Networking for Digital Preservation
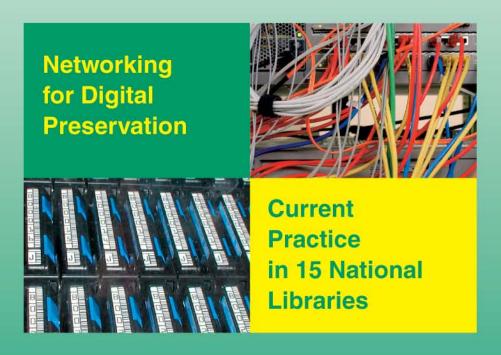
Current Practice in 15 National Libraries

Ingeborg Verheul

K · G · Saur

International Federation of Library Associations and Institutions

Fédération Internationale des Associations de Bibliothécaires et des Bibliothèques

Internationaler Verband der bibliothekarischen Vereine und Institutionen

Международная Федерация Библиотечных Ассоциаций и Учреждений

Federación Internacional de Asociaciones de Bibliotecarios y Bibliotecas

Networking for Digital Preservation: Current Practice in 15 National Libraries

Ingeborg Verheul

IFLA Publications
edited by Sjoerd Koopman

⊗

Printed on permanent paper
The paper used in this publication meets the minimum requirements of
American National Standard – Permanence of Paper
for Publications and Documents in Libraries and Archives
ANSI/NISO Z39.48-1992 (R1997)

Printed / Bound by Strauss GmbH, Mörlenbach

Foreword

Increasingly, libraries are having to deal with digital materials that need to be safeguarded not only for our generation, but also for the generations to come. Digitised images and born-digital objects need to be preserved for future access and use. For national libraries, safeguarding the digital heritage is a major issue because of their legal task to preserve the national heritage of a country in paper or digital form. One particular problem with digital material is the very short lifespan of the carriers. Moreover the hardware and software needed to render digital materials is undergoing constant technological development so that existing systems rapidly become obsolete. Therefore safely storing the digital heritage whilst still ensuring access for future use, requires that libraries not only need to have a trusted digital repository system in place, but also an ongoing R&D programme aimed at developing preservation strategies.

Despite being a relatively new field in the library sector, digital preservation is becoming increasingly important in the everyday routine of the library. Cooperation and knowledge dissemination activities on digital preservation issues are starting to emerge, but can still be intensified. An overview of recent developments in the field of digital preservation could be a valuable aid when planning digital preservation activities: Does the day-to-day practice in storing and accessing digital objects illustrate a mutual need for certain standards? Are there currently any standards for the development and building of digital repositories, and how are these being applied? Are there common standards in research on permanent access? Or is it still too early to speak of standards, and is it only possible to distinguish best practices?

In 2004−2005, *Koninklijke Bibliotheek* conducted a survey for the IFLA-CDNL Alliance for Bibliographic Standards (ICABS) on the use and development of standards in digital archiving within the international library world. The survey resulted in this overview of the current state of affairs in 15 libraries (baseline July 2005). The libraries involved are the national libraries of Australia, Austria, Canada, China, Denmark, France, Germany, Japan, the Netherlands, New Zealand, Portugal, Sweden, Switzerland, the United Kingdom and the United States of America.

The survey addresses both operational and R&D activities aimed at digital preservation. The main focal points are the use of standards in operational safe place environments and the state of affairs on permanent access strategies, such as migration and emulation. Apart from providing information on the status, function and organisational embedding of digital repositories in the library organisations, the survey also gives an overview of the current national and international R&D projects.

A comparison with two earlier surveys on developments in digital preservation (Neil Beagrie for CLIR and the Library of Congress in 2003 and the survey report of the PREMIS Working Group in 2004) leads to the conclusion that digital preservation in the year 2005 is becoming increasingly integrated in the day-to-day library activities. A growing number of national libraries consider it their mission to safeguard not only the paper cultural heritage, but the digital cultural heritage as well. Some of them are still focussing on the national heritage, whereas others have a more international perspective. Digital preservation is a rapidly developing discipline and although there are not many official standards for it yet, emerging best practices might develop into standards within a few years.

Since 2003, considerable progress has been made in building digital repositories, or networks of interconnected computer systems. A second development is the emergence of large networks for national and international cooperation. At first, the focus of these was limited to the cultural heritage sector, but now the perspective is broadening, and cooperation between the cultural heritage sector and the science sector is becoming increasingly important. Although such cooperative projects or platforms often started with a focus on knowledge dissemination and knowledge sharing, it is likely that they could be used in the near future for the joint development of tools for permanent access as well.

In 2004−2005 the National Library of Australia (NLA) also carried out a survey for ICABS on digital preservation. This survey focused on the availability of suitable guidance documents for preserving digital materials. During the ICABS Session of the IFLA World Library and Information Congress 2005 in Oslo, the KB and NLA reports were presented together in one lecture. The NLA report will only be available online. However, since both surveys emerged from the ICABS Alliance, a summary of the NLA report can be found in the Appendices of this volume.

In 2003, *Koninklijke Bibliotheek*, together with five other national libraries, has been one of the co-founding partners of ICABS, the strategic IFLA-CDNL alliance on bibliographic standards. Now that we are more than half way through the first term period of ICABS and the first evaluative thoughts on continuation and enhancement of the alliance emerge, it is important to stress the advantages and opportunities an international cooperative like ICABS offers to knowledge sharing within the library world.
Within its mission, ICABS forms a framework to stimulate the development of new strategies and to promote different aspects of the long-term preservation of electronic resources. With this survey on best practices in digital preservation *Koninklijke Bibliotheek* hopes to provide a worthwile contribution to ICABS.

Ingeborg Verheul prepared both the survey and the report on behalf of *Koninklijke Bibliotheek.* I am extremely grateful to her for providing us with this broad international perspective on digital preservation.

Dr Wim van Drimmelen

Director General
Koninklijke Bibliotheek

December 2005

Foreword

The IFLA-CDNL Alliance for Bibliographic Standards (ICABS) − an alliance founded jointly by the International Federation of Library Associations (IFLA), the Conference of Directors of National Libraries (CDNL) and the national libraries of Australia, Germany, the Netherlands, Portugal, the United Kingdom, and the United States of America − is a continuation of the late UBCIM Core Activity (with respect to Bibliographic Standards), parts of the Universal Dataflow and Telecommunications Core Activity, and the CDNL digital initiatives which involved preservation and digital resource management, access mechanisms, interoperability and much more.

The alliance has a strategic focus and offers a practical vehicle for improving international coordination and steering developments in these key areas. The alliance aims to maintain, promote, and harmonise existing standards and concepts related to bibliographic and resource control, to develop strategies for these, and to advance the understanding of issues related to the long-term archiving of electronic resources, including the promotion of new and recommended conventions for such archiving.

Within ICABS, *Koninklijke Bibliotheek* (KB), the national library of the Netherlands, and the National Library of Australia have been exploring the requirements and conditions for the long-term archiving of electronic resources. Moreover both libraries have been exploring and promoting strategies, methods, and standards for migration and emulation.

In this context KB conducted an international survey on recent developments in digital preservation in 15 national libraries. This report presents the outcomes of this study based on recent publications, information about ongoing projects and survey results.

We hope that this report will serve as a useful guide for other libraries and cultural heritage institutions as they face the future challenges of long-term archiving and preservation.

Renate Gömpel

Chair of ICABS Advisory Board
Die Deutsche Bibliothek

Table of contents

Websites: ICABS (http://www.ifla.org/VI/7/icabs.htm)
PADI (http://www.nla.gov.au/padi/)

All URLs in this publication were valid as of December 1, 2005

Acknowledgements

Introduction

Context

The research has been carried out within the framework of the IFLA-CDNL Alliance for Bibliographic Standards (ICABS).[1] It has been funded by structural funding from *Koninklijke Bibliotheek*, the National Library of the Netherlands, (KB) has received from the Dutch Ministry of Education, Culture and Science (OCW) for cooperation between KB, the National Archives, and the Netherlands Institute for Cultural Heritage (ICN) in the field of 'preservation' in the broadest sense. This cooperation concerns both paper and digital preservation. In 2004 part of this funding was earmarked for KB research on recent international developments on standards and best practices in digital preservation within the library sector, which resulted in this ICABS survey.

ICABS
ICABS is a strategic alliance of national libraries that focuses on improving the international coordination of bibliographic standards. The mission of ICABS is to maintain, promote and harmonise current standards on bibliographic and resource control. The efforts undertaken within the framework of ICABS have to stimulate the development of new strategies and as such promote different aspects of the long-term preservation of electronic resources and the promotion of standards.

ICABS was founded in 2003 to continue and expand the coordination work formerly done by the IFLA Universal Bibliographic Control and International Marc (UBCIM) and Universal Dataflow and Telecommunications Core Activity (UDT). Since the 1970s, UBCIM has focused on coordinating activities to develop systems and standards for bibliographic control at a national level and on the international exchange of bibliographic data. UBCIM cooperated closely with UDT. The IFLA Core Activity UDT mainly focused on analysis and research of technologies and standards for interoperability, information retrieval, digitising and metadata. The activities of UBCIM and UDT have been continued within ICABS. UBCIM and UDT received structural financial support from the Conference of Directors of National Libraries (CDNL). After the disbanding of these two Core Activities, the Committee on Digital Issues (CDI) was founded within CDNL. The committee's work on bibliographic standards and digital preservation is now being incorporated into the ICABS mission, while the Committee's work on deposit agreements will be continued separately by the National Library of Australia (NLA).

---

[1] ICABS: See: http://www.ifla.org/VI/7/icabs.htm

ICABS and digital preservation

Cultural heritage institutions are now becoming increasingly aware of the urgency of digital preservation: the long-term preservation of and permanent access to electronic objects. Several institutes have taken action to design and build trustworthy and permanent electronic repositories (sometimes referred to as 'safe places'). International projects such as NEDLIB (Networked European Deposit Library)[2] have improved cooperation, knowledge dissemination and standardisation in this field.

ICABS aims to promote knowledge dissemination on the long-term preservation of and permanent access to digital objects. Within ICABS, KB has assumed the responsibility for this task and is carrying it out in close cooperation with NLA.

---

[2] NEDLIB: See: http://www.kb.nl/coop/nedlib/

Aim, scope and methodology

Aim

This survey is the result of a research project KB carried out for ICABS in 2004−2005 on the use and development of standards in digital archiving within the international library world. This has resulted in an overview of current practice in digital preservation in 15 national libraries.

The main focal points in the survey are the use of standards in operational safe place environments (the current state of affairs of initiatives on electronic repositories), and the current state of affairs with respect to, and the need for permanent access strategies (such as migration and emulation). The survey also provides an overview of current national and international projects on digital preservation. The study made use of desk research, interviews and an analysis of new developments.

Scope

This KB survey can be seen as a continuation of two previous surveys: 1) the outcome of the research of Neil Beagrie on national digital preservation initiatives in 2003 at the request of the Library of Congress and the Council on Library and Information Resources[3] and 2) the outcome of the research by the OCLC/RLG PREMIS Working Group Preservation Metadata: Implementation Strategy (PREMIS) in 2003, the results of which were published in 2004.[4] The design of this survey combined subjects from both publications.
The report of Neil Beagrie was written to provide NDIIPP, the national digital preservation programme in the USA, with input and reference material to outline specific points of interest in an international context. It described current practice in digital preservation in Australia, France, the Netherlands and the UK in March 2002. The report provides a good perspective for designing a topical study. It also provided a good comparison of what had changed in three years, not only due to its international perspective but also the set of forecasts it provided about developments in digital preservation in the near future. Beagrie's report formed the basis for the sections on legal deposit, funding and current activities in this ICABS survey.

---

[3] Beagrie, Neil. 2003. National Digital Preservation Initiatives: An overview of the developments in Australia, France, the Netherlands, and the United Kingdom and of related international activity. See: Appendices – References.

[4] OCLC/RLG PREMIS Working Group. 2004. *Implementing Preservation Repositories For Digital Materials: Current Practice And Emerging Trends In The Cultural Heritage Community*. See: Appendices – References.

The PREMIS survey was carried out to provide a background study to examine and evaluate strategies for managing and exchanging preservation metadata. The examination would provide the context for the development of a preservation metadata scheme and a data dictionary. These were published in May 2005.[5] The PREMIS survey offers the perfect layout for describing the current developments in the field of digital preservation. Its set of questions and choice of subjects was the result of intensive international cooperation and detailed discussions between experts in the field. Therefore it provides a thorough coverage of all aspects involved in current digital preservation activities. The sections on digital repositories and preservation strategies in this ICABS survey are modelled on the PREMIS approach. As this survey is based on these two previous reports it provides an adequate opportunity to visualise the recent changes and progress within the field of digital preservation.

Whereas Beagrie and PREMIS both focus on national libraries and other cultural heritage institutions, the ICABS survey only considers national libraries. This is because ICABS primarily serves this type of library. The emphasis is on the common activities of all national libraries: the handling of the electronic equivalent of print and born-digital materials. The handling of audio or visual equivalents, which is also an important task of some national libraries, falls outside of the scope of this survey.

Methodology

This report contains two parts. Part one provides an analysis of the current state of affairs at 15 national libraries. Part two contains the detailed overviews per library. Each overview is divided into four sections: General; Digital repository; Preservation strategies; and Current activities.

Within the overviews, the general section contains information on recent developments with respect to deposit laws for digital objects (online and offline); the embedding of digital preservation activities (building of the repository and research) in the library organisation, and how digital preservation activities are financed.

The second section, on digital repositories, contains information on: the status of the digital repository; services provided; depositing; software and OAIS; materials; metadata and metadata schemes; and access.

---

[5] OCLC/RLG PREMIS Working Group. 2005. *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group.* Dublin, O.: OCLC Online Computer Library Center Inc. See: http://www.oclc.org/ research/projects/pmwg/premis-final.pdf

The third section, on preservation strategies, describes which strategies are currently applied in the processing of digital objects and the plans for the future. The fourth section, on current initiatives, gives an overview of current projects and working groups in which the national libraries are involved or in which they participate (national and international projects) and a general impression of institutes active on digital preservation in other cultural heritage sectors (museums, archives and audiovisual institutions).

The fifteen libraries involved in the study are (in alphabetical order of country names) the national libraries of: Australia, Austria, Canada, China, Denmark, France, Germany, Japan, the Netherlands, New Zealand, Portugal, Sweden, Switzerland, UK, and USA.

Initially a selection of the national libraries to survey was made on the following basis: i) ICABS partners; ii) national libraries known to be active in the area of digital preservation. Obviously a selection inevitably leads to a somewhat subjective interpretation. The survey does not pretend to be complete and from a practical point of view is limited in size. A limit of fifteen libraries seemed feasible within the framework and time-span of this research.[6] In the next stage an overview was made of all libraries involved, based on the websites of the national libraries, the information provided by the research of Beagrie and the survey of PREMIS, and additional literature. All of the topics were then completed and updated using information gathered by phone or e-mail. The overviews were then sent to the contact persons in the respective libraries for commenting, correction and addition. Lastly the final version of the report was sent for review to the ICABS partners, the contact partners in the national libraries surveyed, and the authors of the CLIR and PREMIS reports.

The overview represents the current state of affairs in July 2005. In the analysis part, the libraries mentioned to illustrate the general observations are used as examples. This has been done to avoid endless enumerations of the information that can be found in the separate overviews per library.

Within the framework of ICABS, NLA also carried out a survey in 2004−2005. This concerned the availability of guidelines for digital preservation activities and various digital objects. This report will only be available online and a summary of it is included in the Appendices of this report.

---

[6] Other national libraries presently known to be active in the field of digital preservation but not considered in this study due to scalability and other reasons are, for example, the National Libraries of Finland, Norway, South Korea, Italy, Scotland, and Slovakia.

Practical definitions

In 2005, a wide range of concepts and definitions are still used in the field of digital preservation. Despite the many sources available, a standard definitive glossary has yet to be provided. This might impede communication and knowledge sharing. To illustrate how specific terms and concepts are used in this survey, the main basic concepts are explained below. These definitions are partly based on daily practice in the Digital Preservation Department of KB, and have been compared with and complemented by the definitions used in a selection of the existing glossaries.

The list only serves to provide a standard reference point for this survey and does not pretend to be complete or to form the basis for a final glossary. For practical reasons the list is subdivided in three themes: the discipline, the material and the system.

The discipline

Digital preservation or *long-term preservation* is the general term for all activities concerning the maintenance and care for/curation of digital or electronic objects, in relation to both storage and access. *Long-term* means five years or more; *short-term* is less than five years. Within digital preservation the main activities can be divided into *digital archiving* and *permanent access*.

Digital archiving means the process of backup and ongoing maintenance of digital objects and the associated software and hardware, as opposed to strategies for digital preservation. (Source: *DPC Handbook*).

Permanent access is usually paired with the term *digital preservation,* indicating that preservation is only half the battle. Within the digital environment, providing permanent access and adequate rendering of the digital object will be one of the greatest challenges, given the technological changes that have and will continue to occur. (Source: *CENDI report*).

Day-to-day activities are those activities concerned with the daily operational workflow of handling the digital objects in the framework of digital preservation within the library.

Preservation strategies (such as migration and emulation) are methods for keeping stored material permanently accessible. It refers to all techniques that provide more than would be obtained by merely storing the digital objects and never looking at them again.

Developing **tools** means, among other things, developing a device that provides a mechanical or mental advantage in accomplishing a task. Examples are: UVC, JHOVE, PANIC, Preservation Manager. (Source: Wikipedia).

The material

Digital objects is a general term used for the body of digital material that is subject to digital preservation, both *digital publications* and *digital records*; both *online* and *offline*; both *born-digital* and *digitised.*

Digital publications (or e-publications) are digital materials which have been released for public access (publication) and made available free of charge or for a fee.

Digital records refer to the type of electronic or digital materials the archival sector usually has to deal with. Digital records are created digitally in the day-to-day business of the organisation and assigned formal status by the organisation. They may include, for example, word-processing documents, e-mails, databases, or intranet web pages. (Source: *DPC Handbook*).

Online refers to digital objects that are connected to some larger network or system.

Offline An offline publication is not connected to or accessible through a network, but stored on a stand-alone carrier (CD, DVD, tape, optical, hard disk). It means that a digital object can be held in the hand and put on a shelf, without transforming it into any other form (paper or microfilm, etc.).

Born-digital refers to materials which are not intended to have an analogue equivalent of the object, either as the originating source or as a result of conversion to analogue form. (Source: *DPC Handbook*).

Digitised implies the transformation of the information of the original physical, analogue carrier into a digital form. Digitised refers to the mode of production.

The system

Digital repository or *electronic repository* is the system (or combination of systems) that provides long-term storage and preservation of and permanent access to digital objects.