United Nations
Educational, Scientific and
Cultural Organization

Memory of
the World

The UNESCO/PERSIST Guidelines for the selection of digital heritage for long-term preservation

By the UNESCO/PERSIST Content Task Force

March 2016

Authors:

Sarah CC Choy (Legislative Council Archives, Hong Kong)
Nicholas Crofts (ICOM CIDOC)
Robert Fisher (Library and Archives Canada)
Ngian Lek Choh (National Library Board of Singapore)
Susanne Nickel (Eskilstuna stadsmuseum)
Clément Oury (ISSN – International Centre)
Katarzyna Ślaska (National Library of Poland)

Chair of the Content Task Force: Ingrid Parent (University of British Columbia)
Secretary of the Content Task Force: Julia Brungs (International Federation of Library Associations and Institutions - IFLA)

Endorsed by:

ica international council on archives

IFLA

ICOM international council of museums

United Nations
Educational, Scientific and
Cultural Organization

Memory of
the World

Contents

United Nations Educational, Scientific and Cultural Organization

Memory of the World

Introduction

Heritage institutions – libraries, archives, and museums – traditionally bear the responsibility of preserving the intellectual and cultural resources produced by all of society. This important mission is now in jeopardy around the world due to the sheer volume of information which is created and shared every day in digital form. Digital technology, in dramatically easing the creation and distribution of content, has generated exponential growth in the production of digital information. The digital universe is doubling in size every two years and will grow tenfold between 2013 and 2020.[1] Preserving this vast output is difficult, not just for its extent, but because much of it is ephemeral. Digital information does not have the same longevity as physical objects, documents, and books, which often will survive for centuries. Digital file formats, storage media, and systems are ever evolving, jeopardizing the future readability and integrity of digital heritage over much shorter timeframes than does the deterioration of paper and physical objects, and its availability for capture is fleeting. The survival of digital heritage is much less assured than its traditional counterparts in our collections. Identification of significant digital heritage and early intervention are essential to ensuring its long-term preservation.

To assist heritage institutions in undertaking this vital task, the UNESCO/PERSIST (Platform to Enhance the Sustainability of the Information Society Transglobally) Project has created these *Guidelines on the selection of digital heritage for long-term preservation*. PERSIST arose out of the Memory of the World Conference in Vancouver, Canada (September 2012), which issued the *UBC/UNESCO Vancouver Declaration, The Memory of the World in the Digital Age: Digitization and Preservation*, a call for action to preserve the world's digital heritage before it is too late. In response, PERSIST was launched at an international conference in The Hague, The Netherlands (5 and 6 December 2013), as a collaborative venture of UNESCO, the International Federation of Library Associations and Institutions (IFLA), the International Council on Archives (ICA), and other partners. PERSIST is organized into three task forces (policy, technology, and content) each addressing different challenges to long-term digital preservation. These Guidelines have been prepared by the Content Task Force for discussion by UNESCO and the heritage community.

The aim of the Guidelines is to provide an overarching starting point for libraries, archives, museums and other heritage institutions when drafting their own policies on the selection of digital heritage for long-term sustainable digital preservation. Existing institutional policies may be assessed against the Guidelines and revised if required. The Guidelines address a diverse audience. As digital heritage may differ widely between communities, regions and countries, its preservation requires engagement and cooperation of both the public and the private sectors, as well as content creators. While public institutions may have the primary legal

[1] EMC, "Digital Universe Invaded by Sensors", press release, 7th Annual Study on the Digital Universe, 9 April 2014.

responsibility for managing heritage collections, the private sector must also face the challenge of preserving and ensuring access to its digital information. It may be for regulatory requirements and responsibilities to shareholders, but there is also a corporate social responsibility to support the sustainable growth and development of their communities, regions, and the world by retaining valuable digital heritage and making it accessible for future generations. The Guidelines acknowledge that libraries, archives, museums and other related organizations are different in mandates, operation and methods used in collecting and managing their holdings. However, with the development of information technology and the growing user demand and expectation of fast and integrated access across library, archives and museum collections, all heritage institutions and information providers face similar challenges in the selection and preservation of digital materials. The Guidelines thus target institutions, professionals and administrators on every level and in every region of the world in order to review existing material for selection, highlight important issues, and offer guidance when drafting institutional policies. The Content Task Force also acknowledges that there is an ethical dimension to the issue of selecting heritage for preservation but will not explore this in detail at this stage.

The role of National institutions and networks

National institutions should play a vital role providing leadership to their heritage communities on issues of digital selection and preservation. In many countries, designated national institutions have legislation related to the legal deposit of cultural heritage for example of published material and/or the deposit of official records of their governments. Existing legislation, where it covers only physical material, should be adapted to cover digital material.

Acquiring and collecting digital heritage, through diverse channels and platforms, requires significant effort and resources. National web domains for example can range from thousands to millions of websites; on which millions to billions of files are posted, updated or deleted daily. Even offline there is a great amount of digital heritage worth being kept for the future (for example, raw data from research, government records, private digital files of organizations and individuals). The scope of this challenge makes it natural for large, national institutions to take a leading role, whether by establishing proper policies and systems to collect and manage digital material or leading collaborative networks in the adoption of shared selection and preservation models. We recommend that national institutions and networks develop national selection strategies in consultation with their heritage communities. It is critically important that other partners, such as governments, academic institutions, research centres, non-profit and private sector organisations take on digital stewardship roles and responsibilities. For example, national institutions could jointly define the standards and processes of collecting, organising and preserving digital information while various other partners can complement these

standards and processes. International organizations (such as ICA[2], IFLA[3] and ICOM[4]) should be involved in this work as preservation of digital heritage is a global challenge.

The impact of the Legal environment on selection

The legal environment has important implications for the selection and preservation of digital heritage. International and national laws, which vary widely, regulate the dissemination, duplication, access and use of digital heritage, but the Internet transcends territorial boundaries, often making it difficult to identify rights-holders and which or whose laws apply. Government laws and regulations respecting the protection of intellectual property rights, privacy of persons, confidentiality of "state secrets", and public access to information all affect what and how digital heritage can be preserved, and if and when it can be made accessible to the public. Copyright legislation, save for specific exceptions and limitations may prohibit the making of copies and raises new issues in a digital environment in which duplication may be necessary for long-term preservation. Moreover, digital materials are often software dependent for search and retrieval, and this software may also be protected by copyright. Some countries have enacted laws to prevent circumvention of technological protection measures used to prevent copying and redistribution, which could inhibit preservation and impede future access to digital heritage. Legal impediments to preserving or making accessible digital heritage will weigh heavily on selection decisions.

There is a strong risk that the restrictive legal environment will negatively impact the long-term survival of important digital heritage. Cooperation or specific agreements with rights-holders may be the only way to protect and preserve certain types of heritage (e.g. the Library of Congress/Twitter agreement). We recommend the adoption of international and domestic legislation aimed at overcoming barriers to the selection and preservation of digital heritage for public access.

Thinking globally: Digital selection issues for heritage institutions

The challenge of long-term preservation in the digital age requires a rethinking of how heritage institutions identify significance and assess value. The proliferation and abundance of digital heritage and information, and the ephemeral nature of much of it, means that heritage institutions must be proactive to identify digital heritage and information for long-term preservation before it is lost. What should be preserved for the long-term benefit of humanity? Traditional forms of cultural heritage – books, periodicals, government records, private correspondence, personal diaries, maps, photographs, film, sound recordings, artefacts, and works of art, to name a few – now have digital equivalents, which often fit well within our

[2] International Council on Archives: http://www.ica.org/3/homepage/home.html
[3] International Federation of Library Associations and Institutions: http://www.ifla.org/
[4] International Council of Museums: http://icom.museum/

existing practices and mandates. But the digital environment has created new forms of expression, ranging from web pages and interactive social media sites to private research databases and online gaming environments that blur boundaries and lines of responsibility and challenge past approaches to collecting.

Existing mandates and collecting policies often do not cover these new forms of digital heritage. Our collective neglect of these new forms raises the risk of creating large gaps in our cultural heritage for future generations. As an example, even though the value of individual posts on blogs or social media may be marginal, collectively they constitute a unique record of contemporary society, the discussions, thoughts, and achievements of billions of individuals. If preserved, they will represent an incomparable source of knowledge for future generations. Focusing only on the "best" part of this output would introduce biases and prevent any analysis of contemporary digital creation as part of a whole. But few, if any, heritage institutions have the resources, and in some cases the rights, to harvest and preserve *en bloc* this social output in digital form. This is the paradox of selection in the digital age. Selection is as essential, as it is economically and technically impossible, and often legally prohibited, to collect all current digital heritage. Selecting for long-term preservation will thus be a critical function of heritage institutions in the digital age.

Though some of the traditional boundaries between libraries, archives, and museums are blurring in the digital age, these different communities still have core interests common to each preserve heritage. As a result, particular issues are of more or less relevance in each community:

Libraries will face the challenge of digital selection with respect to e-publications, harvesting web sites, and proprietary content in social media sites like Facebook and YouTube. National libraries striving to build a comprehensive collection, often with a strong tradition of legal deposit, will have to adopt selection for more ephemeral publications in digital form. In the past, selection was done, in effect, by publishers who "curated" creative output through editorial choices that determined what would be published. In the democratized world of self-publishing and e-books, national libraries will have to modify past comprehensive approaches and adopt criteria to select for long-term preservation. But not all libraries are "memory" institutions, many only have a mandate to support its community's contemporary user or research needs. Selecting digital heritage for long-term preservation for such institutions may focus primarily on evaluating publications already in their collection, originally acquired for short-term use, rather than assessing new publications for acquisition.

Museums with strong and well-developed collections of physical material culture generally acquire for permanent preservation and make collections development decisions in this context. This material culture is now increasingly digital (machines which are driven by computer software for example, born digital works of art, digital documentation of archaeological sites, etc.). Research information related to the physical holdings of museums

is also increasingly digital. Digital heritage in museums thus can be divided into the following categories: born digital items in the collection, digital or digitized information about the collection and digital representations of physical artefacts in the collection (digital images or 3D scans for example). Due to this categorization, museums normally should prioritize the first and second categories for long-term preservation. The second and third categories also include institution-generated administrative records.

The importance of metadata (information about physical and digital heritage) to museums cannot be overstated. This metadata includes contextual information created about the physical and digital heritage before it enters the museum and contextual information created during its life in the museum. The principle of provenance is also important to museums.

Archives traditionally also acquire original or unique records for permanent preservation and have relied on the passage of time between their creation and their acquisition by an archive to lend historical perspective in making selection decisions. Rapid obsolescence in digital formats, storage media, and system hardware and software systems, however, is collapsing the window of opportunity of selection, with the increased risk that records are lost or that those acquired early might not have yet "proved" their significance over time. Whereas, the published material acquired by libraries is distributed in multiple copies or widely available online, digital archival records generally reside offline in private systems, servers, and networks not accessible to the public. Access and selection must be negotiated with the owners, often even with government institutions where a legislated mandate exists for records transfer. Archives focus on the importance of authenticity, provenance, and context in the appraisal of archival records for acquisition, but the ease of manipulation and duplication of digital records makes it more difficult to evaluate these factors in selection. The legal environment often dictates what digital information must be acquired by an archive and how, or if, it can be made accessible for public access and research.

These factors which influence the selection environment are not necessarily exclusive to each of the library, museum, and archives communities. Indeed, there will almost certainly be some overlap. But reviewing the diversity of our communities helps illuminate the range of issues to be faced by institutions in identifying and selecting heritage for long-term preservation.

Acting locally 1: Strategies for collecting digital heritage

Heritage institutions must adapt their existing approaches to the digital environment. It is likely that most will adapt one or a mix of the following strategies or approaches in fulfilling their collecting mandates:

Comprehensive collection

Comprehensive collecting is used to acquire all of the material produced on a given subject area, time period, or geographic region. This approach requires significant institutional resources, or a narrow focus. Legal deposit of publications is perhaps the most familiar comprehensive approach, in which a national library attempts to collect the entire publishing output of the nation through a legal requirement for publishers to deposit copies of each publication they produce. But museums too might seek to gather all works created in a particular time period or an archive might attempt to collect everything related to an influential public figure.

Representative sampling

Sampling is another approach used to identify material for long-term preservation. It is often used when an institution does not have the resources or capacity to collect comprehensively and differentiating the material by specific selection criteria is problematic. In these cases, sampling offers a means of capturing a representative picture, making selection and preservation more manageable and less resource-intensive. For example, a national library might perform regular crawls of a whole national web domain (such as .dk or .fr) to preserve at different times a representative portrait of its national presence online. An archive might select government case files by using a sampling methodology, such as keeping only the most-documented cases, or those from a given year, or that begin with a given letter of the alphabet.

Selection

An approach based on selection is used when heritage professionals – archivists, librarians and curators – identify material for addition to their collections based on specific criteria. These criteria can vary widely depending on the type of institution, its collecting mandate, its resources, and the type and extent of material available for acquisition. The selection criteria are generally expressed or defined in a collecting or acquisition policy, and may be based on the following criteria (which may also be combined):

- *Subject/Topic*. An institution will focus on and attempt to document one or more subject areas. For example, all websites dedicated to a specific painter or locality, or a web crawl to document a specific event like a political election or arts festival.
- *Creator/Provenance*. An institution will focus on particular creators of heritage or provenance. For example, an archive might acquire the digital records of authors of a particular region; or a museum might collect works of the artists of a particular movement.
- *Type/Format*. An institution might collect by the type or format of content (e.g. digital photography, music recordings, film, video games and others).

In some cases, institutions may choose to capture all the digital heritage material now and apply selection criteria at a later date, in a form of deferred selection.

Acting locally 2: Developing selection criteria for a single institution

How should individual libraries, archives, and museums select, identify, and prioritize digital heritage before it is lost? Existing institutional missions, mandates, and collection development policies for physical collections, in most cases, will provide the starting point and essential guidance for assessing and selecting digital heritage. These should now be adapted to embrace new forms of digital expression.

Evaluating and assessing digital heritage should be based on many of the same principles that underlie traditional selection, but must also now consider new issues of long-term accessibility, use, and preservation in making selection decisions. An institution should answer these questions by evaluating the relative **significance** of the digital heritage to its mandate and public; by assessing its **sustainability**, that is, its capacity to preserve it for long-term access and use; and by considering its **availability** in other heritage institutions, that is, its prospects for preservation elsewhere, and what is the most appropriate institution to preserve it. The concepts of significance and sustainability in this environment must be evaluated in light of institutional mandates and resources. Availability looks outward to other institutions in the heritage community to assess the level of risk to the digital heritage's long-term survival. Particular attention must be paid to heritage that is at risk of being lost over the short to medium term. Digital heritage with broader human significance, such as the expressions of indigenous peoples in digital form, must be identified and preserved before it is at risk.

Bearing in mind that every memory institution is unique in its mandate, collections policy, and resources, we offer below a series of steps and questions that can frame digital selection decisions. This approach can be scaled to the individual needs of institutions which are diverse in purpose and size. Even if not adopted, these steps can form the point of departure for an institutional discussion about the selection of digital heritage for long-term preservation.

Decision Tree for Selection in an individual Institution

This approach consists of four steps, posed in a series of questions, to support consistent, evidence-based evaluation:

Step 1: Identification

Identify the material to be acquired or evaluated. What is its title, creator, provenance, extent, condition? What kind and which amount of metadata is available? Define the parameters of the project, if appropriate. Is a simple yes/no decision sufficient, or is relative evaluation (high, medium, low, or even a numerical valuing?) required to compare it against other material?

Document/record your decision made in this step (and how you got it) and keep the record. Ensure the record is kept up-to-date and accessible.

Step 2: Legal framework

Does the institution have a legal obligation to preserve the material? Does the institution's mandate or policies on digital preservation and collections development require the preservation of this material?

- If yes, preserve. A positive decision is confirmed, no further steps are required.

Document/record your decision made in this step (and how you got it) and keep the record. Ensure the record is kept up-to-date and accessible.

Step 3: Application of Selection Criteria

If the institution does not have a legal obligation to acquire the digital heritage, it can assess the material using three selection criteria: significance, sustainability, and availability to identify if it should be preserved. These criteria should be assessed in whichever order is most efficient or effective for the institution, generally beginning with the easiest criteria to evaluate and proceeding until a final decision is reached.

3(a) Significance

Does the long-term value of the digital heritage justify its preservation? Does it have significant social, cultural, historical, or artistic value for the community served by the institution? Does it have significant information, content, use, exhibitions, or research value? How closely do these values support and align with the institutional mission and mandate? Does the material's provenance, rarity, uniqueness, or representativeness affect these values? How will the institution's stakeholders (clients, sponsors, society) be affected if this digital heritage is <u>not</u> preserved?

- If the digital heritage is significant in relation to the institution's mandate, consider preserving.
- Document/record your decision made in this step (and how you got it) and keep the record. Ensure the record is kept up-to-date and accessible.

3(b) Sustainability

Does the institution have sufficient budget and resources to preserve the digital heritage material over the long-term? Does the institution have technical capacity to read, migrate and preserve the digital heritage? Are specific rights required to transfer or migrate the material to different file formats and physical carriers? Is sufficient metadata available to access and preserve the digital heritage? Can the institution make it accessible for research, exhibition, or other use to meet its public's expectations?

- If the answers are mainly no, consider not preserving.
- Document/record your decision made in this step (and how you got it) and keep the record. Ensure the record is kept up-to-date and accessible.

3(c): Availability

Consider the general availability of the digital heritage in other institutions in the heritage community or network. Is this institution the only one preserving this material, or are exact duplicate copies held by other institutions? Is it rare or unique, or is it widely duplicated? Where will it receive the most use or be of the most benefit to the public? Is it at risk at other institutions? Is this institution the most appropriate or best-placed to preserve and make accessible this digital heritage? Keep in mind that a certain amount of redundancy is necessary to secure proper preservation of digital heritage[5].

- If the answer is no, perhaps the digital heritage should be preserved by another institution. But the other criteria should still be weighed against this answer.
- Document/record your decision made in this step (and how you got it) and keep the record. Ensure the record is kept up-to-date and accessible.

Step 4: Decision

Compile and review all the records made during the process and make a decision based on the results from steps 1 to 3. Document and record the rationale and justification for the evaluation or decision. This is vital, both for governance and to capture important information for potential reappraisal in the future. Prepare a written statement of the digital heritage's significance and its technical preservation issues, incorporating the answers to the questions in steps 1 to 3. The arguments behind the decision are often more important than the evaluation itself. A standard institutional evaluation form or appraisal document should be created to capture these arguments and be a record of the decision. Document and record all data produced during the process and keep it accessible. The documentation of the rationale and justification for the evaluation or decision is vital, both for governance and to capture important information for potential reappraisal in the future.

This approach is flexible, not every question will be applicable to each institution. Nor is the order of the criteria set in stone; for example, in some cases step 3(c) might be better assessed before steps 3(a) and 3(b), particularly if it is clear that another institution is more appropriate. But following this approach should support heritage institutions in making better decisions in selecting digital material for long-term preservation.

Conclusion

The long-term preservation of digital heritage is perhaps the most daunting challenge facing heritage institutions today. Developing and implementing selection criteria and collecting

[5] See Redundancy in Appendix 1

policies is the first step to ensuring that vital heritage material is preserved for the benefit of current and future generations.

Appendix 1: Management of long-term digital preservation and metadata

Selection of digital heritage is closely connected with issues related to long-term preservation and access. Some losses of important digital heritage may be unavoidable, but the risk can be mitigated by following best practices in digital preservation, including redundancy, active management, and metadata management.

Redundancy
Important digital heritage, including master files with associated metadata, should exist in multiple copies that are stored in at least two different physical locations. Heritage institutions can use a mix of on-site, off-site, and distributed cloud-based storage, but digital originals should be backed-up in at least one other location. Storage sites should be chosen to diminish the risk of loss due to natural or man-made disasters and economic or political crisis.

Active management
Heritage institutions should actively manage their digital heritage assets to ensure their accessibility and integrity over the long term. Digital heritage should be preserved in open and well-documented file formats, without encryption, and at least lossless compression. This method is strongly recommended for heritage institutions in the active management of digital objects. Storage should use two or more different types of storage media, ranging from institutional servers to portable media (e.g. magnetic disk, optical media, or magnetic tape).

Systems failure over the long-term can cause vital information loss to stored digital heritage. Many institutions guard against this by using a periodic media refresh, comprising reading in the digital data, checking for errors using error correction techniques, and rewriting on new media. To avoid software failure digital data owners often use standards-based protocols for access to data storage, where different storage sites are running different implementations of the storage software. Therefore the integrity and reliability of data does not depend on the integrity and reliability of any single implementation.

Metadata management
Metadata is usually defined as “data about data”, which though accurate, is not very precise. In heritage institutions, the required metadata should be considered as any information (in digital or physical form) that is essential to ensure that the digital material being preserved is, and remains, accessible, intelligible and usable over time. Metadata provides the institution with the information required to access and preserve digital heritage into the future.

Heritage institutions generally preserve three key types of metadata associated with digital heritage that are crucial to long-term preservation:

- Structural (required for the technical capacity to read digital content)

- Descriptive (containing bibliographic, archival, or museum contextual information, which can be system-generated or created by heritage professionals, content creators, and/or users)
- Administrative (documenting the management of a digital object while in its collection).

If the digital heritage is the “content”, then the metadata provides the “context”.

There are five basic functional requirements for digital metadata:

Identification: The metadata must allow each digital object to be identified uniquely and unambiguously. This will usually require a globally unique identifier to be assigned to each item.

Location: The metadata must allow each digital object to be located and retrieved. The long-term validity of this location data needs to be ensured so that items are not lost when systems are migrated or updated.

Description: A description of digital object is needed to facilitate recall and interpretation. Descriptive metadata falls into two categories: data about content and about context. Data about the content of an item can often be re-created by examination and consultation. It is nonetheless useful as a finding aid for resource discovery. Data about context, where, when and by whom an item was created, what it was used for, its place in relation to a general corpus of material, is much more difficult to recreate once lost.

Readability: Metadata about the structure, format and encoding of digital objects is needed to ensure that they remain legible over time. This functional requirement is particularly important for digital objects as they cannot be read without mediating technology. This metadata should identify the relevant standards and provide references to the technical documentation, authority files and other related material needed for a complete rendering of the digital resource. Care needs to be taken to ensure that all the multiple layers of a digital object can be interpreted: from the encapsulating file format to the representation and codification of the data itself.

Rights management: rights, conditions of use and restrictions applicable to each digital item need to be recorded in the metadata. This metadata should identify the applicable laws and conventions and provide references to relevant legal documentation, contracts, etc. as well as the rights holders.

Storage of metadata

Many digital file formats allow metadata to be embedded within the file itself. This has the advantage of ensuring that the data and metadata remain linked. However, metadata also

needs to be stored independently from the digital resource that it describes; this is essential to meet the functional requirements set out above. An encoded digital item, for example, cannot be read if the code is only to be found embedded in the item itself.

Metametadata

Some data about the source of the metadata and how it was compiled is needed to establish its reliability and authenticity; when was the metadata compiled and by whom? Was the metadata harvested automatically or manually? What tools and techniques were used? For future retrieval and understanding of the digital information, contextualisation is essential.

Appendix 2: Definitions of terms

The definitions are taken from sources listed in Appendix 3.

Authenticity of digital heritage refers to the trustworthiness of a record or an item, i.e. the quality of being what it purports to be, either as an original object or as a reliable copy derived by fully documented processes from an original.

Content selection refers to the process of deciding which items or materials are worthy for preservation through evaluating their significance and lasting cultural, scientific, evidential or other value as well as the feasibility of preservation and access based on defined principles, policies, procedures and standards.

Digital heritage means heritage which is made up of computer-based materials, whether born digital or digitized from other formats, which emanates from different communities, industries, sectors and regions and requires active preservation approaches to ensure its authenticity, accessibility and usability through time.

Heritage refers to legacy from the past, what we live with today, and what should be passed from generation to generation because of its significance and value.

Metadata means information that describes, explains, locates or otherwise makes it easier to understand, retrieve, use, manage, control or preserve an item or information resource through time.

Appendix 3: References

The following list of references is a selection and is in no way meant to be exhaustive:

2015 National Agenda for Digital Stewardship, 2014

CENL/FEP Statement on the Implementation of (Statutory and Voluntary) Deposit Schemes for Non-Print Publications, FEP and CENL, 2012

Digital Legal Deposit, an IPA [International Publishers Association] Special Report, 2014

Bonnel, Sylvie and Oury, Clément, "Selecting websites in an encyclopaedic national library: a shared collection policy for internet legal deposit at the BnF", 2014

Caroline Brazier, British Library, "http://library.ifla.org/222/1/198-brazier-en.pdf.", 2013

British Library, British Library digital preservation strategy 2013 – 2016, 2013.

Bailey, Catherine A. "Past Imperfect? Reflections on the Evolution of Canadian Federal Government Records Appraisal", Archivaria 75, 2013.

Canada, Canadian Heritage Information Network, Digital Preservation Toolkit, 2013.

Kenneth Crews, "Study on Copyright Limitations and Exceptions for Libraries and Archives", Standing Committee on Copyright and Related Rights, World Intellectual Property Organization, Twenty-ninth Session, Geneva, December 8 to 12, 2014

Cunningham, Adrian, "Recent Developments in Standards for Archival Description and Metadata" IASA Journal no 16, 2000.

Digital Preservation Coalition, Digital Preservation Handbook Definitions and Concepts, 2008

Duff, Wendy and van Ballegooie, Marlene, Archival Metadata, DCC Digital Curation Manual, 2006.

Harvey, R. "Appraisal and selection" DCC Digital Curation Manual (2006) (which aims to offer a generic selection framework)

IFLA, IFLA Statement on Legal Deposit, 2011

InterPARES 2 Project, "Preserver Guidelines - Preserving Digital Records: Guidelines for Organizations", 2007

InterPARES 2 Project, InterPARES 2 Project Dictionary, 2015.

ISAD(G), "General International Standard Archival Description 2nd Edition", Ottawa: ICA, 2000.

ISO 15489-1 Information and documentation – Records management – Part 1: General, ISO 2001.

ISO, ISO 16175-1:2010(E) Information and documentation – Principles and functional requirements for records in digital office environments – Part 1: Overview and statements of principles, 2010.

ISO, Information and documentation — Statistics and Quality Indicators for Web Archiving , 2013 (draft available).

Koerbin, Paul, Colin Webb, and David Pearson, "Oh, you wanted us to preserve that?!'

Statements of Preservation Intent for the National Library of Australia's Digital Collections", D-Lib Magazine, 2013, Volume 19, Number 1/2

Lariviere, Jules, "Guideline for Legal Deposit Legislation", UNESCO, 2000

Library of Congress Digital Preservation Webpage

NYARC Reframing Collections for a Digital Age Report from Consultant No. 2, 2012, USA

Netherlands, Cultural Heritage Agency, "Assessing Museum Collections: Collection Valuation in Six Steps", 2014.

Niu, Jinfang "Appraisal and Selection for Digital Curation", International Journal of Digital Curation, 2014 Vol. 9 Issue 2, 65-82.

Niu, Jinfang, "An Overview of Web Archiving", D-Lib Magazine, 2012, Volume 18, Number 3/4.

Noonan, D.W. "Digital preservation policy framework: A case study', Educause Review 2014

Paradigm Project, Workbook on Digital Private Papers, United Kingdom, 2005-2007

Sheridan, John, Amanda Spencer, and David Thomas, "UK Government Web Continuity: Persisting Access through Aligning Infrastructures", The International Journal of Digital Curation Issue 1, Volume 4, 2009.

Society of American Archivists, A Glossary of Archival and Records Terminology, 2005

The State of Digital Preservation: An International Perspective, Conference proceedings, Washington D.C., April 24-25, 2002. Washington, DC, 2002

UBC/UNESCO Vancouver Declaration The Memory of the World in the Digital Age: Digitization and Preservation, Canada, 2012

UNESCO, Charter on the Preservation of Digital Heritage, 2003

Van der Werf, T. and B., 'The paradox of selection in the digital age', 2014
Helmus, W., "Survey on selection and collecting strategies of born digital heritage - best practices and guidelines", 2015

Verheul, I. "Networking for Digital Preservation: Current Practice in 15 National Libraries", München, 2006