

# NORSK MEDIEARKIV

© Børre Ludvigsen, Høgskolen i Østfold  
March 15, 2012

## Norsk sammendrag

Dette dokumentet beskriver en forstudie for etablering av et multimediaarkiv kalt *Norsk MedieArkiv* for frivillig deponering av digitale dokumenter som ikke dekkes av "Lov om avleveringsplikt for allment tilgjengelege dokument".

Økt tilgjengelighet av digitale hjelpemidler som f.eks. datamaskiner, fotografi-videoapparater, scannere og musikkopptagere har gjort det vesentlig enklere å produsere dokumenter av alle slag i digital form. Men det er nettopp den digitale formen som gjør de så veldig mye mer sårbare enn analoge dokumenter.

Tiden frem til etableringen av en permanent ordning for frivillig deponering til permanent lagring av private digitale dokumenter som bilder, text og multimediafiler vil bli preget av et stort tap av personlig kildemateriale. Lemfeldig behandling av lagringsmedier og omgang med datamaskiner preget av uhell og uvisshet vil i all hovedsak forårsake slikt tap og dette prosjektets mål er å undersøke hvordan denne tapstiden kan gjøres så kort som mulig og å avdekke betingelser for etablering av et Norsk Mediearkiv.

# Contents

<b>Contents</b>	<b>3</b>
<b>1 NORSK MEDIEARKIV</b>	<b>1</b>
Overview . . . . .	1
1.1 Archiving . . . . .	1
1.2 Activities . . . . .	3
1.3 Participants . . . . .	7
1.4 Links . . . . .	8



# 1 NORSK MEDIEARKIV

**Keywords:** Metadatamodelling and infrastructures for large multimediarchives to be used in research, teaching and digital libraries, metadata, multimedia, digital archives.

## Overview

Beginning with the work of Eilert Sundt, the recognition of the importance of recording everyday culture has grown steadily over the last century to become the subject of study and preservation by both academic and cultural institutions in Norway. Collections of photographs such as that of Anders Beer Wilse<sup>1</sup> and the private albums saved by Ragna Solleid<sup>2</sup> provide the historic visual record essential to the understanding of everyday culture.

The Norwegian law of *Legal Deposit of Generally Available Documents* provides for the obligatory deposit of all publications made for circulation outside private spheres, irrespective of media. This project is a feasibility study into a publicly run repository of private images and other multimedia artifacts *not covered* by the above law, preserving people's personal documentation of Norwegian contemporary popular culture. Nor is it meant to compete with the digitization efforts of Norwegian cultural institutions, but to preserve the nation's private views of their everyday lives.

## 1.1 Archiving

The law of obligatory deposit<sup>3</sup> obliges all who publish documents irrespective of media which are "made for distribution beyond the private sphere" to deposit copies with the national library. Section 1 of the act states: "The purpose of this Act is to ensure that documents containing generally available information are deposited in national collections, so that these records

---

<sup>1</sup>Deposited with *Norsk Folkemuseum* and now digitized by the National Library

<sup>2</sup>Librarian at the Bergen public library (Larsen and Lied, *Norsk Fotohistorie*, Det Norske Samlaget, Oslo 2007, p. 9)

<sup>3</sup>[http://www.nb.no/fag/for\\_utgjevarar\\_og\\_trykkeri/pliktavlevering/legal\\_deposit](http://www.nb.no/fag/for_utgjevarar_og_trykkeri/pliktavlevering/legal_deposit)

of Norwegian cultural and social life may be preserved and made available as source material for purposes of research and documentation.”

The act does not therefore cover privately recorded and distributed documents such as digital photographs.

In consideration of the following:

- historical albums of photographs have traditionally been our most important source of visual knowledge about our recent history,
- digital photographs are under greater threat than the analog inasmuch as they are easier to destroy permanently and are dependent of more complicated backup routines for preservation,
- digital cameras have significantly simplified photography and thereby increased the possibilities of documenting a greater multiplicity of aspects of our contemporary culture,
- our lives are more frequently and richly documented through digital media.

This projects aims to investigate the feasibility of establishing a massive mediaarchive available to the Norwegian public for the purpose of preserving digitally recorded material that documents our contemporary culture. The purpose of the archive is to serve the roles previously filled by private, analog collections of cultural information such as photographic albums, amateur film, audio tape recordings, diaries and the like, ensuring that these highly volatile digital collections are available for future generations of public and researchers alike.

Modern methods of digital documentation makes recording the everyday very much easier than even quite recent analog means. In photography for example, the basic system of loading film, then developing and printing either commercially or at home has changed very little since the early days of Kodak box cameras.<sup>4</sup> In terms of preservation however analog photography is much less volatile than digital. Other than damage by direct physical destruction, film negative or prints in the average home have reasonable archival lifespans when stored in the normal album or card board box in the attic.

Digital images, especially in these early years of the common use of the technology are much more volatile. Apart from the likelihood of loss through

---

<sup>4</sup>The various “instant” systems pioneered by Polaroid are an exception, but the vast majority of analog photographs have been film-based.

sloppy backup routines, hard disk crashes, CD-malfunctions and technology changes, there are the problems of data lost to defunct proprietary formats and simple ignorance causing inadvertent file erasure. Even worse is the quite common procedure of asking for the camera's memory chip to be erased at the photo shop after prints have been made. A surprising number of people treat the memory chip simply as a replacement for film without the hassle of having to care for the negatives! Furthermore, color prints made with consumer quality ink-jet technology have very limited lifespans, almost fading as one watches.

A publicly available repository for albums of private digital images (and other digital material) would be an invaluable source of information for contemporary and future researchers of everyday life and culture. It is envisioned that such an archive would have a simple, easy to use interface for depositing, annotating and retrieving information. There would also be no limit on the size or number of files.

## 1.2 Activities

There are obviously a number of questions that arise out of such a vision that have to be addressed. The aim of this project is to investigate the feasibility of establishing and running a public media archive and the most apparent problems entailed. The three primary goals in the short term are:

- Proposing an infrastructure architecture
- Investigating appropriate metadata models storage
- A seminar to launch a debate on the feasibility and need for a public media archive

Main activities:

- Mapping similar projects

One of the first comments that has arisen during initial discussions of the project is the existence of a variety of similar sites such as Flickr, PBase, etc. which offer space in which to exhibit photographs. Common drawbacks are severe limitations on image sizes and restrictions on media types that can be deposited. These sites appear often to be focused on photography as art rather than repositories of the whole range of photo-imagery.

The closest to a general and comprehensive collection of images and other media types would be Google images. This however, is a global

index of images available on the net, which necessitates personal websites of some sort, severely restricting the number contributors.

The greatest objection to the commercially run “community” sites in the World Wide Web is their inherently commercial agenda. With no public welfare mission there are no guarantees that the content will be preserved in any way, nor that it or information about the contributors will be used for purposes other than those advertised at the initial establishment of the service.

Possible connections to and experience from UNINETTs *NorStore* project will also be investigated.

A number of similar projects have been proposed in the EU, France, US and UK<sup>5</sup>. The project will attempt to survey those that might be relevant to our aims in order to gather experience and avoid unnecessary duplication of potential mistakes. (See 1.4, page 8)

- Target users

Typical users of a public media archive will be:

- Researchers in the humanities and imaging sciences.
- Teachers at all levels
- Museums,
- The general public

The above study of similar projects will hopefully uncover other, equally interesting user groups.

- Scope, size and scale.

For a media repository to have full value there should be no restrictions on the technical quality of the deposited material. The study will therefore have to make qualified estimates of necessary space and scalability over a foreseeable future.

A study of the scope of the material to be deposited must also be made. This would include questions such as type of documentation, geographic provenance, digitized historic material vs. contemporary, definitions of “privacy” in terms of permissible contribution, how to limit access territorially and so on.

---

<sup>5</sup>The links provided are only a very small selection provided for illustration purposes.

A primary focus of the project will therefore be a study of necessary infrastructure both for a prototype and projected full scale archive.

- Intellectual property rights, fraud and ttl.

In order to encourage as wide a scope of contribution as possible a range of restrictions on access to deposited material might have to be offered. A survey of possible means of restrictions and protections will be made.

Fraudulent material, malicious attacks and annotation will need to be considered along with identity protection policies as set out by Datatilsynet<sup>6</sup>.

Time-to-live (TTL) policies will also be proposed, such as a write-only system where corrections will be allowed, but all previous versions preserved and trackable.

- Access, interface, administration

A key to the success of a cultural media repository is a simple but fully functional access management and administration system including interfaces for deposit and retrieval of material. The project will endeavor to propose the necessary mechanisms and minimum data necessary to accommodate satisfactory data administration over the long term with due attention to open formats and source code.

The results of such a proposal will also focus on possible application to other UNINETT projects in which we are engaged.

- Metadata and classification

A decisive advantage of modern digital media formats is the possibility of embedding metadata in individual files structured as EXIF, IPTC, etc..

The proposal will investigate and propose a minimal structure of metadata to be embedded and used in classification and retrieval. We will also look into the possibility of continual user annotation of files by means of Wiki-like systems. Recent developments in image recognition are also promising with respect to search and retrieval.

The main goal of this part of the project will be to explore the potential, methods and technology for harvesting, structuring, curating data and

---

<sup>6</sup>[http://www.datatilsynet.no/templates/Page\\_\\_\\_\\_\\_194.aspx](http://www.datatilsynet.no/templates/Page_____194.aspx)

its harvested metadata from multimedia files used in research and education today. This will include stand-alone files like images, sound and video with their currently included IPTC/EXIF metadata to the more complex composite file formats.

The three goals of the study will be:

1. To explore current standards and platforms for working with metadata in media-rich file-types, and select a relevant set as they apply to the goal of the NorStore project<sup>7</sup>
2. A test implementation solution of a harvesting and structuring solution based on software that fits into the NorStore solution, and do a trial to gain more experience on:
  - Relevant metadata standards (DC, IPTC), encoding (XML, binary) and structuring (RDF)
  - How to meet the issues of mass and scalability using available metadata
  - Enriching existing metadata through annotation and communities of participants
3. To propose how the experiences from the trial can be implemented

The usefulness of the findings will not only be limited to NorStore and its members, but is likely to provide valuable input to the Norwegian national archive and other archive institutions on digital initiatives. Working to prepare solutions to curate digitally born content, especially as it applies to media-rich file-types, the findings of from a project like this would likely be of significant importance.

- Ownership

It is envisioned that in the long term the mediaarchive will be administered by a public agency in order to guarantee its perpetual availability and preservation. The proposal in form of a near complete feasibility study will be submitted for discussion to relevant cultural preservation authorities for comment.

---

<sup>7</sup>P.19: <http://www.norstore.no/docs/norstore-20071030.pdf>

- Open standards

A subject that would need to be discussed is the use of open standards and possible embedding of software necessary for display of proprietary formats. Some media formats such as QTVR<sup>8</sup> which has been widely used for interactive panoramas has only recently been replaced by HTML5. Some thought would need to be applied to the feasibility of preserving material in such formats. Otherwise it would seem prudent to limit all deposits to open data formats. The actual definition of what would constitute open formats within the scope of the project must also be a subject of investigation.

- Prototyping

A limited prototype for testing and illustration purposes will be made. The prototype will attempt to simulate scope, quality and size and to illustrate possible interfaces for deposit and retrieval.

The prototype will be populated with a series of "albums" of digital material, primarily photographs annotated with embedded metadata. We will attempt to provide large numbers of images in a range of sizes ranging from small JPEG files of 2-10 Mpx through numerous 20 Mpx images with automatically embedded GPS information to gigapixel-sized multirow and spherical panoramas. Also video, sound and text will be included to illustrate the potential diversity of the archive. The "albums" will range from plain collections of related and unrelated (but annotated) media files to assemblies embedded in pages of explanatory texts.

### 1.3 Participants

- Børre Ludvigsen: Høgskolen i Østfold, project leader. Works on digitizing cultural multimedia, analog and digital photography.
- Minh Than Quang. Masters degree student at Høgskolen i Østfold whose thesis subject is "A Prototype for a Digital Norwegian Personal Archive". (2011-2012)
- Eirik Skogh. Masters degree student at Høgskolen i Østfold whose thesis subject is "Policies, Procedures and Interface to a Prototype for a Digital Norwegian Personal Archive". (2011-2012)
- Therese Roksvåg Nilsen. Masters degree student at Høgskolen i Østfold whose thesis subject is "Private digital archives". (2010-2012)

---

<sup>8</sup>Quick Time VR

- Olaf Scjeldrup: UNINETT, “reader” and commentator. Involved with UNINETT’s *NorStore* project.
- Robert Cailliau: Geneva, “reader” and commentator. Co-inventor of the World Wide Web and works on digital albums. ([http://en.wikipedia.org/wiki/Robert\\_Cailliau](http://en.wikipedia.org/wiki/Robert_Cailliau))

## 1.4 Links

- Similar projects
  - Memories for Life  
<http://www.memoriesforlife.org/>
  - A Lifetime Personal Web Space  
<http://connect.educause.edu/Library/EDUCAUSE+Quarterly/BeyondtheElectronicPortfo/39884>
  - Personal Archives Accessible in Digital Media  
<http://www.paradigm.ac.uk/>
  - MyLifeBits Project  
<http://research.microsoft.com/barc/mediapresence/MyLifeBits.aspx>
  - Continuous Archival and Retrieval of Personal Experiences  
<http://research.microsoft.com/CARPE2004/>
- Libraries and digitization initiatives
  - Nasjonalbiblioteket  
<http://www.nb.no/>
  - Digital collections at the Library of Congress  
<http://www.loc.gov/library/libarch-digital.html>
  - Nordiska Museet  
<http://www.nordiskamuseet.se/fotosekretariatet/>
  - Minerva Europe  
<http://www.minervaeurope.org/publications/globalreport.htm>
  - Numérisation du patrimoine culturel : Coopération européenne  
[http://www.culture.gouv.fr/culture/mrt/numerisation/fr/f\\_minerva.htm](http://www.culture.gouv.fr/culture/mrt/numerisation/fr/f_minerva.htm)
- Private examples
  - <http://abdallah.hiof.no/>

- <http://ddc.aub.edu.lb/projects/saab/tamir-nassar/>
- <http://valhall.hil.no/Innhold/foto/index.html>

