Contents

1: Purpose ........................................................................................................................................ 2
2: Background .................................................................................................................................... 2
3: Context ........................................................................................................................................... 2
4: Scope ............................................................................................................................................. 3
5: Principles ....................................................................................................................................... 3
6: Critical Success Factors ..................................................................................................................... 4
7: Responsibilities ................................................................................................................................. 4
Appendix 1: Definitions ......................................................................................................................... 5
Appendix 2: Collections .......................................................................................................................... 7
  A2.1: Repositories ............................................................................................................................... 7
  A2.2: New Zealand Electronic Text Collections .................................................................................... 7
  A2.3: Other Digital Collections ............................................................................................................ 8
Appendix 3: Linkages ............................................................................................................................. 8
Appendix 4: Stakeholders ....................................................................................................................... 9
Appendix 5: Formats .............................................................................................................................. 9

Date: April 2014

Version: 5

Author: Michael Parry

Approved By:

Local Digital Collections Group - Date: April 2014

Library Leadership Team - Date: May 2014
1: Purpose
The Digital Preservation Framework articulates the Library’s principles for preserving its locally created digital collections. It is intended as a broad framework that will create a robust yet flexible environment where new collections and technologies can be adopted and included. The framework will provide steering for projects such as the creation of a Digital Object Management System.

2: Background
The overarching purpose of the Library is to support and enhance learning, teaching and research at Victoria University. Digital preservation is a key component, by ensuring there is continued access to digital resources through the provision of resilient, flexible policies and systems. This ensures the continued preservation of digital collections to enable access in perpetuity to the Library’s digital collections.

This Digital Preservation Framework is part of a range of documents [Appendix 3] related to the development and management of digital collections.

3: Context
The Library has a growing local digital collection developed both through its digitisation programmes and through the acceptance of digital donations. As the Library creates and acquires more digital content, both born digital and digitised, the Library must consider how it best preserves access to that content.

Digital preservation is a difficult and critical component in managing digital collections. The knowledge that Victoria University Library is not the only organisation grappling with digital preservation provides some pointers for us. The National Library of New Zealand and Archives New Zealand have developed their Digital Preservation Strategy\(^1\) which has been useful in the formulation of this document.

The key problems any organisation faces when considering digital preservation include:

- Format degradation – where either time or malicious software renders an object inaccessible.
- Format obsolescence – where a proprietary format is no longer supported.

---

• Loss of context – where important information about the digital object is lost.

Within the context of the University, the Library must also consider a number of other factors when dealing with digital preservation.

For example, the University ITS disaster recovery system which acts as a limited solution to digital preservation in that it only ensures the ongoing existence of the digital object. External vendors for digital preservation activities could be used to provide digital preservation. However the risk inherent in this approach is around continued access if the service provider becomes incapable of offering the service.

The Library needs to understand and monitor what is happening externally in this area. One example is from the National Library which has been doing ‘whole of domain harvests’ and ‘targeted harvests’ of a number of websites. These however cannot be relied on for preservation purposes because they can be limited in their capture.

4: Scope
This framework deals with Local Digital Collections. These are collections that are not purchased by the Library, but rather created and maintained by the Library, or on the behalf of the Library. These currently include:

• Digital repositories such as the Institutional Repositories, AVArchive and Exams database
• The New Zealand Electronic Text Collection (NZETC) a collection of physical works that have been transformed into digital objects.
• Other collections of digital objects that exist outside of the repository infrastructure and/or the NZETC infrastructure.

The framework excludes commercially acquired digital collections.

5: Principles
There are a number of principles that underpin this Framework. These are:

1. Digital Preservation is a critical Library process.
2. The Library adopts a three level view of digital preservation; primary preservation, secondary preservation and descriptive preservation.
3. The Library minimises the number of systems and environments it maintains.
4. Each new proposal for preservation should be considered against the existing framework and solutions, and new options should be explored.

5. The preservation of metadata around the digital objects is as important as the preservation of the digital objects themselves.

6. Internal digital preservation is the preferred option.

7. Where possible current digital objects can be considered either archival copies, or access copies.

8. The Library will implement best practice standards for archival and access copies of all digital objects.


10. The Library will endeavour to capture the best quality/industry standard metadata, both technical and descriptive, when creating new digital objects and accepting the deposit of digital objects.

11. The Library establishes procedures around quality control of existing digital objects.

12. Emulation is considered on a case-by-case basis.

13. The Library maintains a registry of collections and related knowledge base.

14. Future planning is central to the process of digital preservation.

6: Critical Success Factors
The Library will be successful in its digital preservation activities when it is able to show:

- Where every digital object is.
- What each digital object is, where it originates from.
- That every digital object is uncorrupted and accessible.
- Researchers can easily access any digital object.
- Each digital object is fully described.
- That there is a plan for maintaining continued access.

7: Responsibilities
The Local Digital Collections Committee has oversight of the Digital Preservation Framework. The Digital Initiatives team, within the Library Technology Service Group, is responsible for operationalising the Digital Preservation Framework.
## Appendix 1: Definitions

For the purposes of the Digital Preservation Framework the Library defines, using the Digital Preservation Coalition / *Digital Preservation Handbook* as a starting point\(^2\), the following:

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access</td>
<td>Is assumed to mean continued, ongoing usability of a digital resource, retaining all qualities of authenticity, accuracy and functionality deemed to be essential for the purposes the digital material was created and/or acquired for.</td>
</tr>
<tr>
<td>Access Copy</td>
<td>A copy made from a collection item for use so that the Archival Copy can be preserved and protected from damage.</td>
</tr>
<tr>
<td>Archival Copy</td>
<td>A copy of a document, data, software or other object that is maintained in a long-term storage media. Also referred to as a preservation copy.</td>
</tr>
<tr>
<td>Born Digital</td>
<td>Term is used to differentiate digital materials which have been created as a result of converting analogue originals (digitisation); and digital materials which have originated from a digital source.</td>
</tr>
<tr>
<td>Dark Archive</td>
<td>A dark archive is an archive with very restricted access. Items may be deposited but not accessed.</td>
</tr>
<tr>
<td>Descriptive preservation</td>
<td>The process of describing and recording both the descriptive and technical metadata around the preservation process.</td>
</tr>
<tr>
<td>Digital Objects</td>
<td>A broad term encompassing digital surrogates created as a result of converting analogue materials to digital form (digitisation), and &quot;born digital&quot;.</td>
</tr>
<tr>
<td>Digital Preservation</td>
<td>Refers to the series of managed activities necessary to ensure continued access to digital materials for as long as necessary. For the purposes of this framework the length of time is set as in perpetuity.</td>
</tr>
<tr>
<td>Digitisation</td>
<td>The process of creating digital files by scanning or otherwise converting analogue materials. The resulting digital copy, or digital surrogate, would then be classed as digital material and then subject to the same broad challenges involved in preserving access to it as &quot;born digital&quot; materials.</td>
</tr>
<tr>
<td>Emulation</td>
<td>A means of overcoming technological obsolescence of hardware and software by developing techniques for imitating obsolete systems on future generations of computers.</td>
</tr>
</tbody>
</table>

| Metadata | Information which describes significant aspects of a resource. This can be broken into two distinct groups, descriptive and technical. Descriptive metadata describes the digital object, for example title, author, and date. Technical metadata describes the means by which the digital object was created; for example, camera type, and settings. |
| Local Digital Collections | These are collections of digital items that have been created by Victoria University and/or are maintained by the Library. This can include collections where digital objects are created by, or on behalf of, the Library and the maintenance of the collection is outsourced. These do not include collections where the Library pays a subscription for access. These do not include other digital collections managed and owned by other groups within Victoria University. |
| Primary preservation | The process of ensuring continued access through backups and duplication |
| RDA | Resource Description and Access (RDA) is a standard for cataloguing that provides instructions and guidelines on formulating data for resource description and discovery. |
| RDF | The Resource Description Framework (RDF) is a family of World Wide Web Consortium (W3C) specifications originally designed as a metadata data model. It has come to be used as a general method for conceptual description or modelling of information that is implemented in web resources, using a variety of syntax notations and data serialization formats. |
| Secondary preservation | The process that ensures continued access through the creation and management of differing manifestations of digital objects. This revolves around the creation of either archival or access copies. |
Appendix 2: Collections

Several of the Library’s collections have been identified for digital preservation. These include:

A2.1: Repositories

The Library runs a number of different digital repositories including the Institutional Repositories, AVArchive and Exams database. These repositories contain digital objects, which have either been born digital or been digitised, and currently run on Dspace software. These collections, especially the Institutional Repositories, are identified as critically important to preserve.

The Institutional Repositories contain the record of the University’s research outputs. These include mandated deposit of electronic copies of research theses. There is a growing trend internationally to stop the physical deposit of theses and only accept digital copies. With the digital copy of the thesis then likely to be considered the preservation copy it essential that robust processes exist to preserve all theses deposited in the repositories.

The Pacific Island Public Policy Repository (PIPPR) is a collection maintained by the library in the Dspace environment, but with the collection maintained by (role) Graham Hassell. As the Library maintains the infrastructure that hosts the digital objects, it should be included in this Digital Preservation framework.

Current practice relies on backups run by ITS and does not meet best practice standards for digital preservation.

A2.2: New Zealand Electronic Text Collections

The New Zealand Electronic Text Collection (NZETC) is a collection of physical works that have been transformed into digital. There are a range of types of digital objects associated with each work. These include EPUB3, PDF4, TEI5-XML6 and DAISY7 audio books. Due to the nature of the NZETC, with its inconsistent format use and display, the preservation of the NZETC is going to be complex.

---

3 EPUB (short for electronic publication) is a free and open e-book standard by the International Digital Publishing Forum (IDPF). Files have the extension .epub.
4 PDF (Portable Document Format) is a file format used to represent documents in a manner independent of application software, hardware, and operating systems
5 TEI (Text Encoding Initiative) is a form of XML
6 XML (Extensible Markup Language) is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable
7 DAISY (Digital Accessible Information System) is a technical standard for digital audio books, periodicals and computerized text.
Current practice relies on backups run by ITS and does not meet best practice standards for digital preservation.

The University Library is in the process of formulating a pilot project with the National Library of New Zealand to establish parameters for a preservation process using the NDHA. The Salient corpus within the NZETC, encompassing the years 1960-1969, have been targeted as the test collection for the project. The project calls for the digital objects from this corpus to be included in the NDHA.

**A2.3: Other Digital Collections**

Other Digital Collections that exist outside of the repository infrastructure and/or the NZETC infrastructure. They are digital objects held on Library websites. Examples of these include the Dan Long poster collection, the digitised videos (VHS tapes) hosted by ETV on behalf of the Library, and the Open Journal Systems administered by the Library.

**Appendix 3: Linkages**

This framework is linked to the following documents:

1. Terms of reference for Digital Collections Group
2. Collection Development and Management Policy
3. Institutional Repository Collection Level Statement
   [http://library.victoria.ac.nz/library/about/policies/slcs/institutionalrepository.html](http://library.victoria.ac.nz/library/about/policies/slcs/institutionalrepository.html)
4. New Zealand Electronic Text Collection (NZETC) Collection Level Statement
   [http://library.victoria.ac.nz/library/about/policies/slcs/NZelectronictextcollection.html](http://library.victoria.ac.nz/library/about/policies/slcs/NZelectronictextcollection.html)
5. Metadata Framework [to be written]

---

8 NDHA (National Digital Heritage Archive) is the National Library of New Zealand's business and technical approach to managing the collection and preservation of digital heritage and providing ongoing access to the library's digital collections

Appendix 4: Stakeholders
The following are stakeholders who either could be consulted or have an interest in the framework.

- National Library of New Zealand
- School of Engineering and Computer Science
- Information Technology Services

Appendix 5: Formats
Best practice formats currently adopted can be found in the Formats Register.