ideal for long term digital preservation.

Metadata - Essentially data about data. More specifically it is information about the content, context and the processes that manage and maintain records. Metadata ensures that records can be found, read and understood, both in the immediate and long term. Examples of records management metadata include the record's title, who has seen it and when, the level of security around the record, and when it should be destroyed or transferred.²⁴

²³ InterPARES 2 Project 'General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation', 2006 <[http://www.interpares.org/display_file.cfm?doc=ip2_file_formats\(complete\).pdf](http://www.interpares.org/display_file.cfm?doc=ip2_file_formats(complete).pdf)> Accessed Jan 2011.

²⁴ The National Archives of Australia, 'Describing Records Using Metadata', <http://www.naa.gov.au/records-management/create-capturedescribe/describe/index.aspx>, Accessed Jan 2011.

Further Advice

For more detailed advice, please contact:

Government Information Strategy Unit
Tasmanian Archive and Heritage Office
91 Murray Street
HOBART TASMANIA 7000
Telephone: 03 6165 5581
Email: gisu@education.tas.gov.au

Acknowledgements

- This document draws from the State Records New South Wales Guideline 22, 'Managing Digital Records'⁵, Library and Archives Canada 'File Format Guidelines for Preservation and Long-term Access'⁶ and includes input from the National Archives of Australia.
- Archives New Zealand>Advice on Records & Archiving>Continuum Resource Kit>'S6:DigitisationStandard' 2007, <http://archives.govt.nz/sites/default/files/s6.pdf>
- Arms, C and Carl Fleischhauer, compilers, Sustainability of Digital Formats: Planning for Library of Congress Collections. Washington, D.C.: Library of Congress, updated March 6, 2006. <http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>
- 'Capture Your Collections' Glossary, <http://hosting.collectionsaustralia.net/capture/course/glossary.html>
- Enotes.com, 'Lossy Compression', http://www.enotes.com/topic/Lossy_compression
- Guenther, R & R Wolfe, 'Integrating Metadata Standards to Support Long-Term Preservation of

Digital Assets: Developing Best Practices for Expressing Preservation Metadata in a Container Format', 10 May 2009 <http://escholarship.org/uc/item/0s38n5w4>

- Hodge, G and N Anderson, 'Formats for digital Preservation: A review of alternatives and issues, 2007, <http://iospress.metapress.com/content/9v61373118049755/fulltext.pdf>
- International Association of Sound and Audiovisual Archives, 'Guidelines on the Production and Preservation of Digital Audio Objects', Second Edition 2009, <http://www.iasa-web.org>
- InterPARES 2 Project: International Research on Permanent Authentic Records in Electronic Systems, 'General Study II Final Report: Selecting Digital File Formats for Long-Term Preservation', Dec 2006 [http://www.interpares.org/display_file.cfm?doc=ip2_file_formats\(complete\).pdf](http://www.interpares.org/display_file.cfm?doc=ip2_file_formats(complete).pdf)
- Library and Archives Canada, Home>About Us>Digital Initiatives>Digital Policies, Guidelines and Tools>Local Digital Format Registry (LDFR) File Format Guidelines for Preservation and Long-term Access Version 1.0 at LAC 22-10-2010 <http://www.collectionscanada.gc.ca/digitalinitiatives/012018-2220.01-e.html#n> Accessed May 2011.
- Library of Congress, 'Sustainability of Digital Formats Planning for Library Congress Collections', <http://www.digitalpreservation.gov/formats/index.shtml> Accessed May 2011

- Millar, L., International Records Management Trust, 'Training in Electronic Records Management: Preserving Electronic Records', 2009
http://www.irmt.org/documents/educ_training/term%20modules/IRMT%20TERM%20Module%204.pdf
- National Archives of Australia, 'Digital Recordkeeping: Guidelines for Creating, Managing and Preserving Digital Records', May 2004 http://www.naa.gov.au/images/digital-recordkeepingguidelines_tcm2-920.pdf
- National Archives of Australia, 'Storing Digital Records',
<http://www.naa.gov.au/recordsmanagement/secure-and-store/NAA-storage/e-storage/index.aspx>
- National Archives and Records Administration, 'Records Management Self-Assessment', 2009,
<http://www.archives.gov/records-mgmt/resources/self-assessment.pdf>
- National Archives and Records Administration, 'Records Management>Initiatives>Transfer of Electronic Records' Nov 2009 <http://www.archives.gov/records-mgmt/initiatives/transfer-records-tonara.html>
- Paradigm, 'Digital Preservation Strategies: Bitstream Preservation', Jan 2008
<http://www.paradigm.ac.uk/workbook/preservation-strategies/degree-bitstream.html>
- Preservation Metadata: Implementation Strategies (PREMIS), *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group*. United States: On-Line Computer Library Centre and Research Libraries Group, May 2005, p. 237. <http://www.oclc.org/research/projects/pmwg/premis-final.pdf>
- Queensland State Archives, 'Digital Archiving Discussion Paper: Informing an approach to the longterm management and preservation of government digital records', May 2010
<http://www.archives.qld.gov.au/downloads/QSA%20Digital%20Archiving%20Discussion%20Paper.pdf>
- Queensland State Archives, 'Electronic Recordkeeping Issues', Oct 2010
<http://www.archives.qld.gov.au/government/RecordKeepingIssues.asp>
- State Records NSW Guideline 22 'Managing Digital Records' 2009,
<http://www.records.nsw.gov.au/recordkeeping/government-recordkeepingmanual/guidance/documents/recordkeeping-digital/Managing%20digital%20records.pdf>
- State Records New South Wales, 'Policy on digital records preservation', December 2007
<http://www.records.nsw.gov.au/recordkeeping/government-recordkeepingmanual/documents/recordkeepingpolicies/Policy%20on%20digital%20records%20preservation%20Final.pdf>
- State Records New South Wales, 'Standard No. 10 Standard on digital recordkeeping' Sept 2008
<http://www.records.nsw.gov.au/recordkeeping/government-recordkeepingmanual/rules/standards/standard-on-digital-recordkeeping/standard-on-digitalrecordkeeping/?searchterm=standard on digital recordkeeping>
- Technical Advisory Service for Images 'Advice Paper: Choosing a File Format' May 2006
<http://www.cab.unipd.it/system/files/format.pdf>
- Technical Advisory Service for Images 'File Formats and Compression'. United Kingdom: March 2005.
<http://www.tasi.ac.uk/advice/creating/fformat.html#f3>
- The National Archives (UK), 'Digital Preservation Guidance Note 1 - Selecting File Formats for Long-Term Preservation', 19 June 2003
http://www.nationalarchives.gov.uk/documents/selecting_file_formats.pdf

- The National Archives (UK), 'Digital Preservation Guidance Note 2 – Selecting Storage Media for Long Term Preservation' Aug 2008 <http://www.nationalarchives.gov.uk/documents/selectingstorage-media.pdf>
- The National Archives (UK), 'Digital Preservation Guidance Note 3 – Care, Handling and Storage of Removable Media' Aug 2008 <http://www.nationalarchives.gov.uk/documents/removable-mediacare.pdf>
- The National Archives (UK), 'Digital Preservation Guidance Note 5 – Image Compression' Aug 2008 <http://www.nationalarchives.gov.uk/documents/image-compression.pdf>
- Wikipedia http://en.wikipedia.org/wiki/Free_software_Unicode_typefaces
- Wikipedia http://en.wikipedia.org/wiki/Unicode_typefaces#List_of_Unicode_fonts

Appendix I - Table of Recommended File Formats²⁵

Recommended - The Tasmanian Archive and Heritage office recommend that high value, permanent and long term records are either created in a recommended format or migrated shortly after their creation or shortly after their active business use has ceased.

Acceptable for Transfer - While these formats are not actively recommended, we recognise that many are in widespread use and that it may not be feasible to migrate them into a recommended format. Please note that there is no implied migration path from the **acceptable for transfer** formats to the **recommended** for preservation formats. The selection of a preservation format will be based on the degree to which the significant properties of the source format are retained in the target preservation format (and the relative importance of specific properties).²⁶

Content Type	Recommended	Acceptable for Transfer
Text	<ul style="list-style-type: none"> • Extensible Hypertext Markup Language (XHTML) • Extensible Markup Language (XML) • Hypertext Markup Language (HTML) • Multipurpose Internet Mail extensions (MIME) • Open Document Format (ODF) • PDF for long-term preservation (PDF/A) • Rich Text Format (RTF) • Standard General Markup Language (SGML) • Text (plain text) 	<ul style="list-style-type: none"> • Microsoft Office including: Word Document Format, Excel Spreadsheet Format, PowerPoint Presentation Format • Portable Document Format (PDF)
Still Images	<ul style="list-style-type: none"> • Joint photographic experts group JPEG 2000 (JP2) (uncompressed) • Tagged Image file format (TIFF) (uncompressed) • TIFF - GeoTIFF 	<ul style="list-style-type: none"> • Encapsulated postscript (EPS) • Graphics interchange format (GIF) • Joint photographic experts group (JPEG)

²⁵ Library and Archives Canada, Home>About Us>Digital Initiatives>Digital Policies, Guidelines and Tools>Local Digital Format Registry (LDFR) File Format Guidelines for Preservation and Long-term Access Version 1.0 at LAC 22-10-2010 <http://www.collectionscanada.gc.ca/digital-initiatives/012018-2220.01-e.html#n> Accessed May 2011.

²⁶ Library and Archives Canada, Home>About Us>Digital Initiatives>Digital Policies, Guidelines and Tools>Local Digital Format Registry (LDFR) File Format Guidelines for Preservation and Long-term Access Version 1.0 at LAC 22-10-2010 <http://www.collectionscanada.gc.ca/digital-initiatives/012018-2220.01-e.html#n> Accessed May 2011

Content Type	Recommended	Acceptable for Transfer
	<ul style="list-style-type: none"> • Portable network graphics (PNG) • SVG (Scalable Vector Graphic) 	
Audio	<ul style="list-style-type: none"> • Broadcast Wave Format (BWF) (for newly digitised content i.e., creating) • FLAC_1_1_2 Free Lossless Audio Codec Version 1.1.2 (FLAC) • Waveform Audio Format (WAV) (for migrating born digital audio content) 	<ul style="list-style-type: none"> • Audio Interchange File Format (AIFF) • Mpeg-1 layer-3, Mepeg-2 layer-3 (MP3) • Mpeg-4 aac - advanced audio coding (AAC)
Web Archiving	<ul style="list-style-type: none"> • Internet archive format (ARC) • Web archive format (WARC) 	
Digital Video	<ul style="list-style-type: none"> • Motion JPEG 2000 • Dirac • Raw bitstream 	<ul style="list-style-type: none"> • MPEG-1, MPEG-2 (*.mpg, *mpeg, wrapped in AVI, MOV) • MPEG-4 (H.236, H.264) (*.mp4, wrapped in AVI, MOV) • Windows media video (WMV)
Databases	<ul style="list-style-type: none"> • Software independent Archiving of Relational Databases (SIARD) • Delimited Flag file with DDL 	<ul style="list-style-type: none"> • dBase Format (DBF) • SQL Structured Query Language3 (export)
Spreadsheets	<ul style="list-style-type: none"> • Comma-separated values (CSV) • ODS (Open Document Spreadsheet) 	<ul style="list-style-type: none"> • Microsoft Excel
Computer Aided Design - Technical Drawing	<ul style="list-style-type: none"> • Drawing Interchange File format/Data eXchange Format (DXF) 	<ul style="list-style-type: none"> • Computer Graphics Metafile (CGM)

For technical information on the details of each file format visit Library of Congress Sustainability of Digital Formats Planning for Library of Congress Collections and search Format Descriptions.²⁷

²⁷ <http://www.loc.gov/>

Appendix 2 - Guidelines for creating Word Documents, Spreadsheets and Emails²⁸

Documents

Choose an open format to begin with: The single greatest effect on the long-term availability of a record begins when you click the save button. Create your documents in a free format, such as ODF (Open Document Format, ISO/IEC 26300), and send read-only versions to others in another open document format such as PDF 1.7 (Portable Document Format, ISO/IEC 32000-1). Microsoft Word has had native support for saving in ODF since Office 2007 SP2.

Always begin with an empty template: Do not create a new document by changing an existing document based on the same template. Templates often include metadata regarding context. Copying an existing document means you may save incorrect metadata from the original document to the new document or risk that not all relevant new metadata will be completed.

Minimise the use of Macros within your document: Macros in proprietary file formats (and some open formats like OOXML ECMA-376) are platform and file format specific. Use of Macros in these formats greatly reduces interoperability and greatly increases the risk of loss to the record in the future. Free formats such as ODF tend to use cross platform languages such as Python.

Use styles to impart structure to documents: Format styles in templates (eg Heading 1, Heading 2, Heading 3) impart structure to documents which not only increases standardisation and readability, but also helps preserve meaning.

Do not use passwords to protect documents: If you forget your password you will not be able to open the document and the information will no longer be accessible. Passwords for editing the document can be set, however, and do not compromise the digital sustainability of the document. Where sensitive material is required, follow internal security guidelines on safeguarding data

Use standard fonts: Unconventional fonts can be lost in migrations and reduce the probability of authentic preservation of word processed records. Use free fonts²⁹ (such as Liberation or DejaVu) if possible, as these are free of licensing issues. Unicode fonts such as Arial, Tahoma and Times New Roman are also recommended.

Use headers and footers to include suitable metadata: Headers and footers are ideally suited to the capture of metadata such as the name of the file, version number, date etc. Prompts to record this information can be built into templates

²⁸ Digital Preservation Testbed, *From digital volatility to digital permanence*, Part 3, Preserving text documents, pp.48-52, <http://www.digitaleduurzaamheid.nl/bibliotheek/docs/volatility-permanence-textdocs-en.pdf> as cited in State Records NSW Guideline 22 'Managing Digital Records', <<http://www.records.nsw.gov.au/recordkeeping/government-recordkeeping-manual/guidance/documents/recordkeepingdigital/Managing%20digital%20records.pdf>> pp. 5-9.

²⁹ Wikipedia, 'Free software Unicode typefaces', http://en.wikipedia.org/wiki/Free_software_Unicode_typefaces Accessed Feb 2011

Avoid date/time insert fields: The automatic capture of dates and times in templates is undesirable as they are updated every time the document is opened. The only automatic date field which may be suitable to use is Print Date. If used, it should clearly indicate it is the last printing date.

Insert any images or illustrations in suitable formats: Images and illustrations should be inserted in formats that are more likely to be preserved in the long term, such as PNG (Portable Network Graphics).

Do not use text boxes when a table would be more appropriate: Many documents use spaces to vertically align information into columns but during migration this layout could be lost. The best way to create columns is to use a table or a specific alignment.

Give preference to object embedding over linking: The information contained in linked objects is only updated when changes are made to the source file. The target file only stores information about the location of the source file and displays an image (icon) of the linked file. Linked objects should only be used when it is necessary to restrict file size. If using links, the link between the source and target file should be removed when the text document has acquired its definitive form and can no longer be changed. Embedded objects are incorporated as part of the target file. With an embedded object, the information in the target file is not changed when the source file is changed because the object no longer makes use of the source file.

Additional Guidelines for Spreadsheets

Assign meaningful names to rows and columns (titles or labels): This provides important contextual information to the spreadsheet to enable it to be understood, even with the passing of time.

Avoid automatic date and time functions like “=NOW()”: The result of the NOW() function is the current date and time in the form of a serial number that is automatically recalculated each time the spreadsheet is opened, which is undesirable.

State currency in a separate cell and in full: When making use of currency amounts state the relevant currency in the title or name of the column. An integrated currency symbol eg typing \$AUD1,000 into a column can be lost in migration or replaced by a different currency symbol.

Emails³⁰

Always use the address book in your email application: The address book contains extra information about the people to whom you are sending messages. This information is stored together with the email message for context. Address book information should be filled in as completely as possible.

Be careful when using distribution lists for email messages if addresses are needed to provide contextual information: Names and address and other details of recipients of emails are not always

³⁰ Digital Preservation Testbed, *From digital volatility to digital permanence*, Part 3, Preserving text documents, pp.48-52,

<http://www.digitaleduurzaamheid.nl/bibliotheek/docs/volatility-permanence-textdocs-en.pdf> as cited in State Records NSW Guideline 22 'Managing Digital Records', <<http://www.records.nsw.gov.au/recordkeeping/government-recordkeeping-manual/guidance/documents/recordkeepingdigital/Managing%20digital%20records.pdf>> pp. 5-9.

registered on distribution lists. Depending on the email application and type of email, the name of the list and people on the list will not appear in the email. Distribution lists are dynamic (ie names are added and removed constantly) and it is nearly impossible to tell who is on a distribution list at the particular time an email was sent. Information about who the email is being sent to can be added to the text of an email sent via a distribution list. Whether this is necessary depends on the information being sent and the importance of knowing who is receiving it eg if you are circulating a draft you may need to prove it has been sent to all relevant people for their comment before being released.

Always give email messages a subject: This is contextual information and enables the email to be sorted, evaluated by recipients and accessed in the future. The subject line should be relevant and useful.

Only add 'flags' to messages such as urgency or sensitivity flags when necessary: Although flags can be useful, not all email applications can reproduce them correctly and they may not be seen by the recipient. It is better to include this information in the subject line or body of the message.

Wherever possible make and send email messages in plain text or HTML: Messages in plain text are suitable for simple emails. More formal and official emails eg with images or logos can be made in HTML. Beware of using Microsoft Outlook's Rich Text Format (RTF) because this is specific to Outlook and when sent is coded in a file that may contain attachments and layout information. This data has to be translated to be read by other email applications. MS Exchange does this translation and what is sent depends on the Exchange settings.

Do not use automatically updating fields in email messages: Always enter this information as 'hard' text. Automatic fields are unstable and may update every time you open an email causing the content and context of the email to be lost. Seek to send attachments in open formats and do not paste images into an application like MS PowerPoint or Word. These applications produce files in proprietary format.

When replying to an email message do not insert text into the original message: If you wish to respond add your comments at the top above the headers of the original message. This will keep your comments separate from the original. Do not forget your email may be read by someone in twenty or thirty years, when you cannot explain what you added.

Use a signature block: This provides important contextual information about you as the sender. It is worth using two signature blocks: one for internal emails and one for external use for more official correspondence. Internal signature blocks should contain name, position, project or department, organisation. External signature blocks should also contain: address, telephone number, email address and website.