Yale University Library's Digital Preservation Policy Framework

November 2014

Version 1.0

Note: The single-page list of principles that form the core of this framework is on page 4

Contents

Overview

This policy framework has been developed to provide a common understanding of how digital preservation activities will be organized, managed and undertaken at Yale University Library (YUL) and to outline the specific commitments YUL is making in relation to those digital preservation activities.

The framework is a living document that will be updated over time as conditions change within the YUL operating context. In its current version the framework consists of a Digital Preservation Strategy made up of ten guiding principles and five supporting policy statements. The strategy's principles provide a summary of the commitments that YUL is making in relation to undertaking digital preservation activities. The policy statements provide further details to clarify the principles where there may be ambiguity due to their (intentional) brevity.

The framework will be supported by numerous additional documents including documented digital preservation and access plan for specify types of digital objects. An example digital preservation and access plan is included in Appendix 4: Example Structured Data Digital Preservation and Access Plan

Scope, Applicability, and Audience

- This policy framework applies to all digital preservation activities undertaken within the YUL system.
- This policy framework is intended to outline what can be expected from the digital preservation services of YUL and to place reasonable limits on those expectations given expected operating constraints and technological limitations.
- The principles in this policy framework apply to digital preservation activities only. They do not, and are not, intended to restrict the ability of curators, selectors and acquisitions staff to acquire whatever digital content they are empowered to collect.
- The policy framework may form the basis for service level agreements if YUL provides digital preservation services to non-YUL partners in the future.
- The intended audience for this framework is wide and includes (but is not limited to:
    - The YUL community and all related stakeholders
    - Content owners within YUL
    - Users of digital information at YUL
    - Digital Preservation staff
    - IT staff who are involved with digital preservation at YUL
    - Prospective digital preservation partners
    - Digital preservation practitioners and policy developers

Digital Preservation Strategy

Yale University Library (YUL) will adhere to the following principles in order to ensure the preservation of digital content that it is responsible for.

Commitments to the provision of digital preservation services

Principle 1: Provision of services

YUL will provide digital preservation services to ensure the preservation of digital content for which it is responsible, for as long as it is needed

Principle 2: Collaboration

Whenever appropriate and efficient YUL will work with third parties, such as other universities or digital preservation service providers, to preserve digital content

Principle 3: Compliance and accountability

YUL will endeavor to be standards compliant and will implement regular audits of digital preservation services in order to monitor this commitment

Commitments to the use of digital preservation services

Principle 4: Agreement on services

At point of ingest the YUL and the depositor(s) will agree on the content to be preserved and the preservation services to be provided for preserving the agreed content

Principle 5: Efficient approaches

YUL will use standard sets of services to enable economies of scale to be realized.

Principle 6: Technological openness

YUL accepts digital content of all types (including all digital formats) for preservation provided funding is available to provide the agreed level of service for the agreed preservation period.

Principle 7: Security and availability

YUL will ensure preserved digital objects are securely stored and appropriate access controls are in place

Commitments to the preservation of digital content

Principle 8: Preserving integrity and authenticity

YUL will ensure that the integrity and authenticity of preserved digital objects can be proven from point of ingest

Principle 9: Obsolescence monitoring

YUL will monitor preserved digital objects to identify when content is in danger of becoming inaccessible due to technological change

Principle 10: Preservation approaches

YUL will support the use of multiple digital preservation approaches including content migration, hardware emulation, and software and hardware dependency preservation.

Supporting digital preservation policies

YUL's digital preservation strategy and its principles are supported by a number of policies that further detail how the strategy's principles will be implemented in practice.

Supporting policy 1

YUL will ensure the digital preservation services are compliant with the following standards and will audit the repository against those standards at least once every 5 years:

- ISO 14721:2012: Space data and information transfer systems -- Open archival information system (OAIS) -- Reference model
- ISO 16363:2012: Space data and information transfer systems -- Audit and certification of trustworthy digital repositories

Supports:

Principles 3 and 7

Rationale:

YUL needs to ensure it is compliant with relevant international standards covering digital preservation activities in order to:

- Be trusted by content owners and users of digital content preserved by YUL
- To ensure YUL's digital preservation services meet minimum practices
- To match the commitments of peer institutes

Supporting policy 2

YUL will not limit digital objects accepted for preservation solely based on technological challenges

Supports:

Principles 4, 5 and 6

Rationale:

Decisions about what content to acquire/select and therefore preserve are not covered by this policy framework. However some digital objects may be technically challenging and thus expensive to preserve and YUL cannot be expected to commit to preserving these objects without a clear understanding of what that entails and how it will be funded.

Conversely, YUL should not limit its commitment to preserve technically challenging content just because it is technically challenging as in many cases the content's value may justify additional expense. Therefore YUL will not limit content accepted for preservation provided the following are agreed upon at point of deposit:

- What content in the submitted digital objects needs to be preserved
- What dependencies the digital objects have

- What preservation activities will need to be undertaken to preserve the content
- How the preservation of the content will be funded

Additionally, YUL cannot be expected to commit to preserving any digital content if the Library:

a) Does not know what needs to be preserved
b) Cannot judge the extent of that commitment, e.g. does not know what dependencies may need to be maintained in order to preserve access to deposited content, or what preservation strategies might need to be resourced and used in order to preserve the content.
c) Has not been funded to preserve the content

These factors are key responsibilities that need to be documented and in place for any digital repository that seeks to be OAIS and ISO 16363 (Trustworthy Repositories Audit and Certification) compliant.

Example of application:

An obscure piece of digital artwork may be submitted to YUL for preservation. To preserve this artwork may require:

- creation and maintenance of replica digital hardware over time (e.g. CRT monitors or input devices such as motion sensors, or game controllers)
- preservation of original software environments and supporting documentation (e.g. a Mac OS 9 installation running custom software code on it)
- access to and support of emulation tools to emulate old hardware to run the original software on (e.g. a PowerMac G3 emulator).

Provided this is known in advance, and funds are allocated to account for this at point of ingest for the period the artwork is to be preserved, then YUL will not object to accepting the content for preservation.

Supporting policy 3

YUL will use the information in the depositor agreement to decide:

- What preservation approaches can be used to preserve the content
- What storage technology to use to preserve the bits of the content
- The acceptable level of risk of loss of the content
- What content in the files submitted as part of a digital object needs to be preserved and what content can be discarded if necessary (e.g. through a content migration process)
- Which dependencies of the object need to be maintained over time and which can be discarded or replaced
- What access mechanisms should be provided by default
- Security and availability restrictions and allowances

Supports:
Principles 4, 5, 6 and 7

Rationale:
Decisions about what content to acquire/select and therefore preserve are not covered by this policy framework. However YUL cannot be expected to commit to providing preservation services for these without having a clear understanding what that will entail. Agreeing on services to be provided is essential in order to:

- Manage depositor expectations
- Ensure adequate funding is available to support the preservation activities for the full preservation period
- To be OAIS compliant
- Enable efficient approaches to be used that do not require expensive manual intervention at the single-object level

Examples of application:
a. A depositor of digital rare books may agree to have migration processes applied to them that change the look and feel of the interaction environment used to access the books, provided those processes maintain all of the textual content, formatting and visual elements of the original digital objects. Without an agreement in place at point of deposit (when the content is well understood) it may be impossible to make this judgment in the future without a great degree of expensive manual intervention.
b. A depositor of interactive digital objects may not agree to have migration processes applied to them if those processes change the look and feel of the interaction environment, or if some of the content of the objects is provided as part of the interaction environment. In this case the only agreed preservation approaches may be to use hardware emulation or original/replica hardware to preserve the content. Without an agreement in place at point of deposit (when the content is well understood) it may be impossible to make this judgment in the future without a great degree of expensive manual intervention.

Supporting policy 4
YUL will ensure access to content migration or format normalization tools and make available support for undertaking content migration.

Supports:
Principles 3 and 6

Rationale:

As technology changes digital content will become inaccessible on modern computing systems. In order to ensure that content can still be accessed (and reused where appropriate) it may be necessary to move content that is dependent on obsolete technologies into new systems and storage mechanisms that are compatible with modern computing systems.

Migration tools may vary in cost and usability. In order to ensure YUL does not over commit its resources this policy does not commit YUL to either running these tools or supporting them over time. Instead YUL is only committed to ensuring such tools are accessible and providing support to *use* the tools when needed.

Supporting policy 5

YUL will ensure access to hardware and software dependencies of digital objects and emulation or virtualization tools by:

- Providing access to original digital hardware or replica digital hardware to interact with content that depends on it
- Providing access to available hardware emulation tools that enable the emulation of original digital hardware for use in interacting with content that depends on it
- Preserving, or providing access to preserved software (applications and operating systems), and pre-configured software environments, for use in interacting with digital content that depends on them
- Assisting users in the use of these tools and services
- Preserving or providing access to preserved technical documentation to support the use of hardware and software dependencies of digital objects

Supports:

Principles 3 and 6

Rationale:

As technology changes digital content will become inaccessible on modern computing systems. Migration tools are often limited in how much content they can preserve and can introduce unacceptable changes to content over time. Emulation tools and the use of original hardware can prevent such loss to a much larger degree. For these reasons provision of access to emulation tools and access to original hardware will be essential to ensuring the preservation of many digital objects.

In order to ensure YUL does not over commit its resources this policy does not commit YUL to either running Emulation tools or maintaining original hardware over time.. Instead YUL is only committed to ensuring such tools and dependencies are accessible and providing support to *use* the tools and dependencies when needed. If a software dependency is inexorably intertwined with a digital object such that it is essential to preserving the content over time then this dependency should be deposited along with the other digital files at point of ingest.

Appendix 1: Glossary of terms

Bit
The fundamental unit of digital information storage, which can have a binary value of either 1 or 0.

Bitstream
A sequence of *bytes*, which has meaningful common properties for the purposes of preservation. A bitstream may be a *file* or a component of a file.

Byte
A unit of digital information and measure of data volume, normally equivalent to eight *bits*.

Content Migration
Content migration is the process of transferring content between storage types, formats, or computer systems.

Data object
A technological instantiation of an *information object*, composed of one or more *bitstreams* and dependent on a specific technical environment to provide access

Digital Content
The conceptual object of preservation. Digital Content is realized as meaningful information by interpreting a *data object* through its associated *representation information.*

See also: “Information Object”.

Digital object
A digital object is an entity which is the target for preservation. Depending on context, it may refer to an *information object* or a *data object*.

Digital Preservation Service (DPS)
A combination of people, processes and technologies, which together provide the means to preserve *digital objects*.

File
A *bitstream* which is managed by a file system as a single, named entity.

Hardware Dependency
A hardware dependency is a relationship between a piece of hardware and something else (e.g. other hardware, software, or a document file) in which the other thing requires the hardware in order for one or more of its features to be interacted with.

Hardware Emulation

Hardware emulation is the process of imitating the behavior of one or more pieces of hardware with another piece of hardware, typically a special purpose emulation software system.

Information Object

The conceptual object of preservation. An information object is realized as meaningful information by interpreting a *data object* through its associated *representation information.*

See also: “Digital Content”.

Representation Information

The set of information required to interpret a *data object* as a meaningful *information object*, or a component of a technical environment that supports interpretation of that object (such as a software tool or hardware platform).

Software Dependency

A software dependency is a relationship between a piece of software and something else (e.g. hardware, other software, or a document file) where the other thing requires the software in order for one or more of its features to be interacted with.

Appendix 2: Sources reviewed and referenced

1. In developing this policy the policy development group reviewed a number of policies from peer institutions around the world. Links to these policies are available here: https://web.archive.org/web/20140901013628/http://wiki.opf-labs.org/display/SP/Published+Preservation+Policies
2. The policy development group also reviewed YUL's previous (never formally approved) policy that is available here: https://web.archive.org/web/20140822184322/http://www.library.yale.edu/iac/DPC/revpolicy2-19-07.pdf

Appendix 3: Policy Development Group Members

This policy was developed by a policy development group consisting of:

Euan Cochrane, Digital Preservation Manager, YUL

Daniel Dollar, Director of Collections Development, YUL

Michael Dula, Chief Technology Officer, YUL

Kevin Glick, Head of Digital Information Systems & University Archives, Manuscripts & Archives, YUL

Ellen Hammond, Director of International Collections and Research Support, YUL

Michelle Hudson, Science and Social Science Data Librarian, CSSSI, YUL

Peter Leonard, Librarian for Digital Humanities Research, YUL

Francesca Livermore, Arts-Area Digital Librarian, Arts Library, YUL

Robert Klingenberger, Head of Digital Reformatting & Vendor Services, YUL

Youn Noh, Digital Information Research Specialist, YUL

Gabriela Redwine, Digital Archivist, Beinecke Lib Manuscripts, YUL

Appendix 4: Example Structured Data Digital Preservation and Access Plan

The following is an example digital preservation and access plan that is intended to give an idea of what the supporting digital preservation and access strategies may look like. These plans will further detail the specific commitments Yale may make to preserving particular types of digital objects and making accessible content from those objects. This example is not current YUL policy.

Structured Data

Preservation Strategy for Structured data

YUL commits to using the following techniques to preserve structured data:

1. Preserve received files in original formats
2. Create flat-file versions of all data tables where possible
3. Preserve metadata as received
4. Ensure interaction/analysis software is available for emulation

Access enablement strategy for structured data

1. YUL will attempt to convert all structured data into SIARD format and maintain this as the access version of the received data. This conversion process may introduce undocumented change to the content but will enable a wider range of users to be able to interact with the data using common tools and infrastructure.

2. YUL will attempt to have the received metadata converted into the Data Documentation Initiative (DDI) format to enable standard discovery tools to be used across all preserved data.