



68th IFLA General Conference and Council

August 18th - 24th 2002, Glasgow, Scotland

Code Number: 176-145-E
Meeting: WS
Simultaneous Interpretation: -

Creating Common Ground to Stand On: Image Collections and Digitisation Principles

Kate Parson

Project Leader, *Image Databases and Digitisation - platform for ALM-collaboration*
Royal Library of Sweden
Stockholm, Sweden

Abstract:

*In applying their signatures to the document "Funding Application for the Project **Image Databases and Digitisation - platform for ALM-collaboration**" (2000-03-02), the directors of Sweden's National Library, Nationalmuseum, National Heritage Board and National Archives embarked upon a joint venture project that aimed to build a common platform for collaboration between archives, libraries and museums where standards, norms and regulations for digital imaging practices and registration may be developed. The project aims towards the creation of a common working language and agenda for four cultural heritage institutions of differing traditions.*

Introduction

A survey from 1999 charted the activity of more than 200 digitisation projects in the ALM-sector in Sweden. In recent years, a growing number of ALM-institutions (archives, libraries and museums) have begun to digitise parts of their collections and present them in database environments, with or without linked digital image files. In many cases, cultural heritage institutions maintain collections with similar types of objects used for the purposes of research, education and enjoyment. Full and seamless accessibility, as seen from the end user's perspective, would require that these materials are made accessible through automated searches from a number of entry points. For the user searching for information, it is of little relevance which type of cultural heritage institution it is that physically maintains the sought-after material. Despite a number of common points of interest due to the overlapping of the collections maintained, ALM-institutions work independently and often without much communication

regarding digitisation efforts. As a result, the institutions work in isolation from one another with problems that are mutually significant.

Through better organisation, the tasks involved could be realised more effectively and the numbers problems solved together. Co-operation regarding registration principles and the applied technology would lay a foundation for providing future users the opportunity to conduct searches from a single platform. This opens up the possibility, in turn, for new learning opportunities. Among other things, complex searches might then be possible, resulting in relevant “hits” amongst different types of materials in the collections.

Background

In March of 2000 a project proposal was drawn up between members of the Royal National Library [1], the National Archives [2], the National Heritage Board [3] and the Nationalmuseum [4] of Sweden. The joint venture project aimed to create a starting point: a common platform for collaboration between archives, libraries and museums where standards, norms and regulations for digital imaging practices and registration may be developed. The project would aim towards the creation of a common working language and agenda for four cultural heritage institutions of differing traditions. In this paper I will describe our work in the project - started in March 2001 and to be completed in December of 2002 - which has grown out of the need and desire to pool intellectual resources and tackle a number of digitisation issues over institutional boundaries.

Participating institutions

- The Royal Library, Sweden’s National Library was established in 1661 with legislature concerning legal deposits and today is assigned by the Department of Education to acquire, describe, preserve and to make available all materials published in the country. The present volume of the book collection is estimated at 3.5 million; the poster collection, one half million; maps, 300 000; portraits 750 000; and assorted image-based materials (drawings, photographs, engravings and other prints) number another half million.
- The National Archives was established in 1681 and today is the overseeing organisation for the state’s national, provincial, municipal and numerous private archives, as well as the custodian of the archives of the Department of Defence. According to legislation for archives, and under the supervision of the Department of Culture, the law stipulates the responsibilities of preservation, organisation, and care of records for the purpose of making them available according to the public’s right of access. The archives house a wide spectrum of different types of materials: text-based documents but also considerable holdings of older manuscripts, maps, drawings, microfilm, sound and image-based materials.
- The National Heritage Board has its roots in the 1630’s and later in 1786 was established as an academy of antiquities with the responsibility of recording and caring for ancient monuments. Today the Board is regulated by the Department of Culture and is the central organisation in the country responsible, together with the provincial governments and regional museums, for seeing to it that Sweden’s unique cultural heritage and diversified cultural landscape is preserved and used. Historical maps, topographical and architectural registers, together with image-based documentation of monumentally important sites make up the larger portion of the Board’s physical collection of materials.
- The Nationalmuseum, also under the supervision of the Department of Culture, is the national depository for older works of art in Sweden. (Modern and contemporary works are housed at Moderna museet.) The collections consist mainly of older painting and sculpture (16 000 works), art handicrafts (30 000 works), drawings and graphic works (500 000 works).

A significant thing to note is that these four institutions had not previously worked together as a group before. As my colleague Wolfgang Undorf pointed out yesterday, ALM-collaboration has been quite sporadic and may be described as “informal action” [6]. The idea of ALM as an organisational solution for co-ordinating strategies of the different (but in many ways similar) types of cultural heritage institutions has existed in Sweden for a little over a decade. However, to this date, most efforts have been at the local level and otherwise have been limited to concrete, goal-oriented projects between archives and libraries.

After months of brainstorming, a project committee consisting of representatives from these four institutions (A - L - M and ArtM) set the wheels in motion in creating the aforementioned “common platform” upon which shared strategies in collection management and digitisation could be discussed. A number of specific, pressing needs and concerns of the four participating institutions regarding both registration issues and digital imaging issues were identified. Additionally, for the purpose of streamlining the project work, the area of concern for the project was limited to the custodianship of specifically image-based objects maintained commonly between all four participating institutions: maps, drawings, graphic works, photographs, and posters. The project was granted funding by the Bank of Sweden Tercentenary Foundation and one year later, in March of 2001, the organisational structure was put in place. With a budget of 330 000euro and 10 500 projected man-hours of work divvied up by 25 project participants over 21 calendar months, the project was finally lifted off the ground. Title of the project: *Image Databases and Digitisation - platform for ALM-collaboration* [5].

At the start of the project period a number of working groups was established to share current practices and research best solutions (both the possible and the “impossible” solutions) to better long-term collections management. These committees comprised of the in-house experts from each institution and who agreed to spend a portion of their time alongside their regular responsibilities to project work. Issues regarding terminology, collections description in a database environment, the implementation of shared authority files, and standards of technical quality for visual, digital documentation of the collections were pinpointed and prioritised as major areas of concern. Other issues that earlier had been identified (such as selection criteria for digitisation projects, authority file collaboration regarding topographical names and geographic places, legal issues such as copyright, and a programme of educational seminars and workshops geared towards smaller institutions and private-interest organisations) were put on hold for the interim, with the hope of addressing them at a later date and with renewed funding.

The first activity of the project was to establish a common working language that would enable participants to communicate on an even level, and to avoid unnecessary future misunderstanding. Experts from the Swedish Centre for Terminology (TNC) were brought in to help map out the standard terminological working process of unearthing the concepts at hand and the assignment of terms to represent those concepts, together with definitions for those terms agreeable to all. At any given point during the 30 weeks of our terminology activity, we had 50+ terms on the table up for “excerpting” or definition. Of those terms, some were alternately discarded and added to the list, according to the group’s findings along the way. We started with a list of terms we deemed critical for smooth communication in the project and divided those terms essentially into two categories, registration terms (ex. “thesaurus”, “indexing”, “classification”, “controlled vocabulary”, “controlled term”) and technical terms (“data”, “data carrier”, “image capture”, “digital image”). Our working committee of 13 people was set up, and the members assumed responsibility for the category of terms central to their area of work. Other experts employed at our institutions were called upon to provide additional expertise. To begin with, we explored the concepts and terms used locally at each institution and sought to pin down on paper the existing definitions for those terms to share with each other. A number of terms we found were traditional only for the “type” of institution it was found in. We began to formulate our discoveries, saying “In the archive world we would say *archive builder* rather than *collector*” or “In the museum sector we would never use the library’s terms *cataloguing* and *controlled vocabulary*, instead we would say *registration* or *key word lists*”. Next, we explored general lexicons and other reference and terminological research materials,

together with ISO documents featuring terminological reports. Our final task was to pare down our research, and, where no ISO definition was present, agree on a formulation taken from our findings that would best suit our purposes for the project. As our initial project activity -- and in retrospect an invaluable first step in our collaboration -- the terminology activity allowed the project participants to get to know one another and become more familiar with the working language environments and professional traditions of the archive, library, and museum sectors respectively.

With a stable vocabulary as our point of departure, the project work was then divided into two parallel work-packages that were to span the remaining fifteen calendar months of the project period: on the one hand, activities concerned with digitisation standards; and on the other hand, activities regarding registration issues.

The working committee concerned with digitisation standards set out in their work by following a plan that outlined key areas to be discussed and researched. The goal was to establish best-practice recommendations primarily for the benefit of the four participating institutions, however even other ALM-institutions and organisations active in the digitisation of cultural heritage were intended to benefit from the findings. These key areas were:

- the selection of file formats for varied purposes (long-term storage, presentation in a web environment, and delivery formats to be made available to users) and including a study on long-term preservation and maintenance/migration of data;
- technical metadata essential for documenting image capture;
- quality standards for visual, digital documentation of objects in the collections;
- principles and guidelines in colour management;
- and the principles and guidelines with regards to authenticity.

Subcommittees were set up to tackle each key area, and in an initial phase each institution provided insight into and compared current working practice. Following this, each subcommittee delved into researching existing working methods and models in use at other various leading institutions and organisations for the basis of their summary recommendations. This is the point we are at now, in late august of 2002. Additionally, as a sort of concluding exercise, this committee will also provide a summary of compiled working methods and routines when setting up digital and photographic imaging projects before the project period is complete. Practical hands-on visits to the photo and digital imaging studios of the participating institutions has provided a great deal of material to work with (insight into the tricks of the trade) that might otherwise have gone unobserved and unmentioned in our regular activity report. The working committee has decided to present this guideline of working methods and routines as a website where the entire activity's findings (both the report and the guideline) will be gathered. Not entirely unproblematic in the context of digitisation standards and working methods is the fact that the technology changes rapidly. As we all know, over time file formats are bettered and replaced by new formats, data carriers are exchanged for different carriers, digital storage is becoming more compact and the technology to hold it compact more affordable. "Best practices" in this field have a relatively short shelf life. Ideally our results will be maintained and updated in a sort of after-the-project-life.

Alongside the first work-package concerning technical quality guidelines and best practices there is, as I mentioned, a second work-package running parallel that has to do with registration principles. More specifically, it consists of four individual activities run one after the other during the course of the same fifteen months:

- one prefatory activity to explore and propose a common way of classifying/labelling the different types of objects in our collections in order to better organise and retrieve information in the database;
- a second activity which explores the possibility of developing shared name authority files between institutions;

- a third activity to compile a proposal delineating minimum-level obligatory data elements for the identification and description of objects;
- and a fourth and final activity that in effect is a research stint on the indexing of images according to motif (concept-based image retrieval, rather than content-based image retrieval).

Again, the committees' work was entered into by presenting existing practices at each institution. The next step was to debate what optimal practices might look like, that is, exploring what kinds of expectations we have as records custodians and also what kinds of expectations our users have as members of both a general public and as niche researchers. At this point, our working committees have proposed recommendations for the first three registration activities. At this time, we are in the introductory phase of our last - and most difficult - activity for this work-package. During the fall we will explore and debate image description and indexing according to such abstract properties as "motif", and have planned a series of meetings with art historians and art theorists on the subject. In addition, we have made contact with a number of professionals in the private, commercial image bank agency sector in Sweden and planned a meeting where we can share experiences with indexing tools and discuss our various needs regarding image retrieval.

Conclusion

I find it difficult to make conclusive remarks about the project results as such, as all of our current results are being considered internally before publication at the end of this year with the end of the project. I am happy, however, to be able to say that our project has been favourably evaluated for its initiative in an investigational governmental report on ALM-collaboration published in March of this year [7]. Among the final recommendations, the report calls for a permanency of precisely those issues - in addition to other digitisation issues - that our project has opened doors for in discussing in a trial format, ie. project form. As to open doors, it has been said more than once in our meetings: even if the only thing we were to achieve in our project time was the gaining of a clearer understanding and respect for our colleagues at neighbouring cultural heritage institutions, then indeed that must be worth something in itself.

[1] Royal National Library of Sweden: <http://www.kb.se>

[2] The National Archives of Sweden: <http://www.raa.se>

[3] The National Heritage Board of Sweden: <http://www.raa.se>

[4] Sweden's Nationalmuseum: <http://www.nationalmuseum.se>

[5] Collaboration project *Image Databases and Digitisation - platform for ALM-collaboration*: http://www.kb.se/ABM_plattform/Default_Start.htm

[6] Wolfgang Undorf. (2002). *Means before purpose - the development of cooperation between cultural heritage institutions in Sweden*. Stockholm.

[7] Magdalena Gram. (2002). *ABM-samverkan mellan arkiv, bibliotek och museer: En lägesrapport på uppdrag av regeringen från Kungl. biblioteket, Riksarkivet, Statens kulturråd och Riksantikvarieämbetet*. In Swedish only. Translated title: *ALM Co-operation between Archives, Libraries and Museums. A Report by Order of the Swedish Government on the Situation at the National Library, the National Archives, the Swedish National Council for Cultural Affairs, and the National Heritage Board*. Stockholm.