Digital Preservation Strategy
for
the State and University Library, Denmark

Version 4
February 2016
Table of content
1. Introduction ............................................................................................................. 3
2. Objectives ................................................................................................................ 3
3. Preservation Framework ........................................................................................................ 4
   3.1 Roles and Responsibilities ............................................................................. 4
   3.2 Knowledge Sharing and Competency Development ........................................... 5
   3.3 Digital Infrastructure ................................................................................... 5
   3.4 Trustworthy Digital Repository ....................................................................... 6
   3.5 Registry of Recommended and Accepted Data Formats ........................................... 6
4. Requirements for Preservation of Individual Collections............................................ 7
   4.1 Legislation and Guidelines ............................................................................. 7
   4.2 Finances .................................................................................................. 8
   4.3 Risk Management ....................................................................................... 8
   4.4 Access to Digital Collections .......................................................................... 8
   4.5 Bit Preservation ......................................................................................... 9
   4.6 Functional Preservation ................................................................................ 9
      4.6.1 Acquisition of digital material ............................................................... 9
      4.6.2 Ingest ............................................................................................... 10
      4.6.3 Continuous Monitoring ........................................................................... 10
      4.6.4 Preservation Actions ............................................................................ 10
   4.7 Metadata ................................................................................................ 11
   4.8 Quality Control in Digitisation Projects ............................................................ 12
   4.9 Resource Persons ...................................................................................... 12
5. Sources of Inspiration and Related Documents .......................................................... 12
   5.1 Sources of Inspiration .................................................................................. 12
   5.2 Internal Documents related to the Digital Preservation Strategy ....................... 13
6. Definitions ............................................................................................................. 14
1. Introduction

The Digital Preservation Strategy applies to all digital collections at the State and University Library, Denmark. The strategy has been developed to ensure that the library’s digital collections are preserved in the best possible way for the future. The strategy outlines principles and priorities for carrying out digital preservation at the State and University Library. Preservation activities concerning analogue materials, including physical media with digital information such as CDs and DVDs, are not included in this strategy, but are covered by the document Bevaringsplan (the State and University Library’s preservation plan for analogue materials).

Collections for which the State and University Library shares the curatorial responsibility with another institution, are also covered by the Digital Preservation Strategy. However, an individual collection plan must be developed in close cooperation with the other institution(s) which the library shares the responsibility with.

The Digital Preservation Strategy for the State and University Library must be seen in conjunction with a complex of policies, strategies and guidelines for the work related to handling and preserving digital material at the library. The strategy is closely connected to the State and University Library’s Digital Preservation Policy, which outlines the framework for carrying out digital preservation at the library.

The Digital Preservation Strategy for the State and University Library is reviewed and updated annually. As the official owner of the library’s digital collections, the Head of the National Library Division will initiate and approve the annual revision of the strategy.

The strategy is public and is published on the website of the State and University Library.

2. Objectives

The Digital Preservation Strategy for the State and University Library states the library’s prioritisations and principles for how to carry out preservation of digital material. The strategy thereby creates a basis for deciding how to plan and work with digital preservation at the library. The strategy ensures the taking of qualified choices of methods for preservation of digital collections for the future.

The strategy forms the basis for the ongoing selection of which digital collections must be preserved for the future, and it clarifies roles and responsibilities in connection with this. At the same time the strategy aims to ensure that digital preservation is included from the early planning stages of digitisation projects as well as in connection with acquisition and discarding of collections. The objective of the strategy is also to ensure that digital cultural heritage is accessible to the library users now and in the future.

The State and University Library will implement the prioritisations and principles of the strategy in its daily work on digital preservation. Thus the strategy will help coordinate internal cooperation at the library as well as cooperation with external parties, including international
partners. At the same time the strategy will ensure the retention of decisions and seek to ensure proper financing for digital preservation now and in the future.

Finally, the strategy outlines the guidelines for the creation of an individual preservation plan for each digital collection.

3. Preservation Framework

3.1 Roles and Responsibilities

Roles and responsibilities within the digital preservation area are allocated in accordance with regulations in ISO27001 and the overall organisation of the library.

Roles and Responsibilities on Division level

- The IT Division is responsible for technical implementation and service of preservation systems etc.
- The National Library Division is the owner of the digital collections and as such has the overall responsibility for their preservation. The ownership includes continuous control of the digital collections, including assessment of whether the preservation means decided upon are sufficient. Decisions about preservation formats, access control, etc. are made in consultation with the IT Division.
- The National Library Division is responsible for handling all contact to producers regarding the acquisition of digital collections.

Roles and Responsibilities in the IT Division’s Sections

- IT Services is responsible for the security and running of IT systems.
- IT Development is responsible for the development and maintenance of preservation systems and systems for authorisation.
- IT Development is responsible for the development and maintenance of access tools aimed at external users.
- IT Development is obliged to participate with two members in the Digital Preservation Group.

Data loss, errors and abuse in all systems are prevented by the use of back up, regular data logging and access control as described in the State and University Library’s ISO27001 implementation.

Roles and responsibilities in the National Library’s Section

- The owner of the collections prioritises collections for preservation and approves the level of preservation and strategy for the collections.
- The Digital Collections Manager (DCM) creates and maintains preservation plans for each collection in accordance with section 4 in this document.
- The DCM maintains a list of all digital collections.
- The DCM is responsible for coordinating preservation activities between the owner of the collections and the IT Division. This coordination takes place in the Digital Preservation
Group consisting of two members from IT Development and two members from the National Library Division.

3.2 Knowledge Sharing and Competency Development

The State and University Library ensures that its digital preservation staff has the competencies necessary for working in the digital preservation field and staying up-to-date on the topic. Competency development is performed e.g. by involvement in national and international projects and partnerships, tasks that contribute to maintain and develop the competencies within digital preservation.

The State and University Library conducts research in digital preservation as part of the institution’s competency development and knowledge sharing. Research in digital preservation is conducted in agreement with the overall research policy and plans of the library, and prioritises subjects of relevance to its digital preservation systems, e.g. bit preservation, and subjects related to the library’s role as curator of Danish cultural heritage, e.g. audio-visual media and newspaper collections. The library prioritises cooperating in research projects and is involved in a number of national and international digital preservation partnerships and projects to ensure research, competency development and knowledge sharing in the digital preservation area.

Knowledge generated by working with digital preservation is disseminated via relevant websites and projects, e.g. www.digitalbevaring.dk.

The library’s involvement in national and international partnerships is described in the Appendix ‘Projects’ to Digital Preservation Strategy for the State and University Library, Denmark.

3.3 Digital Infrastructure

All collections are preserved in the State and University Library’s Repository. The Repository consists of a data repository in which data is preserved as binary files and a metadata repository for the preservation of metadata.

The library uses the National Danish Bit Repository as its bit repository. The National Danish Bit Repository has been developed by the State and University Library in collaboration with The Royal Library and the Danish National Archives. This system enables us to have a replicated, active, integrity checked archive.

The library’s metadata archive is a self-developed system called DOMS (Digital Object Management System), based on Fedora Commons. Within the system it is possible to maintain and enrich the metadata and the relations between metadata objects. The relation to the files in the data repository is maintained as part of the metadata.

Data is ingested into the Repository as soon as possible and is displayed via other systems as described in the State and University Library’s Strategy for Information Flow. Furthermore, the library’s Repository allows for setting-up software which analyse and enrich data and metadata.

When a collection is preserved in collaboration with other institutions it might be more desirable to use other software solutions. In these cases the principles of the strategy described in this document are used, but not necessarily the same infrastructure.
3.4 Trustworthy Digital Repository

The State and University Library seeks to achieve the status of Trustworthy Digital Repository and thereby meet internationally acknowledged standards. This means e.g. that the State and University Library ensures integrity and authenticity in its digital preservation, including the correct and necessary metadata, and that the library meets the requirements for a digital infrastructure, organisation and documentation. At the same time it guarantees that the library has and updates its policy and strategy for digital preservation.

All components of the digital infrastructure of the State and University Library must adhere to internationally recognised requirements for trustworthiness within their areas. The library will use an internationally recognised audit tool to verify trustworthiness. An audit of the library’s status as Trustworthy Digital Repository must be carried out every second year. The plans and frameworks for the Trustworthy Digital Repository will be adjusted regularly and when needed. The owner of the collections is responsible for initiating the audit and for the follow-up in close cooperation with relevant staff and sections.

Audit of the State and University Library’s preservation is the basis for conducting risk management in accordance with the requirements in ISO27001, and the requirements the library has to meet in order to achieve the status of Trustworthy Digital Repository.

3.5 Registry of Recommended and Accepted Data Formats

It is the responsibility of the owner of the collections in collaboration with the IT Division to create and maintain a list of recommended and accepted data formats respectively for preservation in the State and University Library’s repository (see section 4.6.1). A recommended format is the format, which the library recommends for reception of a specific collection when it is possible to influence the choice. An accepted format is the format, which the library chooses to receive when it is not possible to influence the choice.

The clarification of formats and the update of the list of formats is done in the Digital Preservation Group.

The list of recommended and accepted data formats can be requested by contacting the library’s Digital Collections Manager. The IT Division and the owner of the collections cooperate with internal experts on data formats as well as with other institutions to maintain and announce information about these data formats.

The list of data formats for digital preservation is based on the following criteria:

- Openness of the format: Is the format well described and is documentation available? Is the format subject to any patents? Is a licence or permission required to use the format?
- Distribution of the format: Is the format used widespread? Will many software programmes be able to understand the format?
- Error tolerance of the format: Will a single bit error make the whole file unreadable? Has the format been compressed (lossless or lossy data compression)?
- Acceptance of the format as a preservation format: How is the format evaluated on corresponding lists of recommended formats?
- Dependency of the format on external sources of information, for example fonts or pictures with external references.
 Ability of the format to embed data in other formats, for example embedding a video in a pdf-file.

Based on these criteria the owner of the collections can add a data format to the list as “Recommended” or “Accepted”.

The list will be continuously updated as part of the preservation work. If a formerly accepted data format is no longer regarded as suitable, the owner of the collections must make sure that any collection containing the data format is considered for migration.

When a data format is included in the list as “Recommended” or “Accepted”, the owner of the collections must make sure that enough information about the data format is available, e.g. version numbering, description of the data format and any limitations regarding the suitability of the data format (e.g. a prohibition on embedding of data in other data formats). For each of the data formats a technical profile must be developed. The profile describes how the State and University Library uses the data format, for example what functionalities of the data format are used and which ones are not allowed, as well as other possible recommendations such as quality (for example resolution).

The owner of the collections will publish the list of data formats.

In order to identify criteria and information about data formats, the owner of the collections will aim to enter into co-operation with other institutions working with digital preservation.

### 4. Requirements for Preservation of Individual Collections

For each digital collection a preservation plan will be developed. The DCM is responsible for producing descriptions of the digital collections and composing preservation plans for the collections as they are ingested in the State and University Library’s Repository.

The preservation plan covering the individual collection relates to all elements and requirements described in the following sections of this document. The preservation plans are located in the State and University Library’s wiki for digital collections, which aside from the plans also holds the guidelines for composing and updating preservation plans.

The DCM will involve the Digital Preservation Group as needed in the work with the collections.

#### 4.1 Legislation and Guidelines

The State and University Library’s preservation of digital material follows the legislation and guidelines which the library applies to at all times. The owner of the collections collaborates with the library’s internal adviser on legal questions in order to respond continuously to relevant legislation and guidelines. The preservation plan for each collection contains information about the legislation and guidelines which are of relevance to the specific collection.
4.2 Finances

In collaboration with the IT Division the owner of the collections estimates the costs for preservation of each collection and documents the conclusions in the preservation plan. Together with the library’s accession policies the cost estimates will support the choice of preservation strategy for each digital collection. Parameters which can be used to estimate the costs include choice of data formats, number of copies, ingest requirements, archival storage costs, and more.

The owner of the collections will stay updated on international research and work done in regards to calculating the costs of digital preservation.

4.3 Risk Management

A risk analysis is produced for each digital collection, and the risks are described in the preservation plan for the digital collection. The owner of the collections is responsible for preparing and updating the risk registry for each collection. The risk analysis is based on an evaluation of the condition of the collection and the conclusions from the most recent audit of the State and University Library as Trustworthy Digital Repository.

4.4 Access to Digital Collections

Continuous access to digital material is the basis of digital preservation. Aside from this, use of the digital collections does also implicitly help check preservation quality.

The State and University Library aims to provide online access to all digital collections ready for dissemination via the library’s website and other dissemination channels in a technology and in a format which are of relevance to its users, see the document *Strategy for Information Flow*.

The owner of the collections can decide to make digital collections available for exposure through other dissemination channels. Defining the conditions for obtaining access to a collection is the responsibility of the owner of the digital collections, see section 4.1. Legislation and Guidelines. If the State and University Library decides that a need for broader access exists, the library ensures agreements or contracts to meet this need, see *Strategy for Information Flow*.

Information about the provenance of the digital collections is registered in the library’s metadata repository. Information about the digital material's rights is registered in the library’s license module, and access to the digital material is controlled through this system which ensures that all rights are respected technically and legally so no user will obtain access to material he or she has not been authorised to access.

All known information concerning access to a digital collection is described in the preservation plan for the individual collection.
4.5 Bit Preservation

The State and University Library preserves data in the library’s data repository, see section 3.3 Digital Infrastructure. The library carries out bit preservation on all digital collections as a minimum. The IT Services Department is responsible for all bit preservation activities. Service and maintenance of bit preservation systems are conducted by the IT Services Department and prospective partners.

As a minimum, the State and University Library keeps at least two copies of data. One of these is kept at the library’s main address and the other copy is located at the library’s facilities in Skejby, Aarhus. The two copies are stored by using different technologies, and the library makes sure that both copies are not controlled by the same organisational unit and/or person. When only two digital copies of a collection are preserved a checksum is preserved separately from the two copies.

A risk assessment and a cost benefit analysis are made for each collection by the owner of the digital collections. Based on these analyses the owner of the collections states the requirements regarding the number of copies and how frequently the control of bit integrity must be carried out. The IT Services Department is responsible for generating checksums of all files and for controlling bit integrity.

The State and University Library uses the infrastructure of the National Danish Bit Repository for bit preservation.

All decisions concerning digital preservation are documented in the preservation plan for the individual collection.

4.6 Functional Preservation

The State and University Library aims to preserve the functionality of a digital collection either by migrating data to new data formats or emulating the original systems to be able to access data.

4.6.1 Acquisition of digital material
As far as possible the owner of the collections will attempt to ensure that the digital collection exists in a suitable data format and will document the data format in the preservation plan and in the library’s list of digital formats.

When the State and University Library itself is the source of digital material, for example in case of digitisation (in-house as well as outsourced), the owner of the collections will always choose a suitable data format. When the library receives a collection from an external source, the owner of the collections will attempt to influence the provider to deliver the collection in a suitable data format. If this is not possible, the owner of the collections will decide if the data format, which the collection will be delivered in, can be accepted. If the suggested data format cannot be accepted, the owner of the collections will decide if it is necessary to let the digital collection undergo migration or the like when the library receives it or if the State and University Library will refuse to receive the data.
4.6.2 Ingest
As soon as possible after a collection has been received, all data and metadata are ingested into the State and University Library’s Repository. The system owner of the library’s digital infrastructure is responsible for ingesting the collection.

4.6.3 Continuous Monitoring
Once a collection has been ingested it is the responsibility of the owner of the collections to monitor whether the collection’s data and metadata comply with the chosen specifications and whether the formats are still suitable for preservation. If this is not the case, the owner of the collections will decide which preservation actions should be performed to mitigate errors.

The preservation plan documents how a collection’s data and metadata comply with the chosen specifications and that the formats of the digital collection are suitable for preservation.

If a collection’s data and / or metadata are not compliant with the chosen specifications, the library must seek to mitigate this in one of the following ways: reacquisition of data and metadata, manual amendment of errors, mechanical amendment of errors, discarding etc.

4.6.4 Preservation Actions
Once a collection has been ingested into the Repository a set of preservation actions can be carried out. The owner of the collections and the system owner coordinate the activities. A preservation action is an action which will enhance the preservation properties of a collection. A preservation action can change the preservation object or create derived data in the form of enriched metadata, dissemination copies or the like. Any invalidity is registered in the same way as other characterisation results.

All preservation actions are documented in the collection’s metadata.

Examples of preservation actions which can be carried out on collections in the Repository are shown below:

**Preservation action – Characterisation**
Characterisation is an action which enriches metadata with information about the collection. The result of the characterisation is registered in the metadata as well as in the collection plan. Any invalidity is registered on equal terms as the other characterisation results.

**Preservation action – Refinement of data model**
All data is ingested with a basic and generic data model which can subsequently be refined to be more collection specific in order to be more operable on the individual collection. The preservation action can result in a new and optimised data model, e.g. with new metadata formats.

**Preservation action – Migration**
If the formats of a collection are not suitable for preservation, the owner of the collections can decide that the data should be migrated. If migration is not a possibility the owner of the collections can decide to emulate or use filming (see below). When using migration the owner of the collections must choose a suitable data format for the method. Evaluation of the suitability of the migration data format is based on the following criteria:
• Suitability of the new data format for long term preservation (for more information, read paragraph 3.5 Registry of Suitable Data Formats),
• Preservation of significant properties (for a chosen target audience) (e.g. contents, layout, embedded metadata, possibilities for interaction),
• The accessibility and quality of the migration tools, and
• The cost of migration.

As a rule both the original and all migration copies will be preserved. The owner of the collections can decide to deviate from this principle on the basis of costs compared to the quality of migration, i.e. the preserved properties from the original to the migration copy. The system owner of the Repository is responsible for carrying out the migration and documenting the process.

**Preservation Action - Emulation**

If it is not possible to use migration for a collection (or the threatened parts of it) in a way that the collection can be accessed in the future, the owner of the collections can decide to make use of emulation.

Migration can be opted out if, for instance, it is not possible to find a migration path that preserves the significant properties, or if migration is not cost-effective due to e.g. large amounts of different data formats in the collection. If the owner of the collections and the system owner decide that emulation is a reasonable alternative to migration, the owner will contact a partner with expertise in emulation in order to establish an emulation platform with access to the collection data. The system owner of the State and University Library’s digital infrastructure will cooperate with the emulation partner on the technical set-up of the emulation platform.

**Preservation Action - Filming**

If the owner of the collections and the system owner of DOMS deem it unrealistic to preserve an interactive collection by emulation, but the owner of the collections still finds the collection to be of such value that certain properties of the collection must be preserved, the owner of the collections can ask a business partner if they would like to film the contents and use of the collection. Ingest of these files must be conducted in accordance with the preservation principles which concern ingest of digital collections at the library.

### 4.7 Metadata

The owner of the collections ensures that all digital collections and associated metadata are registered when received by the library.

The revision history is preserved in the metadata.

Metadata for all collections is preserved in the State and University Library’s metadata repository and is kept separate from the data files. The relation between data and metadata is sustained by the metadata containing a reference to the correspondent data in the data repository.
Some files may have embedded metadata, e.g. generated during the digitisation process. This metadata will, if possible, be extracted and preserved with the other metadata, but will also remain in the actual files and will be sought preserved in future migrations.

When an analogue collection is digitised, data generated in connection with the digitisation process is added to the metadata of the collection.

Collections are registered in the metadata repository with the generic data model when ingested. A data model for the individual digital collection is continually developed in collaboration between the owner of the digital collections and the library’s Metadata Team subgroup working with metadata concerning the library’s own digital resources. The Metadata subgroup describes metadata standards used at the library, and monitors the international development in metadata standards including descriptive, administrative (including technical) and structural metadata.

### 4.8 Quality Control in Digitisation Projects

When a collection is digitised at the State and University Library’s request, the validity and quality of the digital collection is investigated at reception by checking the following elements: File structure, data files and metadata. The results of the investigations are documented in the preservation plan for the individual digital collection.

### 4.9 Resource Persons

The preservation plan for a digital collection must list the names and contact information of persons who hold special information about the contents and preservation of the collection. The resource persons for a given collection can be external parties as well as internal members of staff who hold important knowledge of the collection. This can be information about the contents and history of the collection or information about the data formats, preservation history, metadata, and more.

### 5. Sources of Inspiration and Related Documents

#### 5.1 Sources of Inspiration

The following documents have been sources of inspiration for the preparation of the *Digital Preservation Strategy*.


5.2 Internal Documents related to the Digital Preservation Strategy

The following documents and draft texts have been used in the process of developing the Digital Preservation Strategy version 4.

Statsbibliotekets Politik for digital bevaring v.3.0 (2014) (Digital Preservation Policy for the State and University Library)
(https://www.statsbiblioteket.dk/om-statsbiblioteket/filer/politik-for-digital-bevaring_ver3)

Statsbibliotekets IT-strategi (2012) (IT strategy for the State and University Library)

Statsbibliotekets Indsamlingspolitik for pligtafleveret radio/tv (2014) (Acquisition Policy for Legal Deposit radio and television)
(http://www.statsbiblioteket.dk/om-statsbiblioteket/publikationer/indsamlingspolitikfortradiotvfinalaugust2014.pdf)

Statsbibliotekets Indsamlingspolitik for audiovisuelle materialer (analoge og digitale) (2014) (Acquisition Policy for audiovisual material (analogue and digital))
http://www.statsbiblioteket.dk/om-statsbiblioteket/publikationer/indsamlingspolitik-for-audiovisuelle-materialer


Strategi for informationsforsyning (2012) (Strategy for Information Flow)
(http://www.statsbiblioteket.dk/om-statsbiblioteket/filer/strategi_for_informationsforsyning)

The following appendixes can be requested by contacting the Digital Collections Manager at the State and University Library:

Appendix for Strategy for Digital Preservation version 4: Bit Preservation Levels
Appendix for Strategy for Digital Preservation version 4: List of formats
Appendix for Strategy for Digital Preservation version 4: Projects
6. Definitions

The State and University Library’s definitions of the following phrases:

- **Bit Repository**
  (The National Danish) Bit Repository is the State and University Library’s primary system for bit preservation. It is the result of collaboration between the Danish National Archives, the Royal Library and the State and University Library with the purpose of guaranteeing bit preservation of digital material.

- **Copy**
  One among a number of identical duplicates of the same file which is preserved for the purpose of completing regular integrity checks.

- **Data repository**
  The State and University Library uses the Bit Repository infrastructure for its bit preservation (see The National Danish Bit Repository).

- **Digital collection**
  Collection of digital files and metadata which can be seen as a whole, e.g. defined by content or data formats.

- **Digital Collections Manager (DCM)**
  Function in the National Library Division which is responsible for administration of the digital collections at the State and University Library.

- **Digital Collections Wiki**
  Wiki used for collecting information and decisions concerning the State and University Library’s digital collections.

- **Digital infrastructure**
  The Digital infrastructure is the infrastructure which comprise the Repository of the library.

- **Dissemination copy**
  Copy of original digital material used for dissemination purposes. A dissemination copy emphasises the possibility of disseminating content to users.

- **DOMS**
  The State and University Library’s Digital Object Management System which has been developed in order to preserve and administer objects with metadata and the relations between these.

- **License system**
  System for registration of access rights.

- **Metadata repository**
  The State and University Library uses DOMS for its preservation of metadata (see DOMS).

- **Metadata Team**
  The Metadata Team at the State and University Library coordinates cross-section efforts at the library with the purpose of staying up to date with the development in metadata standards and paying attention to changes, amendments and best practice, as well as compose guidelines/rules for registration in order to guarantee a uniform use of the standards.

- **Migration copy**
  A new version of a file in a different data format. A migration copy can be created from both the original and from a previously generated migration copy.
• **Preservation copy**
  Copy of original digital material needed for long term preservation. In a preservation copy features like long lasting formats which enhance long term preservation are prioritised.

• **Original**
  The original file which is used for generating subsequent copies

• **Owner of the collections**
  The role responsible for the digital collections, including maintenance, preservation and dissemination of the digital collections. At the State and University Library the Head of the National Library Division is the owner of the library’s analogue and digital collections

• **The State and University Library’s Repository**
  The digital collections at the State and University Library and the preservation and organisation of these

• **System owner**
  The role responsible for maintenance and service vital for performing digital preservation. The Head of the IT Department is the system owner in relation to systems used for digital preservation

The State and University Library defines the following terms in accordance with the descriptions at www.digitalbevaring.dk (in Danish):

  o Bit preservation (Bitbevaring)
  o Checksum (Checksum)
  o Data format (Dataformat)
  o Digital preservation (Digital bevaring)
  o Emulation (Emulering)
  o Format (Format)
  o Functional preservation (Funktionel/logisk bevaring)
  o Metadata (Metadata)
  o Migration (Migrering)
  o Risk management (Risikovurdering)