



NB-note 10-06: Policy for long term preservation of digital materials at the Royal Library

1. Version history

Version no.	Date	Initials	History	Status
1.0	15 December 2009	BNH	1.0	Approved
1.1	17 February 2010	BNH	Proof corrections added	Approved

This document is updated and approved every year by the Board of Directors of the Royal Library. Responsibility for revision rests with the administrator with the operational responsibility for the activity 'digital preservation', currently the department head of Digital Preservation in the National Library.

The document is published on the website of the Royal Library in both Danish and English versions so that it is available for general international evaluation of policies in this special area.

Introduction

This document describes the Royal Library's policy for the long term preservation of digital materials in the collections. It begins with a discussion of the challenges that digital materials give rise to both in relation to preservation and to continued provision of access to the materials. Following this the purpose of the policy is described and the activities it covers. Next is a section on



the implementation principles on which the policy is based. The policy note concludes with a vocabulary list and a brief list of references.

The Royal Library has had a preservation policy for its physical collections since 2001. This policy is now supplemented by this policy for the long term preservation of the digital collections.

Large portions of the collections of the library will in future be digital and this creates a series of new challenges for the Royal Library. For the preservation of the physical materials the Royal Library already has storage space, a series of thoroughly tested and approved policies, as well as financing dedicated to this area. A corresponding situation must be established in order to be able to preserve access to the digital materials.

Digital materials confront us with these challenges:

- A new type of material, rapidly growing in the coming years, but with an as yet unknown specific volume
- Digital material cannot, as is the case with books, be left passive until needed, since systems, storage media, data formats and programmes are continually changing. Lack of action – also in the early phases – will therefore with certainty lead to loss of the capability to provide access to the preserved data
- There are no ready-made methods and solutions to deal with these problems, but international efforts in this sector are being made to develop strategies and techniques to manage some of the problems for certain portions of the materials
- Nor are there any ready-made models for what these activities cost, neither annual or collectively over the lifetime of the materials, but international efforts are being made to establish early versions of cost models.
- Legislation regarding storage and access to these types of materials is more restrictive than for physical materials

At present no Danish national policy exists for this area. This policy document is thus restricted exclusively to the Royal Library's own requirements in this area.

The preservation policy for internet materials covered by legal deposit has been established in close cooperation with the State and University Library in Århus, which is its own, separate, preservation policy.

This document is part of a collection of documents, which also includes:

- *"Acquisition and discarding policy for digital collection materials at the Royal Library"*
- *"Strategy for preservation of digital materials at the Royal Library"*, which describes how the library will attempt to carry out the approved policy <document in preparation >

The Royal Library normally stores its data in stable IT systems with backup. Backup is a short term solution to re-establish recently lost data in the same structures. Digital long term preservation is intended to secure data over a very long time span and is not the same as backup.

There is a vocabulary list at the end of this document.



2. Purpose

The broad purpose of this policy at the Royal Library is to introduce risk management in relation to the digital collections which the Royal Library intends to preserve for a long time by

- Ensuring that digital materials in the collections that the Royal Library procures or produces within its collection profile will be accessible for researchers and others in up-to-date form for both the short and the long term (> 100 years)
- Ensuring that the digital preservation processes are carried out in such a way that there continues to be a high degree of credibility regarding the authenticity of the collections of the Royal Library
- Ensuring that the library continuously carries out the necessary investments in infrastructure and knowledge to be able to perform the preservation actions which will ensure the collected effort over a long time span
- Ensuring that there are clear priorities which make it possible to carry out an optimal policy in relation to the Royal Library's budgets

3. Activities covered by the present preservation policy

3.1. The main principles in the Royal Library's long term preservation policy for digital material

The main principles are the following

- All acquisitions to the Royal Library's digital collections intended for long term collection building are to be preserved within the financial resources
- Long term preservation must ensure retrieval and possibilities for access now and/or in the future to the information contained in the collections
- Long term preservation covers as a minimum secure storage of the physical bit sequence
- Implementation of this policy must follow international and open standards as far as possible, in cooperation with international and national partners



3.2. Acquisition of material in the collections covered by digital preservation

Acquisition and discarding of collections covered by digital preservation is carried out according to the document: "*The Royal Library's acquisition and discarding policy for digital collection materials*"

Bibliographic, preservation, access and administrative descriptions (including ownership and access rights) - referred to collectively as metadata – are worked out according to locally described norms, which are as close to international standards and 'best practice' as possible.

3.3. Preservation actions

Digital material selected for preservation will be subject to, in order of priority, the following four preservation actions:

- Preservation of the physical bit sequence
- Preservation of the capability of retrieval through the use of (global) identifiers
- Preservation of information or structures, which link complex materials together
- Preservation of the capability to interpret the physical bit sequence

Until financing is obtained, the extent of digital collections to be preserved is limited mainly to digital safety copies and born-digital materials in the collections, see "*The Royal Library's acquisitions and discarding policy for the digital collection materials*".

Preservation strategies must be established for all the Royal Library's digital collections which have been selected for preservation, see "*Strategy for preservation of digital material at the Royal Library*".

The composition and volume of these collections must be continually monitored and specific preservation action plans must be set up for those (sub) collections which might be threatened.

If, for either technical or economic reasons, it is not possible to preserve all functionality in a digital object in connection with carrying out a preservation action, the preservation of the digital object's intellectual content (information) has as a rule the highest priority.

Correspondingly, the external environment (technological development of relevant tools/standards etc.) must be continually monitored, as well as developments in the requirements of users to functionality.

3.4. Access to preserved digital data in the collections

Access to the preserved digital master copies (collections) is, as a rule, regulated, restrictive and fully registered.



There must be access to the preserved digital collections of the Royal Library for the following purposes:

- On-demand from users for materials located, e.g. via REX (OPAC) interface
- Reproduction of up-to-date use copies
- Administration of the collections by the responsible collection department
- Performance of automated processes in connection with carrying out the preservation policy

4. Implementation principles

4.1. Use of standards

4.1.1. Models and framework

The standard "Reference Model for an Open Archival Information System (OAIS)" is universally used in the international library and archive world as a model of reference in the construction of solutions to deal with the tasks connected to management and preservation of digital objects. The Royal Library also uses this OAI model as our reference model in connection with the description and construction of our solutions.

The Royal Library also chooses an international level for other significant parameters: models for solutions, cost descriptions and certification as a "Trusted Digital Repository".

The Royal Library supports and practices the use of open source, which is especially necessary for this area, since no ready-made solutions yet exist.

4.1.2. Data formats

The Royal Library chooses open, standardised data formats with a reasonable assumption that they will be manageable with one of the preservation strategies that the library chooses to use.

When either practical or economic reasons dictate the use of compression of data, lossless compression algorithms are used.

4.2. Legislation

Digital material is currently subject to the same laws as physical collection materials, though with more restrictions. Insofar as the library chooses to preserve the digital materials in the collections, this is to be done in such a way that the more stringent restrictions, which primarily deal with access, can be assured.



4.3. Technical infrastructure

The Royal Library establishes the necessary technical infrastructure to carry out preservation of its digital collections.

This infrastructure supports the necessary monitoring of systems, media and formats to ensure a pro-active effort.

The infrastructure encompasses tools which both support the decision-making processes as well as monitoring and performing specific preservation actions.

Financing

Preservation of digital collection materials is a new core responsibility for the national library function of the Royal Library. A significant part of the task will be to obtain the necessary continual financing to do the job now and in the future.

4.4. Cooperation with other institutions

Division of labour (primarily with regard to responsibility for choice of strategies for data formats) and operational cooperation between Danish state preservation institutions with the necessary competence should be agreed to as far as possible.

The Royal Library should secure a central place in the international network through strategic cooperative projects. This ensures that the institution is legally, politically and, with regard to knowledge, methods and tools, at the level of the medium-sized leading national libraries in Europe, with whom we normally prefer to compare ourselves in this area.

4.5. Personnel

It must be ensured that the library always has sufficient up-to-date expert personnel to maintain and carry out the preservation policy responsibly, whether it is done internally or outsourced.

5. Vocabulary

By **digital preservation** is understood the processes that ensure ongoing access to digital information despite changes in technologies (storage media, data formats and programmes for access) over time. Access in this connection means both that data is preserved without error, but also that they can be interpreted by future user groups in a perspective of a minimum of 100 years. Digital preservation is a task with many different aspects, where there are no ready-made solutions and answers today. Digital preservation is not the same as a backup. A backup is a short term solution to re-establish recently lost data in the same structures.

There are basically two main activities which both must be performed in order to speak of digital preservation:

- **Bit preservation** must be performed, i.e. the completely correct sequence of '0's and '1's must be preserved for posterity



- A **logical preservation** must be performed, i.e. methods must exist so that files can forever be transformed into something understandable by people. This last can basically be achieved from three different preservation strategies:
 - The **museum** approach, in which an attempt is made to save the original hardware and software in the form of the electronic reader apparatus continually developed for this purpose.
 - **Format conversion (also called format migration)**, where it is ensured that the material is always in a form in which the current reader apparatus / interpreters can read. There is always the risk of losing something in format conversion – typically not anything related to the document’s intellectual content, but e.g. something bound up in the functional, such as look-and-feel. Format conversion is the easiest to do, since the task can start immediately with those formats for which there is a practicable solution.
 - **Emulation**, in which either in software or in hardware ensures an emulation of the original surroundings, so that when running the original programmes on the original data the original experience can be re-created. Emulation is the most difficult to achieve because all the challenges must be met as a whole.

A **digital master copy** is the copy of a digital object which is used as an archival copy. It is typically richer in information than the copies made for other purposes and thus normally also is larger. A master copy usually is the basis for the production of derived formats, e.g. for ordinary dissemination use.

If different formats are used for dissemination and long term preservation, a **dissemination copy**, which is a copy of a digital object only used in a dissemination situation, may be made. Typically it lacks some of the information in the master copy and is often adjusted for a particular dissemination purpose. The lifetime of a dissemination copy will often be far shorter than that of a master copy, since the demands on up-to-date dissemination constantly change.

6. Literature:

Erpanet Guidance: Digital Preservation Policy Tool, September 2003:
<http://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf>

National Library of Australia: Digital Preservation Policy, 3rd Edition (2008)
<http://www.nla.gov.au/policy/digpres.html>

Neil Beagrie, Najla Semple, Peter Williams, and Richard Wright : Digital Preservation Policies Study, 30 October 2008, <http://www.jisc.ac.uk/publications/publications/jiscpolicyfinalreport.aspx>

Reference Model for an Open Archival Information System (OAIS)
<http://public.ccsds.org/publications/archive/650x0b1.pdf>



NB-note 10-07: Section of the acquisition and discarding policy for digital collection materials at the Royal Library

1. Version history

Version no.	Date	Initials	History	Status
1.0	15.December 2009	BNH	First accepted version	Approved
1.1	17.February 2010	BNH	Proof corrections added	Approved

This document contains sections which belong in the relevant parts of the Royal Library's acquisitions policy in other departments.

At this point it has not been decided whether the document will remain a separate, independent document or whether it should be incorporated in the existing acquisitions policies.

This document in its current form does not take into consideration whether the collection materials belong to the National Library or to Copenhagen University Library.

Responsibility for revision of this document belongs to the vice director for the National Library and is to be coordinated with the revision of the document "*Policy for long term preservation of digital materials at the Royal Library*".



2. Introduction

An increasing proportion of the library's collections are digital and this creates a series of new challenges for the library.

One of these challenges is to obtain financial resources to manage and preserve the digital collections.

Digital collections, like other collections, involve expenses both for storage, access and administration, as well as operating expenses and wages.

Until separate financing for this purpose is provided, the expenses are covered by re-allocation of resources. This means that the resources are severely limited and it is necessary to ensure strict control and clear priorities regarding which materials are acquired for the collections and under what conditions.

This document is a part of a collection of documents, which also includes:

- *"Policy for long term preservation of digital materials at the Royal Library"*
- *"Strategy for preservation of digital materials at the Royal Library"*, which describes how the library will attempt to carry out the approved policy <document in preparation >

3. Purpose

The purpose of this document is to ensure that there exist clear priorities, which make it possible to carry out an optimal policy within the budgets of the Royal Library with regard to which materials should be acquired for the digital collections, and which there should be access to in the short or the long term.

4. Activities covered by this acquisitions policy

The policy covers the following parts of the digital life cycle:

- Selection of materials to be covered by the acquisition policy
- Selection of and setting priorities for materials, which are not only covered by the acquisition policy, but in addition are included in digital preservation
- Acquisition and incorporation into the digital collections of the Royal Library
- Discarding policy



4.1. Materials covered by this acquisitions policy

Selection of the digital materials, which are collected and managed, is done in accord with the mission of the Royal Library, its priorities and institutional operational needs.

Until the necessary financing has been achieved, the extent of the digital collections which are to be included in digital preservation is limited principally to digital safety copies and born-digital collection materials, with clear priorities as indicated in the following table.

The policy relates to materials in the following groups of materials:

Category of material	Specific material	Priority in connection with digital preservation Scale: 1..5, 1 = highest
Born-digital material (with its metadata), not covered by legal deposit, but which is desired incorporated in the national library's collections for preservation over time.	<ul style="list-style-type: none">- digital texts- digital maps- digital sound- digital still picture collections- moving pictures with sound- the library's own born-digital productions	2
Born-digital material (with its metadata), which is desired incorporated in Copenhagen University Library's collections for preservation over time.		Priority dependent on Copenhagen University Library financing
Born-digital material (with its metadata), not covered by legal deposit but which is desired incorporated in the national library's collections for short term dissemination	Digital materials donated by the library's users or other institutions for dissemination purposes User generated metadata for material the library has made available on our web	None at this time



Born-digital material (with its metadata), which is desired incorporated in Copenhagen University Library's collections for short term dissemination		Priority dependent on Copenhagen University Library financing
Retro-digitalized material (with its metadata), of which the digital copy is the only copy the library has, or the copy is the library's safety copy.	<ul style="list-style-type: none"> - Material created as part of substitution projects, including material created as part of <i>Dansk Folkeminde Samlings</i> media conversion projects - Material which has been borrowed for digitization to complete the library's own collections 	1
Retro-digitalized material (with its metadata), created through the library's ordinary digitizing process, where the physical material is still preserved in healthy condition in our storage facilities	E. g.: books, pictures, maps, manuscripts, and archival materials digitized for dissemination	None at this time
Retro-digitalized material (with its metadata), created through the library's ordinary digitizing process, and thereafter provided with significant further costly enhancement of the digital material	E.g.: Master copies of <i>Arkiv for Dansk Litteratur</i> and <i>Arkiv for danske tidsskrifter</i> , and BCO54 (aerial photographing of DK in 1954)	3
Retro-digitized material	Digital copy of material,	5, on the condition that the



(with its metadata), which has been created via sale of a digital copy in an especially high resolution	which is (frequently) digitized for sale	responsible collection department assumes that it will be incorporated in its collection and thus assumes the extra preservation expenses for the digital copy– otherwise none
Digital material (with its metadata), which is part of the library's services to other institutions	Digital material which is stored or managed for another institution or person as a part of a regulated and described cooperation	Only to the extent that the expenses are covered 100% by the owner of the data according to the general rules for income-generating activity
Digital material (with its metadata), which is covered by legal deposit legislation, and which the Royal Library is solely responsible for preserving over time.	<ul style="list-style-type: none"> - Games - Materials received on physical media such as floppy, cd or dvd together with legal deposit material, which is not in a network - Digital maps 	4, as long as there is not financing dedicated to this purpose
Legal deposit digital material (with its metadata)	Internet material	Covered by the Danish Net archives' digital preservation in cooperation with The State and University Library

The programmes (with their metadata), which influence the quality of data after we have received it at the Royal Library, i.e. programmes used for formatting of data in connection with incorporating them in our systems or later format conversion, as well as programmes to be used for interpretation of data in our systems, must always be preserved. Correspondingly, the log files for format conversions are to be preserved.

4.2. Acquisition and incorporation in the digital collections of the Royal Library

For each group of materials, those responsible for the collection must, together with the relevant vice director for either the national library or the Copenhagen University Library, set the priorities for which specific material should be collected within the economic framework provided for this purpose.

Those responsible for the collections also draw up the agreements with the providers and producers of digital collection materials, including specifications for delivery formats, metadata, packaging, frequency of delivery, quality control goals, rights issues, sanctions etc.

The material with its metadata is as a rule stored in KB/DOMS, where it is managed.

If there is not already a workflow which describes the life cycle of this particular material in the library's systems, an agreement must be made between the IT-Department and the Department for Digital Preservation at the time of acquisition to ensure that such acquisitions can be managed.

For each type of material decisions must be taken on:

- Which data formats are to be managed
- In which data formats materials are to be disseminated (and thus also where and how dissemination is carried out)
- Which routines and programmes, if necessary, are to be used to adapt the acquired data to the systems of the Royal Library
- Which metadata are to be assigned (what, how and when)

This information is kept up-to-date and published on the intranet.

All material stored in KB/DOMS is subject to the general backup routines. This means that copies are routinely made of the files on another media, from which a re-establishment of the files, in the condition and organisation which they had at the point of backup can be made – typically with a cycle of one week. As in ordinary backup systems it is accepted that material added within the last 24 hours may be completely lost.



If material is to be covered at a higher level, it must be included in digital preservation, where established security ensures that the files do not change and can be preserved and interpreted over longer numbers of years. This solution takes into account that material stored in a trusted digital repository cannot be lost.

Until financing has been established that makes it possible to preserve all the digital collections of the Royal Library for the long term, there must be a clear-cut distinction in priorities between what is preserved in order to give access at the present and what is preserved in order to give access both now and in the long term. Material which has priority to be covered by digital preservation is subject to special procedures with regard to storage and management:

- Material which is covered by these procedures is to be stored in the trusted digital repository of the Royal Library
- When a collection is selected for digital preservation at the same time the annual financing with which to carry this out must be in place
- Collections may only be selected for digital preservation if
 - there is an agreement with the vice director who has budget responsibility for the activity
 - a collection department has been named the owner responsible and thus has responsibility for its management
 - the material exists in data formats and a quality that have been approved for digital preservation, see the document: *"Strategy for preservation of digital collection material at the Royal Library"*
 - the material has been provided with the minimum of metadata that makes it possible to manage and provide access to the preserved data for the long term
 - quality control (to what extent and how) has been agreed upon

4.3. Discarding

The building of the library's digital collections according to an explicit policy has no long history at the library and the systems to carry out these tasks are just now being established and will continue to be in rapid development in the coming years. A period is now necessary during which experience is gathered, from both the collections' and the technical viewpoint, before the introduction of a detailed discarding policy.



An actual policy for discarding should be produced for this area at the latest in 2014 (5 years after the approval of the first preservation policy).

Until an actual discarding policy has been approved, the following simple set of rules is in force:

- As a starting point material stored in KB-DOMS may be deleted by the person responsible for the collection, insofar as this is material which we no longer wish to disseminate from the library.
- It should not be possible to delete any digital master copies produced as part of a retro-digitization process and which are not in a trusted digital repository without a prior agreement with the responsible vice director, since the expense of production is often significant.
- Material stored in the library's trusted digital repository cannot be deleted.