United Nations
Educational, Scientific and
Cultural Organization

Information for All
Programme

PRESERVATION OF DIGITAL INFORMATION IN THE INFORMATION SOCIETY

Ministry of Culture of the Russian Federation
Federal Agency for Press and Mass Communications
Commission of the Russian Federation for UNESCO
Russian Committee of the UNESCO Information for All Programme
Interregional Library Cooperation Centre

Preservation of Digital Information in the Information Society

Proceedings of the International Conference
(Moscow, Russian Federation, 3–5 October, 2011)

Moscow
2012

Support for the publication: Ministry of Culture of the Russian Federation, Federal Agency for Press and Mass Communications

Financial support for the publication: Grant of the President of the Russian Federation for the Support of Culture and Art Projects of National Importance (the President's Executive Order of March 20, 2012 № 95-рп)

Compilers: Evgeny Kuzmin and Tatiana Murovana

Translators: Elena Malyavskaya, Tatiana Zverevitch, Tatiana Butkova and Ekaterina Komarova

English text edited by Natalia Bogdanova

Preservation of Digital Information in the Information Society. Proceedings of the International Conference (Moscow, Russian Federation, 3–5 October, 2011). – Moscow: Interregional Library Cooperation Centre, 2012. – 296 p.

ISBN 978-5-91515-046-9

The book comprises reports of the International Conference, *Preservation of Digital Information in the Information Society: Problems and Prospects*, which was held in Moscow, October 3–5, 2011, within the framework of the Russian chairmanship in the UNESCO Information for All Programme. For both Russia and UNESCO, this Conference was the first major international interdisciplinary forum on the subject.

The book covers general approaches to the study of digital information preservation issues and relevant activities in terms of various subject areas, as well as policies and education aimed at preserving digital information. National approaches, solutions, concepts in the field of the preservation of text-based documents and audiovisual materials are highlighted.

The book is meant for directors and staff of governing bodies, cultural, scientific and educational institutions; information officers, and for all those who take interest in the preservation of digital information.

Contents

Plenary Meeting. Information Preservation: Digital Landscape

Plenary Meeting. Policies and Education for Digital Information Preservation

Plenary Meeting. National Approaches, Solutions, Vision

Section 1. Digital Preservation of Text Documents

Section 2. Digital Preservation of Audiovisual Documents

Preface

Digital Information Preservation: Russia's Contribution into Building Knowledge Societies

The book comprises reports and articles by participants in the International Conference, *Preservation of Digital Information in the Information Society: Problems and Prospects* organized by the Ministry of Culture of the Russian Federation, the Federal Agency for Press and Mass Communications of the Russian Federation, the Commission of the Russian Federation for UNESCO, the Russian Committee of the UNESCO Information for All Programme, and the Interregional Library Cooperation Center and held in Moscow on October 3–5, 2011.

It was the first major international cross-subject forum on the digital information preservation ever, both in the history of the Russian Federation and that of UNESCO. It was hosted by Russia, since in 2010 the Russian Federation was unanimously elected Chair State of the UNESCO Information for All intergovernmental programme. The Conference became one of the most significant international events within the Programme framework. Both Russian authorities and the UNESCO administration, as well as that of IFLA, one of UNESCO's major strategic partners, attach great importance to the Conference.

Mr. Sergey Lavrov, Minister of Foreign Affairs of the Russian Federation, Chair of the Commission of the Russian Federation for UNESCO, Alexander Avdeev, Minister of Culture of the Russian Federation, Mikhail Seslavinsky, Head of the Federal Agency for Press and Mass Communications, Jānis Kārkliņš, UNESCO Assistant Director-General, and Ingrid Parent, IFLA President in Office, sent their welcome messages to Conference organizers and participants.

UNESCO considered the Moscow Conference the main preparatory event for the UNESCO World Conference on digital information preservation to be held in Vancouver, Canada, in September, 2012.

The Moscow Conference brought together more than 150 directors and experts of major libraries, archives, museums, research institutes, universities, international organizations, public authorities, mass media, publishers,

sci-tech information centers, ICT industries, and other agencies from 37 countries interested in the development of digital information preservation.

At the opening of the conference held at the State Tretyakov Gallery, the necessity of an active and detailed examination of digital information preservation issues was stressed by Vladimir Grigoriev, Deputy Head of the Federal Agency for Press and Mass Communications; Ekaterina Chukovskaya, State Secretary – Deputy Minister of Culture of the Russian Federation; Svetlana Zhurova, Deputy Chair of the Russian State Duma; Evgeny Kuzmin, Chairman of the Intergovernmental Council and Russian Committee of the UNESCO Information for All Programme; Sanjaya Mishra, Programme Specialist of the UNESCO Communication and Information Sector; Grigory Ordzhonikidze, Executive Secretary of the Commission of the Russian Federation for UNESCO; and Sinikka Sipilä, IFLA President-Elect, Secretary-General of the Finnish Library Association.

The participants in the Conference unanimously adopted the final document, *The Moscow Declaration on Digital Information Preservation*, which is currently one of the fundamental international documents in the area. It defines the subject situation and outlines priority measures to be taken in politics, awareness-raising, education, scientific research, economy, ICT industries, cooperation and coordination.

Participants' statements during the Conference showed the necessity of a further detailed examination of the matter; its outcomes should be reflected at the national and international levels within political and professional contexts. The speakers unanimously expressed their confidence in the importance and relevance of issues raised at the Conference.

The Conference has shown that traditional depositories of analogue information, i.e. libraries, archives, and museums, fail to cope with the task of digital information preservation, and they are the first to blow the whistle. There is no understanding, at the political level, of the urgent need to develop a new infrastructure for digital information preservation, most likely based on the traditional information media infrastructure, through modernization and strengthening.

It became obvious during the Conference that most states lacked legal frameworks that would make professional digital information preservation binding for relevant institutions and structures; likewise, they lacked coordinated policies that would result in the establishment of such a base. However, both developed and developing countries across the world are

in search of answers to the questions: what exactly and for what reason should be preserved; who, where, when, at whose expense and to which extent should preserve it?

The Conference has not only revealed the challenges. It has also opened up new professional approaches towards the preservation of rapidly growing digital information arrays and the ever-increasing multiplicity of information formats, types and varieties of assets, their life span decreasing due to persistent changes in information technologies.

Participants repeatedly proposed that the UNESCO Information for All Programme, which lists information preservation among one of its five priorities, should be used as an international cross-disciplinary and inter-institutional platform to develop a framework for policies, hold discussions, and share experience in the area.

The UNESCO Information for All Programme is the only international programme that specifies and addresses, in a comprehensive way, problems of that are vitally important for building a pluralistic, just and inclusive information society, such as information availability, information preservation, information ethics, information literacy, information for development and multilingualism in cyberspace. All of the above mentioned problems, though each to its own extent, are faced by both developed and developing countries, for North and South, East and West. Even in the "hard-copy" era, none of the countries could successfully tackle this problem on its own. There was an urgent need for international cooperation that led to general international approaches and standards and the international division of labor. Today, in the "digital" world where information easily travels across state borders and goes beyond the limits of national jurisdictions, cooperation becomes all the more sought for. To address these very problems, UNESCO's intergovernmental Information for All Programme was launched.

The Programme outlines most common problems of the information society and helps those interested address the existing and continuously emerging dangerous discrepancies, if not in practice, than at the consciousness level at least.

The main objective of UNESCO's Information for All Programme is to help member states develop policy frameworks for the establishment of knowledge societies and facilitate its implementation in full and in specific subject areas. That is why the Programme should, and can, succeed where

specialized international agencies, UNESCO member states governments, governmental, non-governmental and commercial organizations fail to cope due to various peculiarities or restrictions.

The Programme's priorities set out, on the one hand, emerging burning global problems, and, on the other hand, define those areas of growth that would help educational, scientific, cultural, information and communication institutions better adapt to the challenges of a new information environment and contribute in a more efficient way to the progress of each country and the whole human civilization.

The Programme's priorities, as information policy elements, are interrelated. Universal access to quality up-to-date information is impossible without information preservation, information literacy, effective and development-oriented information policies, elaboration of, and compliance to, the principles of information ethics. Digital information preservation cannot be ensured without science-based policies, information ethics, and the appropriate level of information literacy. Skills and knowledge in the information search and preservation area should become an integral part of information literacy. To encourage information literacy, relevant promotion policies are needed in national educational systems, and so on and so forth.

The Information for All Programme aims at a comprehensive exploration of cyberspace and its impact on individuals, communities, nations and mankind, on the whole. It is future-oriented; it aims to foresee the potential consequences of today's policies and the ways they are implemented.

We hope that facts, opinions, views and experience collected in the book will contribute to the comprehensive development and improvement of activities in the area of long-term digital information preservation, to the creation of deliberate integrated policies at both national and institutional levels.

Vladimir Grigoriev
Deputy Head
Federal Agency for Press and Mass Communications
of the Russian Federation

Evgeny Kuzmin
Chairman
Intergovernmental Council and Russian Committee
of the UNESCO Information for All Programme
President, Interregional Library Cooperation Centre

Messages to Organizers and Participants of the International Conference, Preservation of Digital Information in the Information Society: Problems and Prospects

Message from Mr Sergey Lavrov, Minister of Foreign Affaires of the Russian Federation, Chairperson of the Russian Commission for UNESCO

I am glad to welcome the organizers of and participants in your Conference.

The forum's topic is of particular relevance. Progress in such areas as the implementation of citizens' right to education and decent life, the building of a democratic knowledge-based society, the achievement of modernization imperatives that are on nearly every state's agenda across the world, depends on finding well-balanced solutions in the sphere of information dissemination and use. The overall task is to ensure international information security – and we consistently come out with relevant initiatives in the United Nations.

Tangible results have been reached on a number of issues discussed, including digital information preservation, much to the credit of those experts who are here today. The Russian Committee of the UNESCO Information for All Programme greatly contributes into this work.

With information flows growing, information and communication resources expanding, the preservation of worthwhile information, the development of effective regulations, rules and methods of information selection are gaining exceptional importance. The preservation of world intellectual and cultural heritage for future generations, which is directly linked to the preservation of multiple cultures and civilizations on the planet as a key aspect of shaping modern world order, is of special significance.

I am sure that the Conference will contribute to strengthening cooperation in the area of digital information preservation and become another step towards a new information society.

I wish you fruitful work and all the best.

Sergey Lavrov

Message from Mr Alexander Avdeev, Minister of Culture of the Russian Federation

Dear participants, organizers, and guests of the International Conference, *Preservation of Digital Information in the Information Society: Problems and Prospects*,

The problem of rapidly growing digital information content and providing access to it across the world is understood as a serious challenge to the development of human civilization. We can see that the problem is growing more complex and urgent, its scope increasing, since information content created and existing in digital form only, has exceeded by far the volumes of information recorded on analogue media and stored in libraries and archives. This problem has been insufficiently studied so far due to its novelty.

To preserve information, mankind has created special institutions to store memory – first of all, libraries and archives. To preserve digital information, we need as reliable an infrastructure as the one that helps us preserve information on paper. What will serve as a basis for this future infrastructure? Will that be, again, libraries and archives extensively modernized by means of information technologies?

Here we face numerous unsettled problems – of ethical, legal, or worldview nature rather than of technological one. No universally recognized concept of long-term digital information preservation has been shaped yet. There are no answers to the questions of what should be preserved from the immense and expanding digital universe, who and why needs it today, and who and why will need it tomorrow, and who is to achieve it all. Who will take over from traditional libraries and archives this heavy responsibility before the user, the state, the world, before history? Only one thing is self-evident today: if immediate policy and practical measures both at the national and international levels are not taken, we will soon find out that huge assets of highly important, useful and essential information have not been collected, duly preserved and, therefore, are irretrievably lost for future generations.

Russia has been consistently supporting UNESCO's efforts to find solutions to these problems at the international level, in particular, within the framework of the intergovernmental Information for All Programme. With Russia's current Programme chairmanship, holding an international conference on digital information preservation is a timely and useful

action. It is essential that not only traditional preservers of information – librarians and archivists, but also museum workers, information producers – book and periodical publishers, information distributors, representatives of mass media, and government officials who engage in developing information society policies at national and international levels, take part in the Conference. It is essential that the Conference has brought together representatives of both developed and developing countries.

We hope that the Conference will help us reveal urgent problems and find new professional approaches to long-term digital information preservation, contribute to international innovative experience exchange and intensification of efforts applied by cultural, research, and educational institutions in the area.

I wish all the participants in the Conference prolific professional and warm human contacts!

Alexander Avdeev

Message from Mr Mikhail Seslavinsky, Head of the Federal Agency for Press and Mass Communications of the Russian Federation

Dear participants, organizers and guests of the International Conference, *Preservation of Digital Information in the Information Society: Problems and Prospects*,

On behalf of the Federal Agency on Press and Mass Communications, let me welcome you in Moscow at this most significant gathering which is held under Russian chairmanship in the Intergovernmental Council of the flagship UNESCO Information for All Programme. This is the first major interdisciplinary forum on the issues of long-term preservation of digital information in the history of modern Russia and, as far as we know, in the history of UNESCO.

The issues you are going to discuss are relevant not only for those who have been dealing professionally with the provision of access to and preservation of information and knowledge, and who are now facing new problems of digital information preservation in an essentially new information environment; they are relevant for all intellectuals. This is why your conference has aroused such interest and brought together almost 150 directors of the largest libraries, museums, archives, publishing houses, and mass media, as well as representatives of government bodies, business and civil society from 40 countries.

We believe that this conference will give us an opportunity to discuss new professional approaches to the preservation of rapidly increasing digital information arrays in the context of quickly developing information technologies, with the diversity of formats multiplying and the lifetime of digital documents reducing due to constant changes of technologies.

Concept approaches and facts to be presented at the conference will undoubtedly help libraries, archives, museums TV, radio, and film depositories and all other institutions preserving cultural heritage in their work and contribute to designing concepts of national digital information preservation programmes. Such programmes are critical for the overwhelming majority of countries and should become an essential element of national strategies for the development of information and knowledge societies.

I wish to the participants and guests of the conference fruitful discussions and a pleasant in Moscow!

Mikhail Seslavinsky

Message from Mr Jānis Kārkliņš, Assistant UNESCO Director-General for Communication and Information

Ladies and Gentlemen,

It is a great pleasure for me to address the conference *Preservation of Digital Information in the Information Society*.

Digital records are today the main source of acquiring knowledge and are an important part of the essence of the emerging knowledge societies.

The increasing availability of broadband Internet access, mobile networks and a plethora of data capture and content creation devices offer new possibilities for making and sharing digital information across nations and disciplines as well as stimulating information re-use and re-purposing and the creation of new information and knowledge.

Digital documents facilitate collaborative and more efficient work-flow processes that provide unprecedented opportunities for bringing to bear new and more powerful approaches to problem-solving. They therefore better position us to expand and realize human potential and contribute to improved prospects for achieving peace and sustainable development.

While this growing trove of digital records is central to current and future human progress, the permanence of these digital records is uncertain. Without assurance of permanence, authenticity and integrality of our digital records, this virtual cycle of knowledge-lead progress may halt or even be reversed.

In 2003, UNESCO, through its Charter on the Preservation of the Digital Heritage, sought to mobilize global action aimed at raising awareness and creating a critical mass to begin addressing the complex issue of ensuring the preservation of digital documents. However, a UNESCO survey conducted in 2009 revealed that globally there is still a low-level of awareness about the need for digital preservation and the implications that this has for our civilizations, especially given our dependence on digital content. Few governments in either developed or developing countries have as yet identified this challenge as a priority, or taken the steps of elaborating strategies for preservation and permanent access to digital heritage.

Digital content is also not defined or confined by traditional notions of space and place and impinges upon a range of trans-border interests, legal, ethical, social, political, technical, economic and cultural factors amongst others. Addressing digital preservation requires therefore, multi-disciplinary expertise as well as a multi-sectoral and a global multi-stakeholder approach.

Your conference is thus an especially important preparatory event to support UNESCO's broader strategy of raising awareness and building global partnerships. It will also stimulate the reflections on the technological, legal, cultural, economic, political, security and other dimensions as we build up to UNESCO's international conference on Digital Heritage Preservation that will be held in Canada in September 2012.

On behalf of UNESCO, I extend our deep appreciations to the Inter-regional Library Cooperation Centre and the Russian IFAP Committee for the successful organization of this major event and for their role as an important and active cooperating partner in the preparation of next year's international conference on digital preservation.

I wish you fruitful deliberations and look forward with great interest to the outcomes of the conference.

Jānis Kārkliņš

Message from Ms Ingrid Parent, IFLA President

Dear representatives of the UNESCO Information for All Programme, Dear Representatives of the Russian Committee of the UNESCO Information for All Programme, Dear Representatives of the Interregional Library Cooperation Centre, Dear Conference Organizers, Dear Chair, Dear Participants,

From Vancouver, Canada, I would like to express warm words of greetings to all of you, participating today in the UNESCO IFAP International Conference "Preservation of Digital Information in the Information Society".

It would have been a pleasure and an honour for me to participate in this International Conference, and in that way to represent through IFLA, the international library sector here, but unfortunately, other obligations in Canada caused an overlap.

The theme of the UNESCO IFAP international conference is very well chosen. Digital information is not only transforming the society we live in, it is also transforming the core business of almost every library in the world. The future of libraries is digital. Librarians are moving into digital collections, with eBooks, with databases, with all kinds of new digital materials. Probably, in a couple of years there will be digital works that we have not even thought about today.

For libraries it is important to remain focused on new forms of digital materials. What is the best way to collect the digital materials? How can libraries assure ongoing access to the digital information? What digital materials should be selected for preservation and how should we best preserve that material – not only for the user of today, but also for the user of the future.

Libraries in the digital age have a great future.

For my IFLA presidency I have chosen a leading theme that hopefully guides libraries through the digital future. It is: Libraries: a Force for Change. Libraries play many roles that benefit individuals, communities and society in general. To realize their full potential, libraries must provide inclusive and transformative services, innovate and forge new collaborative alliances.

IFLA, the international Federation of Library Associations and Institutions acts as the trusted global voice of the library and information community, and drives equitable access to information and knowledge for all. Digital content is one of the Key Initiatives for IFLA to work on in the coming year.

I am looking forward to UNESCO's Global Conference on Digital Preservation, which will take place in my hometown Vancouver, in September next year. IFLA will closely cooperate with UNESCO in preparing this conference. I am confident that the outcomes of today's International IFAP Conference will return on the agenda in 2012.

I wish you all a very successful Conference.

Ingrid Parent

First Plenary Meeting

Sanjaya MISHRA
Paris, UNESCO

Programme specialist
UNESCO Sector for Communication and Information
(Presentation on behalf of the UNESCO Knowledge Society Division)

Preservation of Digital Information

The volume of digital information is increasing at an extraordinary rate. Access to quality information is a key driver of economic growth and social wellbeing in the contemporary world. Our intellectual capital is increasingly at risk by the volatile nature of the digital objects and rapid technological developments. The exponential growth of digital contents presents a novel challenge to the society to specially care and make systematic efforts to preserve these for the benefits of the future generations. Let me now present a brief picture of the massive amount of digital content that are produced every day.

According to an estimate by IDC, in 2011, the amount of information created and replicated will surpass 1.8 zettabytes (1.8 trillion gigabytes) - growing by a factor of 9 in just five years. In August 2010, Google estimated that there are about 130 million books (129, 684, 880 to be precise) ever published in the entire world. But, if the information created only in 2011 is transformed into books, the total number would be a staggering 470 milliion books (a rough calculation based on 300 pages per book). However, much of these 1.8 zettabytes would not pass the test for getting published in the normal channel. Yet, these information are digitally available, though may not be fully accessible to all. According to R.R. Bowker, in 2009, of the 1.3 million books published in the United States 77% were non-traditional publications that include primarily e-books and on-demand books. There are over 25,000 journals in science, technology and medicine and the Ulrich's International periodical directory includes over 250,000 titles. The Directory of Open Access Journals (DOAJ) lists over 7000 journals available in digital format with access to over 633682 full-text articles. While these are the normal library materials, the digital video format has been contributing a huge amount of resources for education and learning. For example, the YouTube alone boasts of having more than 13 million

hours of video uploaded during 2010 and it is said that '48 hours of video are uploaded to it every minute'. Users upload the equivalent of 240,000 full-length films every week. There are more than 750 million active users of Facebook, of which more than 250 million accessing Facebook through their mobile devices. This is not surprising, as there are over 5.3 billion mobile phone users at the end of 2010 constituting about 77% of the world population. While the use of mobile has increased and users use Internet through mobile devices, as of March 2011, there are over 2 billion Internet users (2,095,006,005), measuring up to 30% of the world population. On an average 140 million tweets are sent per day in Twitter. As part of its digital preservation programme, the Library of Congress has started also preserving these tweets that has "extraordinary potential for research into our contemporary way of life". The vastness and complexity of issues can well be imagined from these data. UNESCO's Charter on the Preservation of Digital heritage (2003) and the Memory of the World (MoW) Programme started in 1992, address the complex world of preservation and access to information and knowledge.

While the perception of value of digital information may be subjective, digital access to content increases the autonomy of individuals as information-seekers. It may also be difficult to put a clear line of distinction between what is significant and what is not as cultural heritage, but if we consider knowledge as living and growing, then it is important to be 'openly available' to be searched, retrieved, used, re-used and replicated to produce derivatives. Digital cultural heritage is expressed though social, cultural, political, and economic activities undertaken using digital technologies, and it may always be noted that what is considered trivial today may become a valuable object in future. Therefore, preservation of institutional memory of the world is important to prevent collective amnesia of countries and civilizations. Much of the world heritage has disappeared in the past and the remainder is endangered because of ignorance, neglect or deliberate destruction, as well as from natural disasters, chemical decay or technological obsolescence. Increasing access to information and educational materials is at the heart of UNESCO's activities to promote sustainable development and build peace in the minds of men and women. The networked technologies have become a key component of UNESCO's activities in this area to bridge the digital divide and improve access to digital information.

While it is important to digitize printed resources, it is also important to preserve many resources that are created/generated digitally, for example the ebooks and for that matter the books published only on demand. UNESCO Charter on the Preservation of Digital Heritage (2003) recognizes that information and creative expression are increasingly produced, distributed, accessed and maintained in digital form, creating a new legacy – the digital heritage, and these must be preserved. UNESCO, by virtue of its mandate and functions, has the responsibility to:

a. promote the implementation of the Charter within the United Nations system and by intergovernmental and international non-governmental organizations concerned with the preservation of the digital heritage;

b. serve as a reference point and a forum where Member States, intergovernmental and international non-governmental organizations, civil society and the private sector may join together in elaborating objectives, policies and projects in favour of the preservation of the digital heritage; and

c. foster cooperation, awareness-raising and capacity-building, and propose standard ethical, legal and technical guidelines, to support the preservation of the digital heritage.

UNESCO's Charter on the Preservation of Digital Heritage considers that many digital materials which include texts, databases, still and moving images, audio, graphics, software and web pages, among a wide and growing range of formats have lasting value and significance, and therefore constitute a heritage that should be protected and preserved for current and future generations.

The Memory of the World (MoW) Programme started in 1992 as a platform to handle the problem of falling in the trap of collective amnesia intends to:

- preserve the world's documentary heritage.
- promote universal access to documentary heritage.
- raise global awareness of the significance of documentary heritage.

The Programme is grounded in the belief that the world's documentary heritage belongs to all women and men. As such, it should be preserved, protected and permanently accessible to everyone. To attain its fundamental objectives, the MoW functions through a three-tiered structure operating

at the international, regional and national levels with core elements in common. Strategic guidance is provided by an International Advisory Committee (IAC) composed of 14 members appointed in their personal capacities for a four-year term by UNESCO's Director-General.

The Register of the Memory of the World Programme is an outstanding example of cooperation from Member States beyond the convention or inter-governmental programme approach. The programme has been successful in safeguarding the documentary heritage, both analogue and digital mostly due to the Register, its most successful and visible element. So far there are 245 items that have been judged to be of world significance listed on the International Register. The 'significance' of a document can comprise both the content and the carrier on which it is recorded. Carriers can be as lasting as a clay or stone tablet or as ephemeral as a website. Documents can take the form of a film, a musical opus, or a narrative tapestry and be recorded on plastic, papyrus or palm leaves. Decisions to recommend inscription on this Register are taken by the IAC at its biennial meetings and are, in turn, based on the recommendations and assessment undertaken through the work of the Register Sub-committee and its panels of experts covering the various disciplines.

In partnership with the Library of Congress, UNESCO launched The World Digital Library (WDL) in 2009 as a platform that offers advanced search and browse features, including a 'listen to this page' for the visually impaired, in seven languages: Arabic, Chinese, English, French, Portuguese, Russian and Spanish, with content in more than forty languages. Rare books, manuscripts, maps, photographs, sound recordings, and films of cultural and historical significance are among the items available for browsing.

The major objectives of WDL are to:

- Promote international and intercultural understanding and awareness
- Expand multilingual and culturally diverse content on the Internet
- Provide resources to educators and contribute to scholarly research
- Build knowledge and capacity in the developing world

Several Memory of the World register items are already in the World Digital Library as there has been tacit agreement that items on the MoW Register should be included on the WDL. This will allow items on the

register to be easily accessed in a dynamic and interactive way in addition to being easy to use, comprehensive, and of a high quality.

Preservation is a cost intensive as well as political activity. Since the resources to preserve everything is limited, selection of documents becomes a critical issue and criteria based approach is recommended. The next problematic issue is sustainability. While acquiring deposits of relevant materials is not a problem, cataloguing and preserving them requires significant funding. Technology keeps on advancing every day and therefore, we have many different standards and software for undertaking long term preservation. Considering the problem of adequate funding for institutional preservation programmes, the most feasible and efficient way to manage long term preservation seems to be network-based. However, such an approach requires collaboration from project partners, and in order to be successful it is important to:

- Choose appropriate partners who can contribute to the success of the preservation project
- Understand what each partner wants to achieve collaboratively
- Share interests and commitment to preservation through discussion and demonstrable action
- Allocate sufficient resources by collaborating partners
- Set realistic goals, and monitor progress regularly.

In today's world of technological developments, it is logical for us to work collaboratively to gain both efficiency and effectiveness. No single organization can undertake the responsibility to preserve the world's digital information. It is in this context the present event, and the UNESCO International Digital Heritage Preservation Conference planned in 2012 in Canada are important as platform for knowledge sharing and opportunities of collaboration. I, on behalf of the Knowledge Societies Division (KSD), join our Assistant Director-General in congratulating the Inter-regional Library Cooperation Centre and the Russian IFAP Committee for organizing this conference as a lead-up to the international conference. I am sure the deliberations here will be useful to the planning and organizing of the event by UNESCO. We look forward to the success and outcomes of this event.

Evgeny KUZMIN
Moscow, Russia
Chair
Intergovernmental Council, UNESCO Information for All Programme
Chair
Russian Committee, UNESCO Information for All Programme
President
Interregional Centre for Library Cooperation

Information Preservation: Past, Present and Future

Introduction

To store and preserve information has been one of the principal human duties since times immemorial, though few people have been aware of it. Information and communication technology made this job incomparably more complicated because we have to deal not only with traditional analogue media, which make information storage and preservation hard and expensive. Now, we have also to store digital information. Some of it is created in the electronic (digital) form, while other appears with digitizing information on analogue carriers. These are texts, static and moving two- and three-dimensional images, sound recordings, electronic catalogues and other databases, computer programmes, e-books, audio books and other multimedia editions, media art works, web pages, portals, websites, blogs and microblogs, mobile content of many kinds, etc.

Even today, the amount of digital information dramatically exceeds that of information on analogue carriers. It grows further in an avalanche as communications become ever more extensive and complicated, ICT and relevant gadgetry is used everywhere and steadily develops, and the means of production, dissemination, aggregation, storage and preservation of digital information are progressing.

Long-term preservation of socially important digital information resources created by institutional agencies and private persons in most diverse formats and on various carriers is a new global problem – an extremely topical and little-studied one. Russia encountered it later than the developed Western countries.

The preservation of digital information resources aims to guarantee their lasting, if not eternal, accessibility and maintenance of all semantic and functional characteristics of the original material, and opportunities for search, presentation and interpretation for broader access to them and their use by the present and future generations.

Even superficial study of the problem of digital information storage and preservation shows its many aspects – philosophical, socio-cultural, political, managerial, legal, ethical, economic, psychological, methodological, tactical, organizational, technological, pertaining to the personnel, and others. In this, valuable digital information is lost not so much due to technical difficulties and drawbacks – they are mostly evident, and can be corrected. The basic reason for such losses is that society as a whole fails to see the problem and realize its scope, acuteness and versatility. Such realization is absent everywhere, at all social levels, including political, managerial and institutional. As the result, the preservation of digital information is not duly organized at every stage of its life, though its preservation should start even before its creation – as early as it is planned. That is why there is no adequate body of specialists able to develop a comprehensive idea of necessary activities at the present initial stage, launch relevant efforts and efficiently implement them later.

To evaluate the scope and acuteness of the problem of long-term digital information preservation, we should turn to past experience and analyze the activities for printed matter storage and preservation on paper because man has achieved better results in that sphere than anywhere else. We will use libraries as our example. Though not only libraries but also archives and museums preserve information on paper, libraries store, concerve and preserve it and provide general access to it better than any other memory institution – that is, one established to preserve cultural, research and educational heritage. Neither archives nor museums have whatever tangible prospects – at least in Russia – to compete with the smooth library system for information monitoring, selection, acquisition, collection, arrangement, procession, storage, concervation and provision of access to it. In the foreseeable future, there is no chance to develop a serious programme to rival the thoroughly elaborated and structured Russia's National Programme for the Preservation of Library Materials (established in 2000, updated in 2011).

Past Experience

It took man more than two thousand years to learn to collect and preserve information on analogue carriers purposefully and efficiently for everyone to use it, and pass it for temporary use – particularly from one repository to another – with the slightest possible risk of loss. As this experience was acquired, information keepers and the entire community had to pose essential questions and every time to seek answers to them anew: Why do we store, conserve and preserve it? Who should do it? How was it preserved before our time, and what should we improve? Where to keep information? What conditions does it need? Who should provide them? Who should pay for it? What should we preserve – the text or the carrier, or both? Who and how should train professional information keepers, and what curricula should be used? Who should elaborate, launch, evaluate, improve and implement study programmes?

Ancient repositories, such as the Library of Alexandria, initially accumulated manuscripts and museum treasures. That was not hard to do because such values were concentrated in possession of the aristocracy, and they were rather scanty – at any rate, their amount was manageable.

Later on, with the development of communities, the number and diversity of artefacts grew. That was when specialization of repositories – libraries, archives and museums – was introduced. From then on, they developed as independent institutions, and each underwent major differentiation as they evolved.

By the mid-20th century, every developed country possessed a ramified system of comprehensive and specialized libraries and archives. They divided basic functions between themselves. Libraries collected and stored information – mostly published, replicated and potentially available for everyone. Archives specialized mostly in unpublished and non-replicated information, access to which was frequently limited by these same archives or their founders – the state, in most instances. Some countries' national libraries play also the role of national archives preserving the basic and most valuable part of the national documentary heritage.

As the result, we have presently a vast number of types of libraries, museums, archives and other memory institutions – particularly film, television and radio depositories – oriented on the collection of different types/classes of information for varying purposes and various user categories.

Every major memory institution presently possesses elements of other institutions' structure. Thus, many libraries have archives and museums of their own, while museums and archives have independent library branches. For instance, Russia's State Hermitage Museum possesses a huge and famous library on the arts, while the Russian State Library has established its Book Museum. Its Manuscript Department is one of the best-known Russian depositories, which preserves the heart of our national cultural heritage. The library also has a large archive of its own. No less celebrated is the Manuscript Department (archive) of the State History Museum. The State Archive of the Russian Federation possesses a large and valuable library. The St Petersburg Theatre Library is a library, museum and archive in one. Major theatres, like the Bolshoi or the Arts Theatre, have museums, libraries and archives of their own to store video and television recordings of their productions, and verbal transcripts and sound and video recordings of rehearsals. All such information is of tremendous cultural historical value. Photo, cinema, video and sound recordings are presently stored not only in special repositories but also in most diverse institutions, whose archives are subordinate to the State Archive Fund of the Russian Federation and ruled by the federal Law on Archives.

Precious cultural heritage is also preserved in private archives, especially those belonging to outstanding researchers and cultural and political activists. The most valuable of such archives eventually become part of government memory institutions for the nation and the whole world to take pride in.

The ideas of what libraries should collect, store and preserve changed with the evolution. They initially possessed only the most valuable books and manuscripts. At a certain stage of the development of civilization, man realized that libraries should collect and store not only the greatest treasures but also all published documents without exception, including newspapers and magazines and, later on, all media that appeared – sheet music, maps, prints, posters, photographs, records, audio and video cassettes, CDs, etc.

To help the larger libraries with collecting all printed matter available, the world's first law on legal deposit was passed in England in the 17th century. Now, almost all countries have similar laws. The institution of national library was invented, and such libraries are law-bound to collect and eternally store all materials printed in the country.

The establishment of national library associations improved library work to collect, arrange, store and offer to the public everything published not only in a particular country but also abroad. Such associations allow library managers and leading experts to formulate, discuss and tackle the problems they share. Later on, librarians of the world established the International Federation of Library Associations and Institutions (IFLA). There are not only comprehensive national and international associations, such as IFLA – some bring together school, university, research, technical, musical, medical, agricultural and other specialized libraries.

Libraries and, to a smaller extent, archives actively interacted at the national and international level. Russia was among the world's leaders in this respect in the 20th century's second half, and disseminated the methods and examples of its activities to half the world or even broader.

As professional communication developed, a great number of national and international standards were elaborated for the various sectors of professional activities. For instance, the size and format of bibliographic cards was standardized on the basis of international bibliographic description standards, and the production of catalogue boxes of similar sizes was started all over the world to keep those cards in.

Archives and museums resemble libraries in many respects where the development of analogue information storage is concerned.

Institutions, traditions, and forms, methods and standards of activity also emerged and developed in the publishing business, which produced the content that libraries collected and stored. National and international systems and structures were established to register and bring into system all publications in the world according to diverse characteristics – in particular, such agencies as ISBN, ISSN, ISMN, etc. National and international standards were accepted for imprints. Every item of printed matter received such an imprint specifying the author, translator, editor, proof reader, designer and publisher; the place of publication, the number of copies, the copyright holder, the theme, etc.

All those activities in the library, archive, museum and publishing spheres had a firm academic basis, and involved linguists, methodologists and other researchers, alongside custodians.

As the volume of paper publications grew, national library depositories began to have problems with storing new acquisitions. Libraries received

new specially equipped premises but even these could not hold the entire amount of the national printed matter and the most significant foreign publications. Storage crises are an intrinsic part of the development of all major research libraries.

To cope with the crisis, the leading countries' national *depository* libraries (which collected and stored the entire range of their countries' publications) called in the 1980s to establish libraries of a thoroughly new type – *repositories*, to which they began to pass for storage the publications in smaller demand. The load on the depositories of national, paranational and other major academic libraries was thus reduced. The responsibility for eternal storage of the regional and local press began to shift to regional, specialized and even local public libraries that gradually became local publication depositories. This change of library functions took place literally before our eyes – within the two previous decades.

The training of certified specialists – professional librarians and archive and museum experts – began in the developed countries toward the end of the 19th century. Special institutes or university departments were established with specific duties, methods and methodologies, equipment, curricula, teaching aids and educational standards. They also possessed a meta-structure – academic councils on diverse library disciplines, and regularly published teaching recommendations.

Not only developed countries but even the poorest of developing ones employ only certified professionals to form and maintain the library stock. These are not only librarians but also researchers from many spheres. Chemists, biologists, engineers and others are responsible for storage and preservation alongside librarians.

A great many professions are developing in librarianship – specialists responsible for collection, cataloguing, servicing, conservation and restoration, experts on chemical and biological damage to paper, etc. Libraries now have departments for security, and engineering, technical, chemical, biological and organizational protection. Standard instructions and rules have been elaborated to clearly delineate many kinds of responsibilities.

The economics of librarianship is a rapidly developing new discipline, which includes the economics of stock preservation. Requirements have been elaborated for time input and outlays – particularly concerning the bibliographic treatment and storage of documents. We know now the costs of any library activity. We know in detail how a depository should be built

and equipped, what employees should know, how to arrange their training, who should draw their curricula and on whose order, etc, etc.

There are commercial companies specializing in high-tech equipment for library stock preservation, which they design in cooperation with library experts and on library orders.

As follows from the above, goals were set, challenges prioritized, and responsibilities distributed in the library storage and preservation of printed matter on paper. Benchmarks were also set, and cooperation organized between all involved. Some selected content, prepared it for printing, and later distributed it, while others selected its most valuable parts for subsequent short- or long-term preservation, catalogued it, and organized its storage and access to it.

A sophisticated infrastructure was thus formed for the library storage of analogue information. Alongside libraries, it included:

- research institutes and methodological centres;
- higher and secondary vocational schools;
- a ramified multi-level network for postgraduate training;
- a national and local managerial network for the entire library and information sphere;
- national and international professional associations of a broader and a narrower scope;
- a system of professional periodicals (the United States, for one, had about 600 such publications in the early 1990s);
- a system of annual professional conferences, meetings, seminars and roundtables at the international, national, regional and municipal levels;
- the library and archive equipment manufacturing industry;
- standard-setting agencies;
- traders delivering books to libraries; and
- manufacturers of software and other equipment that libraries need.

Information storage was permanently upgraded. In particular, ambitious efforts were made to specify and prioritize:

the specialization of library stock acquisition;

- the methods and criteria of selecting information sources;
- the methods of information description, accounting and storage;
- the goals, methods and curricula of expert training and retraining;
- the policy toward memory institutions;
- the goals posed them by the community;
- conditions created for memory institutions – particularly their funding; and
- the forms of public accountability of all relevant institutions.

Archives and partly museums were in a similar situation.

Libraries suffer tremendous losses even though their stock is preserved by top-notch experts. Rare books are stolen or destroyed by floods, earthquakes and other calamities and man-caused disasters, such as wars, coups, fires and industrial accidents. The wrong temperature, humidity, lighting and chemical composition of the air in depositories also damage books, just as insects and microbes, to say nothing of readers – a majority of books do not withstand more than 15 readings due to organic acids and microbes on readers' hands. Books are also naturally worn with time.

The Ministry of Culture of the Russian Federation responded to major problems with library storage in the 1990s by beginning to draw the comprehensive National Programme of Preservation of Library Collections as an inalienable part of the national cultural heritage and information resources, which was adopted in 2000. In the ten years that followed, the programme played an incomparable role in changing the concept of library stock preservation and library managers' attitudes to them. It also caused major changes in librarians' professional mentality. The programme became an essential part of national library policies, and the forms and methods of library practice. It was Russia's first endeavour to demonstrate the problem of library stock preservation in all its scope, depth and acuteness. It analyzed and prioritized that problem. Despite token funding, the

gradually implemented programme brought enviable fruit. The library world is growing to realize that stock preservation is an *unbroken process* involving all librarians, and that momentary measures do not suffice, however expensive they might be.

But Russia's National Programme of Preservation of Library Collections does involve only traditional material media.

What's Going On?

Despite the excellently drawn and successfully implemented National Programme of Preservation of Library Collections, it no longer suffices as the entire information environment is growing more sophisticated, with many new kinds and types of information and its carriers – I mean not only electronic information created by libraries as they digitize their stock but also purely virtual information that appears and exists solely in cyberspace, i.e. born-digital information. Here, we should concentrate on socially important portals and websites accumulating information essential for the public, and the most interesting blogs, including ones by major political and community activists, as well as their electronic archives, correspondence, etc.

A new sphere of library activity appeared in the previous 15-20 years. That was the acquisition of information on electronic carriers. It was rather soon complemented or even replaced by the acquisition of licenses for access to electronic content on other agencies' remote servers within a time limit stipulated by the license. National libraries became responsible for the preservation (cataloguing, restoration, microfilming and digitization) of their country's heritage and information about it, and its presentation in the Internet.

For lack of room, major research libraries all over the world began to discard foreign paper periodicals which they had been collecting for centuries, and which were justly regarded as unique and priceless information sources. Certain libraries made their microfilms/microfiches and/or digitized them as recently precious originals were scrapped.

Many libraries don't do even this. They merely wait for other countries to digitize their own national press materials and provide paid or free access to them.

Memory institutions also face another problem – electronic information is rapidly becoming ever more sophisticated. One document might contain various multimedia elements – text, sound and images, static and dynamic (video or animation), a database, hyperlinks, etc, etc. Previously, different memory institutions were responsible for the preservation of each element. They used different approaches and methods. In this, we should realize that the preservation of video and sound recordings on analogue media has a far shorter history than the storage of texts on paper. Audiovisual objects appeared as late as the end of the 19th century on fragile and short-lived carriers, and there was no experience of their lasting and successful storage and preservation.

The sources of audiovisual information became ever more diverse. Major government and private television, radio and film studios were its basic suppliers in the pre-Internet era, and it was stored by cultural, research and educational institutions of various kinds – special archives, sound and video recordings publishers, recording studios, libraries, museums, research institutes, schools and universities, theatres and concert agencies. The number of producers and keepers of audio and visual information was finite and manageable.

Institutional structures (content providers, electronic publishers, electronic media, electronic content dealers, and all types of government, commercial and public organizations) and private persons added to the number of manufacturers and at the same time keepers of digital audio and visual content in the age of digital technologies and Internet. The number of producers and keepers of audio and visual content has not merely grown – it is barely accountable. In fact, the entire humankind can join them.

In parallel with the growing sophistication of information objects and an unlimited numerical increase of their producers and keepers, who expressly orient on short-term storage, the communication process itself has changed. It no longer results in the appearance of one information object (text, sound or video recording, etc) but breeds a vast complex of interconnected information products whose content undergoes major changes with time.

Thus, a radio programme broadcast 15 years ago was subsequently stored in the special radio archive – that is, one particular information object was stored by one particular memory institution. Now, a radio broadcast is accompanied by a sound and/or video recording that is cast

on the Internet in a parallel arrangement. A verbal recording, photos, extra information and users' comments may also appear there. All this information changes with time: new comments appear and disappear, removed by the author or the moderator, new links are added, etc. Thus, communication produces a major complex of heterogeneous information objects of diverse origins and with varying life cycles. Who is to provide their long-term preservation – if at all? Should it be a radio or television archive, or an archive of television and photo documents? Or again, a library or museum? Nothing is prescribed on this score.

On the whole, the emerging situation should be described as follows:

In almost all countries without exception, memory institutions produce, collect and store electronic information, and provide access to it solely on their own initiative and at their own discretion.

Though the volumes of electronic information are growing exponentially, an overwhelming majority of countries, including Russia, presently have no relevant infrastructure for its long-term storage and preservation, and almost all its above-listed elements are absent.

More than that, politicians and managers in charge of the development of culture, research, education and communications realize only on extremely rare occasions that such infrastructure has to be established, and the sooner the better.

Memory institutions have not elaborated a policy of long-term digital preservation, and there are no programmes and strategies. No tasks have been posed to such institutions. Conditions relevant for effective collection and storage of socially important electronic information have not been described and, naturally, not provided. The amount of necessary minimum and maximum funding of the institutions to meet those challenges has not been determined. In fact, no methods of such calculations are available. Memory institutions are not accountable to the public, and no forms of such accountability have been elaborated. Last but not least, neither the community nor government bodies demand such accounts.

The specialization of electronic stock to be preserved by various memory institutions is not determined, and the methods and criteria of selecting electronic information and its sources have not been set. Responsibilities for the storage and preservation of specific types and kinds of electronic

information on particular topics have not been distributed, and no responsible persons have been officially appointed.

Information selection for long-term preservation is all the more complicated because librarians and archivists presently have to assess the quality of selected content themselves while that was previously done by specially established agencies – publishers. It was up to them to decide what individually or collectively created content should be prepared for print – that is, edited from the point of subject matter, the manner of its presentation, language and style. Publishers determined what should appear in print and how, and in what number of copies. It depended on them whether information was to be circulated in the interest of the entire community or particular social, age and professional groups. To do that, publishing houses had specially trained professionals – editors, proof readers, peer reviewers, and editorial councils, boards, etc. At present, on the contrary, it is entirely up to the author to decide what should appear in the Internet, and how.

Standard methods of digital information storage and preservation have not been elaborated, and a majority of nations are not even working at them, let alone applying them in practice.

There is no special training for experts – theoretical and practical – on electronic information preservation. The goals of their training are not yet formulated, and there are no curricula for their graduate and postgraduate training. Only few nations are elaborating such curricula.

There is no proper communication between information authors and keepers. All manufacturers of traditional books and periodicals – that is, publishers and printers – knew in the paper information age that they would pass a certain number of copies to national and regional depositories in compliance with the law on legal deposit. At present, digital information producers – the same publishers, only now of electronic books and press, alongside film, television and radio companies – are not ready psychologically, organizationally and juridically to pass their information to libraries and archives for infinite storage. They avoid it most often with references to the legislation on intellectual property. Traditional information keepers, in their turn, are not ready psychologically, organizationally, technically and juridically to accept and keep such information. They are also not ready to monitor and archive information resources that exist only in the Internet as mobile content. At any rate, there are too few of those who can do it regularly and professionally.

Only in several best developed countries do professional library and archive leagues work actively, consistently and purposefully for long-term storage and preservation of digital information and future generations' access to it. They seek an answer to the same old questions of the analogue information age: what to store and why, who should do it, when, where and on what money, and to what an extent.

However, a majority of nations do not have a regulatory basis that would bind relevant institutions and agencies to take up those questions professionally. Neither do they have a special policy whose implementation would gradually create such a basis.

Russia's professional communities of electronic information producers and keepers also have not yet developed collective awareness of the necessity to start proper, purposeful work in that field. Only a few enthusiasts ask these questions in public – much to their colleagues' consternation.

We can thus say that the goals of long-term preservation of digital information are not formulated either in normative acts or in professional associations' regulations. The problems have not yet been described as a whole, duties not properly arranged, responsibilities not distributed, indices not set, and cooperation between all interested parties unorganized. Whatever is being done to digitize analogue depositories is chaotic and utterly uncoordinated, and no one has ever taken stock of gaps, overridings and parallelisms.

Hopefully, however, the Federal Agency for Press and Mass Media and the Ministry of Culture of the Russian Federation are developing an understanding of these problems. Two analytical reports have been drawn and several seminars held with their support, and the present international conference has also received their help.

Warning for the Future

What will happen unless we take urgent political and practical measures?

Very probably, we will soon see that we have:

- not cared to collect, and so irretrievably lost a vast amount of Internet and other information for this and the future generations – some of this information pertaining to Russian history, culture and research (Internet portals, websites, blogs, social network services, email, private and institutional electronic archives, etc); and

- lost a tremendous amount of digitized information because we failed either to give proper backing to its digitization or to catalogue it correctly and in time, and so it cannot now be retrieved, or again, because it was stored incorrectly.

We will certainly not lose everything. Some electronic information will survive. However, we and the future generations will have not a comprehensive picture of the past but a patchwork of information resources, many of which are not valuable for the understanding of the past and the succession of cultural heritage. That is why many world celebrities refer to the near future as the Dark Information Age, and describe it using the terms of "digital Alzheimer's", "digital amnesia", etc.

Also questionable will be the access to stored electronic information due to the problems with hard- and software for its reproduction. With a fatal lag in electronic information storage, memory institutions responsible for public domain information may lose that function. Transnational corporate mammoths who have found their bearings in time may develop into the principal holders of research, educational and cultural electronic information. In that case, public access to it may be lost. Society will have either to pay for access to such information or expropriate it and remove it from commercial circulation.

The public at large is still unaware of these problems. To promote universal knowledge in this field is one of the essential duties of professional information keepers. They are the people to offer solutions for related problems. However, the professional milieu of content keepers – librarians, archivists and museum people (as museums are also developing into mighty information centres) – also have huge problems: the philosophy of long-term digital information preservation has not been elaborated; explicit and realistic goals of such preservation have not been posed, etc. That is why there are drawbacks in the organization of these activities at the institutional and political levels alike. That is why we are not ready for full-scale work in this sphere – in other words, we cannot yet launch sustainable and unbroken modernization of relevant institutions, which do not have sufficient competent personnel, as such specialists are not trained anywhere. The few experts there are lack motivations, rights and resources. Last but not least, there is no public monitoring of their activities.

What Is to be Done

It is impossible to preserve digital information with the necessary content and quality, and in the adequate amount, if we regard the issue in isolation. We should analyze it in a cluster of such problems bred by the development of ICT as information policy, accessibility of information, information ethics, media and information literacy, and multilingualism in cyberspace. All these problems are topical, to varying extents, for developed and developing countries, for East and West, North and South.

Electronic information cannot be properly stored and preserved unless an efficient policy is launched and implemented through concerted efforts at the national and global levels. These ambitious efforts – intellectual, political, practical, organizational and educational – in research, education, culture, technology, telecommunications and other spheres require smooth cooperation of experts in all of them for a comprehensive and well balanced policy of information society development. The formation and implementation of such a policy at the national and transnational levels is an especially acute problem today. Information preservation should be an inalienable part of this policy, which cannot take shape unless it has a fundamental scientific basis and practises transdisciplinary methods and intersectoral/interdepartmental cooperation. Narrow departmental views and stances do not allow detect and formulate system-wide problems, let alone find their effective solution. That is why many nations' policy of building knowledge society is blatantly one-sided, with the principal accent laid on the development of technologies and the telecommunication infrastructure. Only negligible attention is paid to the content, formation, preservation and accessibility of information and the state of information institutions. What is even worse, no sufficient attention is paid to people – that is, the social aspect of development.

Even during the paper information era, no nation could cope with its preservation single-handed. That was why international cooperation was essentially necessary. It led to shared international standards and labour division. International cooperation is even more necessary now in the digitized world, when information easily crosses state borders and overcomes national jurisdictions.

Active transnational interdisciplinary partnership is one of the principal goals in digital information storage and long-term preservation. Such partnership is indispensable if memory institutions are to meet their own challenges and elaborate effective strategies. Success requires cooperation of not only librarians, archivists, museum personnel and other heritage keepers but also of politicians, public servants, researchers, writers, publishers, university professors, schoolteachers, and service and content providers. The cause must represent such intergovernmental organizations as UNESCO, ITU, and the Council of Europe; international professional associations, such as the International Federation of Library Associations and Institutions, the International Council on Archives, the International Council of Museums, the International Association of Recorded Sound Collections and others; specialized European Commission agencies, NGOs, the ICT industry, etc.

Libraries, museums, archives and other memory institutions were established to collect and eternally store information. Their performance is all the more successful the more their personnel, the staff of mass media outlets, publishers, film, television and radio companies, and public servants know about the policies of digital information storage and preservation in the world's leading countries, and foreign and international programmes and projects.

Top priority goals are:

- the study and assessment of the topicality, scope, depth and acuteness of the philosophical, political, economic, methodological, organizational, legal, personnel and technological problems of long-term digital information storage and broad public access to it;
- the promotion of awareness of the importance, scope and acuteness of the digital preservation problems by memory institutions and the mass media. For this, top-notch professionals should inspire relevant public demand and organize the training of competent journalists capable of problem coverage;
- the formation of effective international, national and institutional policies and infrastructures for the elaboration and implementation of science-based approaches to digital information preservation problems and public access to digital information;

- the formation of science-based approaches to the work organization of the leading libraries, archives and other cultural, research and educational institutions in digital information storage and preservation;
- the organization of permanent cooperation of all parties interested and involved (including potentially) in digital information storage and preservation; and
- relevant graduate and postgraduate training of experts and managers, and instilling in the public at large the skills of information preservation.

Peter Johan LOR
Sedgefield, South Africa
Extraordinary Professor
Department of Information Science
University of Pretoria

Ethical and Political-Economic Issues in the Long-Term Preservation of Digital Heritage

Introduction

"Heritage is our legacy from the past, what we live with today, and what we pass on to future generations" (UNESCO 2008:5). This statement is from the *World heritage information kit* published by UNESCO's World Heritage Centre. Heritage encompasses, broadly, the natural environment, the built environment (buildings and monuments, townscapes, archaeological sites), and artifacts (books and documents, objects, pictures) (Feather 2006:4). The latter category is concerned with movable objects and is sometimes referred to as movable cultural heritage. It includes objects of all kinds, works of art, books and documents. Books and documents fall within the definition of 'documentary heritage', a term applied to 'consciously created information carrying artifacts' (Feather 2006:6). That subset of the documentary heritage which is in digital form can be referred to as 'digital heritage'. It consists of born-digital content (content that was created and disseminated digitally) or digitized content (content that was created and disseminated in analogue format and subsequently digitized). In this paper I use the term 'digital preservation' to refer to the preservation of digital content as well as to the digitization of analogue content for purposes of preservation. I also use the term 'heritage' very broadly to include scientific and scholarly publications and data.

Awareness of the need for preservation of digital heritage has been growing for some time. In 2002 the International Federation of Library Associations and Institutions (IFLA) and the International Publishers Association (IPA) issued a joint statement on the archiving and preservation of digital information, entitled *Preserving the memory of the world in perpetuity* (2002) in which, among others, the following points were stated:

1. An increasing amount of information published only in electronic form has enduring cultural and documentary significance and is just as important as information published in more traditional forms.
2. The long-term availability of this information is required and action must be taken now to make this possible.

In the following year UNESCO adopted a *Charter on the preservation of digital heritage*. Article 1 set out the scope and importance of this heritage:

> The digital heritage consists of unique resources of human knowledge and expression. It embraces cultural, educational, scientific and administrative resources, as well as technical, legal, medical and other kinds of information created digitally, or converted into digital form from existing analogue resources. Where resources are "born digital", there is no other format but the digital object.
>
> Digital materials include texts, databases, still and moving images, audio, graphics, software and web pages, among a wide and growing range of formats. They are frequently ephemeral, and require purposeful production, maintenance and management to be retained.
>
> Many of these resources have lasting value and significance, and therefore constitute a heritage that should be protected and preserved for current and future generations. This ever-growing heritage may exist in any language, in any part of the world, and in any area of human knowledge or expression. (UNESCO 2003)

Article 9, "Preserving cultural heritage" emphasizes the universal value of the digital heritage:

> The digital heritage is inherently unlimited by time, geography, culture or format. It is culture-specific, but potentially accessible to every person in the world. Minorities may speak to majorities, the individual to a global audience.
>
> The digital heritage of all regions, countries and communities should be preserved and made accessible, so as to assure over time representation of all peoples, nations, cultures and languages.

Preservation of heritage seems to be inherently "good". In most publications on the preservation of digital content this is either stated explicitly or implied. Explicit recognition of such preservation as being

in the public interest is demonstrated by the language used in a policy document published by the National Library of New Zealand, *Creating a digital New Zealand* (2007). The priorities, outcomes and actions set out in this document have a strong moral flavor, for example the aim "to discover and cherish our languages, cultures, histories and national identity" (p.6), to protect traditional knowledge (p.21), to maintain cultural memory (p.25) and to strengthen the public domain (p.31). The preservation of the heritage of communities and nations, their traditions and cultural memory is, at face value, a good thing, as is digitization, since it helps to combine the twin goals of preservation and access without the difficult trade-offs we have to face when dealing with analogue materials, such as rare and fragile manuscripts, books films, audio recordings. For example, Lor & Britz (2004b:543-544) argued that web archiving serves a societal goal and that the web archive so created can be regarded as a common good.

Everything is not always as simple as it appears to be at first sight, especially when we consider the application of rapid technological innovation. I illustrate this by means of two hypothetical cases. The first deals with the harvesting and preservation in a wealthy country of political websites originating in a less affluent country. The second deals with a project initiated by a wealthy country to digitize the cultural heritage of a less affluent country.

Hypothetical Case #1: Preservation of Political Websites

In the country of Povertopia the ruling party has been in power for over twenty years. Every five years elections are held. They are rigged to ensure that the ruling power stays in power. Nevertheless some opposition parties are tolerated to show that the regime is quite democratic and when an election is held these parties are given some freedom to put forward candidates and promote their platforms. During the electioneering period, half a dozen new websites appear in which the small parties try to put forward their ideas and appeal to voters. Predictably, however, the ruling party wins again with an overwhelming majority. Some of the opposition politicians are arrested. Their offices are raided and their websites are closed down. That's the end of them, for now.

But such opposition websites are of interest to scholars in the wealthy nation of Opulentia who like to study the political systems, processes and communications in other countries. Knowing that these websites will not

be long-lived, scholars at one of the universities in Opulentia arrange to "harvest" these websites, downloading their pages at regular intervals during the election campaign, so that they will be able to study them later. The websites may provide useful material for PhD students in political science, media studies, gender studies, etc. This is done without obtaining the permission of the owners of the websites. It is felt that it is not feasible to obtain prior permission. The website owners are too busy campaigning to reply to requests for permission. By the time their answers arrive, if at all, it will be too late.

The archived material is made available on a workstation in the library of the university for access by bona fide scholars and students. One of the researchers later turns out to be employed by the embassy of Povertopia. He collects images of opposition figures and supporters. These images are later used to round up and imprison some of these people. Is this a far-fetched example? I leave it for you to decide. But it does raise some ethical questions, such as:

- Is it right to download websites from a poor country such as Povertopia without first obtaining the permission of the website owners?
- Would it be better to ask for permission first, even at the risk of losing the material?
- Does it make a difference if this is done for the sake of science and scholarship?
- Who should decide who may have access to the material?
- Can we argue that by downloading the websites we are actually helping the people of Povertopia to preserve a part of their national heritage that they are unable to preserve themselves?

That such questions are of real interest was demonstrated by the Political Communications Web Archive Project, undertaken by the Center for Research Libraries (CRL) in Chicago with funding from the Andrew W. Mellon Foundation, which studied both technical and curatorial aspects of the preservation of copies of the ephemeral websites of political manifestos and statements by political groups in developing countries and Western Europe (Limb 2004; Lor & Britz 2004b). Ethical issues such as those implied by the above questions were also addressed in the project. Unfortunately the final report on this project (Center for Research Libraries 2004) is no

longer available via the CRL web site. Much more recently a rather similar case is presented by Baker (2011), who discusses ethical implications of the donation of the Twitter Company's entire archive of Tweets to the Library of Congress.

Hypothetical Case #2: A Cooperative Digitization Programme

Library O in the wealthy country of Opulentia proposes to enter into a partnership with Library P in Povertopia, a poor developing country, to digitize the archives of a liberation movement which fought the former colonial power to gain independence for that country. The digital files will be made available on the Internet. Library O puts forward the following benefits:

a. It will provide Library P with the latest digital scanning equipment and train Library P's staff to use it. Library P will be able to retain the equipment for its own use after the project ends.

b. Library P will have staff trained in digitization procedures.

c. The archives of the liberation movement will remain in Library P.

d. The archives of the liberation movement will be digitized, which will help in preserving it because, once digitized, the original documents will no longer need to be handled so often.

e. Library P's building is not well maintained and its collections are not well protected against natural or human threats. If material should be stolen or a catastrophe should hit Library P and destroy the archives, the content will not be totally lost to posterity because there will still be the digital archives.

f. Scholars and students served by Library O will be able to access the digitized archives free of charge for scholarly research on the liberation struggle in the Povertopia. This will promote a better understanding of, and respect for, that country.

g. Scholars and students everywhere will be able to access the digitized archives for a modest fee, which will generate funds for further digitization projects.

h. Scholars and students served by Library P will be able to access the digitized archives free of charge.

This seems to be quite a fair deal. In fact, I sent this hypothetical case (in a slightly earlier version) to a convenience sample of 33 librarians in developing countries and asked them some questions. The main question was: "As the Director of Library P, would you accept this proposal?" Ten responded. This is hardly a scientific study, but it does suggest some insights. Seven of the ten respondents answered "Yes". Here are some of their comments:

- "Since I am in a poor country I have to say 'Yes'."
- "For sure."
- "I think it's a fair deal where everyone benefits."
- "Materials in developing countries are already getting lost at an alarming pace."

The "yes" group includes one very highly developed country whose perspective was providing rather than receiving assistance. The three respondents who answered "no" showed a greater level of legal and political awareness. They raised the following issues:

Copyright: the partner in Opulentia would hold copyright on the digital files, have full ownership, do whatever they wished with them, charge access fees, with no royalties going to the local partner.

- Lack of full, prompt or clear disclosure of contract conditions.
- More partnership was needed; local institution should not be seen only as a "beneficiary" but as a full partner.
- Terms of the proposal were limited to short term benefit to the library in Povertopia (equipment of limited lifespan, training rapidly obsolete if not refreshed) in exchange for a perpetual right for Opulentia party to exploit content commercially.
- Inability of parties in Povertopia to make full use of the content they have digitized as part of a project. [For example, they may lack the appropriate scholarship, scholarly resources or academic programmes.]

Moral Choices

These two hypothetical cases illustrate the contention of Hamelink (2000) that the development and application of technology, and the use of its applications, imply moral choices:

> Whatever breathtaking advances technological innovations offer, they are never without trouble. Technology inevitably brings great benefits and awesome risks. This essential ambivalence raises the challenging question about human governance of technological development. Can a balance be struck between progress and plague? What choices should be made to shape technology towards humanitarian aspirations?" (Hamelink (2000:1)

The cases also suggest that issues of resources and control – political-economic issues of power relations among nations, particularly rich and poor nations – need to be considered. Following Woods (2001:1), who uses the term 'political economy' to refer "the changing relationship between political systems ... and economic forces", I understand the political economy of information to refer to the impact of interacting political and economic forces on the creation, dissemination and availability of information, especially, their impact on the flows of information between developed and developing countries. These can be seen as the international dimension of the ethical issues that have to be considered in respect of individual rights.

Aggravating the problem is the sense of urgency encapsulated in such expressions as 'digital dark ages' (used as early as 1997 by Kuny (1997) and most recently in a report to the European Union (Niggemann, De Decker & Lévy 2011) which offers a choice between a new Renaissance and a digital Dark Age (p.7). In scientific circles there is a similar concern about the potential loss of research data (Burton 2007). The expression 'digital dark age' dramatizes the threat of large-scale loss of heritage and research data and suggests that 'something has to be done, quickly'.

In considering moral problems, ethics offers various approaches. For example, virtue ethics is primarily concerned with the individual making the choices rather than with their consequences for those affected. Strict utilitarian ethics is concerned with consequences and emphasizes the greatest good for the greatest number of people. The deontological approach emphasizes rights and duties, but does not always take into account the consequences of moral actions (Lor & Britz 2004a:17-18). In this paper I adopt a broadly rule-utilitarian approach to moral decision-making, where I understand rule utilitarianism as an approach which adopts

> ...certain rules that can guide our actions aiming to ensure that it will lead to a common good for society. It ... asserts universal principles ... but acknowledges the fact that the application is codetermined by the situation. (Lor & Britz 2004a:18)

In applying a rule-utilitarian approach to decisions on a common good (as would be created by a program to digitize heritage materials, preserve born-digital materials, or archive websites) we need a shared moral foundation, one based on a universal moral consensus – or one which is as nearly as possible universal. This implies that it should be inclusive of East and West, and of developed as well as developing nations, of poor rural communities as well as academic elites. Lor & Britz (2004b:544-545) proposed that such a shared moral foundation be based on the twin principles of justice and human rights. The three core principles of justice distinguished by John Rawls (1971, as paraphrased in Lor & Britz 2004b:545) are relevant here:

1. All people are of equal value and should be treated as such.
2. A person ought to get what is due to him/her. Although this may differ from person to person and from context to context, it is considered when people get or undergo something they do not deserve.
3. Although all people are of equal value, justice also recognizes the inequalities between people in certain cases, provided that differences in the treatment of people should not violate the first principle of equal value.

For purposes of application such general principles of justice are commonly articulated as human rights, as for example in declarative and legislative statements such as the Universal Declaration of Human Rights (UDHR) and national bills of rights. A number of human rights were identified by Lor & Britz (2004b:255) as relevant to web archiving. I expand on them here for application in the broader context of information ethics generally and more specifically in relation to digital preservation:

A. The right to personal autonomy (cf. UDHR Article 1: *"All human beings are born free and equal in dignity and rights"*)
B. The right to life, liberty and security of person (UDHR, Article 3)
C. The right to privacy (cf. UDHR Article 12: *"No one shall be subject to arbitrary interference with his privacy..."*
D. The right to reputation (cf. UDHR Article 12: *"No one shall be subject to ... attacks upon his honor and reputation"*.)
E. The right to freedom of expression (cf. UDHR, Article 19: *"Everyone has the right to freedom of opinion and expression; this*

right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.")

F. The right of access to information (cf. UDHR, Article 19: *"... freedom ... to seek, receive ... information and ideas..."*)

G. The right to freedom of communication (cf. UDHR, Article 19: "*... freedom ... to ... impart information and ideas through any media and regardless of frontiers."*)

H. The right to freedom of association (cf. UDHR, Article 20: "*(1) Everyone has the right to freedom of peaceful assembly and association; (2) No one may be compelled to belong to an association."*)

I. The right to participation in the cultural life of the community (cf. UDHR, Article 27(1): *"Everyone has the right freely to participate in the cultural life of the community..."*)

J. The right to share in the benefits of arts and sciences (cf. UDHR, Article 27(1): *"Everyone has the right... to enjoy the arts and to share in scientific advancement and its benefits."*)

K. The right to control the dissemination and use of created or authored work, including the right to maintain the integrity of the work (based on UDHR, Article 27 *"(1) Everyone has the right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he is the author"*; read with Article 17: *"Everyone has the right to own property alone as well as in association with others; (2) No one shall be arbitrarily deprived of his property"*); the emphasis here is on 'moral interests'.

L. The right to recognition as author or creator of a work (based on UDHR, Article 27, as above, read together with UDHR, Article 12)

M. M. The right to own intellectual property (based on UDHR, Article 27, as above, with the emphasis on 'material interests).

Comprehensive as this may appear, something is missing here. The Universal Declaration of Human Rights, for reasons discussed by Mazover (2004), is concerned with individual human rights and has little if anything to say about communal or group rights. This is unfortunate when we are dealing

with the cultural and documentary heritage of groups who do not have a tradition of allocating authorship, priority of discovery, or ownership of artistic and intellectual contributions, to individuals. I have therefore extended the application of certain rights somewhat to communities and groups. This is particularly relevant to the rights of communities and groups to control the dissemination and utilization of their traditional or communal knowledge and cultural expressions (K), to be recognized as the authors of traditional or communal knowledge and cultural expressions (L), and to be recognized as the owners of traditional or communal knowledge and cultural expressions and to benefit from its use (M).

In the discussion that follows, I shall identify the main parties that have an interest, and may be entitled to exercise rights, in relation to digital preservation. I then attempt to analyze the responsibilities that those who preserve digital content have towards these parties, and the political-economic considerations that arise. The parties are the following:

Moral Responsibilities And Political-Economic Implications

The parties considered in this section are as follows:

Authors, creators

1. Originating communities
2. Rights holders
3. Holding institutions
4. Persons depicted or described
5. Digitizing or acquiring institutions
6. Users

Table 1 is an attempt to summarize the information rights that can be attributed to each of these parties. It is work in progress; the notation in each cell is open to debate. However, the scope of this paper precludes a full discussion of each. Hence Table 1 is followed by a discussion of some selected issues affecting the responsibilities of parties involved in digital preservation.

TABLE 1: Information rights of parties involved or affected by digital preservation

Right	Authors, creators	Originating communities	Rights holders	Holding institutions	Persons depicted	Digitizing / acquiring institutions	Users
A Personal autonomy	**X**	**X**	x		X		X
B Life liberty and security	**X**	x	x		X		X
C Privacy	X	x	x	x	**X**	x	**X**
D Reputation	X	X	X	X	X	X	X
E Freedom of expression	**X**	**X**	x	x	**X**	x	**X**
F Access to information	x	x	x	x	x	x	**X**
G Freedom of communication	**X**	x	**X**	x	**X**	**X**	x
H Freedom of association	x	x	x	x	x	x	x
I Participation in cultural life of the community	x	x		x	x	x	x
J Share in benefits of arts and sciences	x	**X**	x	x	x	x	**X**
K Control dissemination, use & integrity	X	X	X	X		X	
L Recognition as author or creators	X	X					
M Own intellectual property	X	X	X	x		X	

Legend:

X: Yes: significant, with potential for conflict

x: Yes: basic right, no obvious or limited potential for conflict

Blank: Not applicable

Authors, creators

In principle the authors and creators of content that is preserved digitally have essentially the same rights as the creators of analogue content, for example the rights to freedom of expression (E) and communication (G), to be recognized as authors or creators of the digitized works (L), to own the intellectual property (M) and to control the dissemination, use and integrity of their work (K). Two factors render the exercise of these rights more difficult and complicate the decision making of those concerned with preservation.

The first is the nature of the digital medium. Digital content is "susceptible to manipulation, interrogation, transmission and cross-linking in ways that are beyond the capacity of analogue media" (Deegan & Tanner 2002:7). Unlike the content of printed media, digital content is easy to transcribe, replicate and alter without trace (Deegan & Tanner 2002:60). Information professionals have a responsibility to watch over the integrity of digital collections and to combat accidental or deliberate alteration as well as plagiarism. The labile nature of digital content has been referred to as instability. However, Feather (2006:12) ascribes this to the "fundamentally dynamic nature of digital documents". This raises the possibility that, in a sense, freezing at a given point in time content that was never intended to be anything but ephemeral, distorts it. Rauber et al. (2008) point out that, just as artists may create one-time installations that are not intended to be preserved, so "...many documents may be may be placed on the Web because it is ephemeral, because they are meant as a temporary statement; a comment that is never intended to be captured and be maintained for eternity."

This brings me to the second factor: ethical issues have become more complex because the nature of authorship has changed. In the digital world there are two major types of creators, with many in between. The first type comprises authors who consciously create for dissemination, and we can reasonably expect that their will realize that their work is likely to be preserved somewhere for future use, as is the case with print and other analogue media. This clearly constitutes publication. But in the world of Web 2.0 there are also many incidental creators, creators of the second type, who post content on blogs and social networking sites in a spontaneous manner, without considering the implications, and without realizing that their posts, comments and photos may be collected by a web crawler, preserved indefinitely, and made accessible to users anywhere. Some of this material may later be a source of embarrassment to them (or

to other people depicted). A case in point is the donation of the Twitter company's entire database of 'tweets' since 2006 to the Library of Congress. The magnitude of this database, which grows at a rate of 50 million tweets a day, is mindboggling, and the potential for data-mining for research, commercial or security purposes is frightening (Baker 2011).

Much of the Web 2.0 content is analogous to postcards. En route to the intended recipient, postcards can be read by anyone who handles them. In that sense they are not *private*. But sending a postcard does not constitute *publication*. Thus the preservation of Facebook pages, Flicker pictures and Tweets risks violating the creator's right to personal autonomy (A) in the sense that creators are not able to make their own decisions about what to preserve and disseminate, to privacy (C), reputation (D) and, in extreme cases, life, liberty and security (B) (cf. Lor & Britz 2004b). In an excellent exploration of ethical issues in web archiving, Rauber *et al.* (2008) have pointed out that the risks imposed by web archiving are greatly magnified by the availability of data-mining techniques and clever software that can perform very effective searches for specific persons, inter alia for use by human resources managers ('web profiling'). Hamelink (2000:126) warned that digitization facilitates government surveillance. In a UNESCO-sponsored survey of the ethical implications of emerging technologies, Rundle and Couley (2007:72) pointed out that modern IT systems may help in extending and democratizing access, but also pose a threat in terms of "...the potential increase in surveillance capabilities" made possible by the processing of "immense quantities of data gathered by search engines". This, they warn, could hamper the promotion of human rights. Conversely, the recent Wikileaks exposures have shown that the unauthorized dissemination of confidential information held by governments is also easy and attractive.

If the informal posting of content on social networking sites does not constitute publication, the use of such material for research purposes should perhaps be considered in the same light as individually identifiable data about human subjects in social science research, where this is dealt with by government regulation and the codes of ethics of various professions. The Association of Internet Researchers (AoIR) has published a policy document, *Ethical decision-making and Internet research* (2002), which deals with ethical issues in the use of data obtained from the Internet. In the field of library science, some attention has also been paid to the ethics of Internet research (e.g. Holley 2006). Information professions have a responsibility to devise measures that will minimize the risks posed to authors and creators by web archiving.

Originating communities

While the aim of a digitization project may be to promote nation-building and appreciation of cultural diversity through an understanding of the history and culture of a particular group, there are a number of ethical pitfalls in such projects (Lor & Britz 2004a). As suggested by Hypothetical Case #2 it is important that the autonomy and self-respect of the group in question be respected (A; D). Where partnerships are entered into between institutions in developed and developing countries, care must be taken to avoid-donor driven projects which are ultimately exploitative and which ignore the priorities and wishes of the recipient institution. It is important to develop an equal partnership, where the recipient institution participates in decision making on what is to be digitized, shares equally in the management of the project, and receives sustainable, long-term benefits (J) as a quid pro quo for making its heritage available. This implies that the opportunity costs of the project be considered. These are the strategic costs incurred through potentially misdirecting resources to a less appropriate strategy (Deegan & Tanner 2002). An example would be directing the limited human resources of a national library in a developing country to a digitization project which will mainly benefit academics in the developed country when staff resources and management overheads might have been better spent on a program to develop community libraries, literacy or indigenous languages. This is true even if all costs of technology, training etc. are borne by a donor country. Opportunity costs are not outlays of funds but lost benefits.

Digitization programs should not lead to the ultimate retention of the original material in the donor country (K). Referring to digitization of African materials, Limb (2004:150) recommends that there should be regular consultation between Western and African scholars and archivists "to ensure that the 'fattening' of Western repositories does not lead to a corresponding decline in visits to archives in Africa by Western scholars."

Originating communities have a right to share in the benefits of their knowledge (J) when this is made available to other parties. Exploitation can result from asymmetrical relationships. Limb 2004) cites patents taken out by western companies on drugs derived from indigenous medicine. Conversely, digitization of ancient texts documenting indigenous knowledge is being used by the government of India as a strategy to combat 'biopiracy' which occurs through the patenting by pharmaceutical

companies of traditional medical knowledge (Brahmachari 2011). There is no moral objection to the South-North flow of information as such. On the contrary, Britz and Lor (2004) have argued, on the basis of the principle of distributive justice, that developing countries also have a moral responsibility to disseminate their heritage to other countries.

In partnerships with developing countries, cultural values and traditions affecting access should be respected (K). Jordan (2006:32) stresses that "[r]espect for cultural sensitivity and individual privacy should be considered when selecting material for inclusion in digital collections". The Assembly of Alaskan Native Educators (2000) has drawn up guidelines for a wide range of affected and interested parties, including tribal elders, authors and illustrators, curriculum developers and administrators, educators, researchers, native language specialists and native community organizations. In Australia the *Aboriginal and Torres Strait Islander Library and Information Resources Network Protocols* (Australian Library and Information Association 2005) constitute a guide to good practice in handling materials with Aboriginal and Torres Strait Islander content, covering *inter alia* the appropriate handling of documentation relating to the intellectual property of Indigenous peoples, appropriate subject headings and indexing terminology, and the treatment of secret, sacred and offensive material. A brief section deals with digitization and the Internet, from which I quote the following passage:

> Digitization provides opportunities to improve Indigenous Australians' access to historical and contemporary cultural and Indigenous knowledge materials which are currently dispersed in institutional collections across the nation.
>
> Easier access provided by digital technologies also increases the risk of breaching Indigenous cultural protocols for the management of Indigenous knowledge and cultural materials.
>
> Intellectual property and technology issues also introduce levels of complexity for the sustainable management of Indigenous materials in the digital domain that pose challenges for both the library and information services profession and for Indigenous communities.

Information professionals involved in digitizing the heritage of indigenous peoples have a responsibility to consult with the peoples concerned, develop a sympathetic understanding of the issues, and adopt and apply relevant policies. One area of relevance is the assignment of descriptive, structural and administrative metadata (cf. Deegan & Tanner 2002).

Cataloguers assigning descriptive metadata must be aware of the possibility of unconscious cultural bias. For example, in assigning descriptive metadata, the target audience will be taken into account. But who is the target audience? Local people or academics in a developing country? Similar questions arise concerning structural metadata, which describes the structure and relationships of digital objects, and administrative metadata, which refers to the data needed to manage all aspects of the life-cycle, including curation, preservation and access. Of particular concern here are rights management and conditions of access, which should reflect respect for traditional restrictions on who may view traditional artifacts or ceremonies.

Rights holders

Rights holders, who are in many cases not the authors or creators of the material, are generally recognized in law as having a right to own intellectual property (M), to control its dissemination and use (K) and to derive revenue therefrom (M). Copyright law can constitute a formidable barrier to digitization projects since digitization involves the making of copies. Since obtaining copyright clearance is very labor-intensive, many digitization initiatives restrict themselves to material that is in the public domain. Given the excessive term of copyright in most developed countries and (in the case of 'orphan works') uncertainties concerning the identity and whereabouts of copyright owners, this leaves a large tranche of the world's documentary heritage beyond reach of those who can least afford to pay.

The ethical codes of information professionals generally promote compliance with intellectual property laws and conventions. Users of digitized content should refrain from abuse of copyrighted materials. Digitizing and holding institutions have a responsibility to educate users and put in place measures to prevent abuse, without inhibiting legally permitted fair use by excessive caution.

However, many, particularly in scholarly circles, consider the relationship between creators, rights holders and users to have become distorted by the commodification of information. It is thought that the migration of scholarly content from analogue to digital carriers and the growing predominance of networked digital resources have exacerbated what was already an unfair system. The rise of the open access movement must be seen against this background. In developing countries especially, open

access is seen as offering possibilities for closing the gap between wealthy and poor nations through more equitable and affordable access to the world's scholarly output (Jain & Nfila 2011). Up-to-date information is seen as critical to national development.

In this light continued support of the current intellectual property regime raises moral issues. Troll Covey (2009:1), discusses the conflict around open access in terms of an epistemic regime "...comprised of two social worlds: a world where knowledge cannot be owned and a world where knowledge can be owned". Authors have an ethic of sharing. Copyright law embodies an ethic of economic right. Troll Covey's point of view is that law and ethics are out of alignment here and she exhorts librarians to resist the current system of copyright through "civil disobedience" (p.13-14). It is beyond the scope of this paper to discuss open access, about which a huge literature has grown, but it is worth pointing out that it constitutes a major theme in the political economics of North-South information flows, and provides an interesting example of a disconnect between law and what is generally perceived as morally justifiable.

Holding institutions

I refer here to institutions holding the originals of material that is to be digitized. In the case of digitization projects agreements between the holding institution and the party undertaking the digitization should be fair to both parties. A wealthier or better informed party should not take advantage of the other. This principle (A) has implications for the all aspects of the project. One aspect is the selection of materials to be digitized: when materials for developing countries are digitized, are the interests of the holding institution and country taken into account, or is the wealthier party 'cherry-picking' material relevant to its own country, that will complement its own collections? How international are 'world' or 'global' digital library projects? Who selects the material? Is it primarily material that holds a special interest for the library in the developing country? Does it reflect a European or an American world-view? How equal are the "partnerships"?

Another aspect is the disposal of the originals. In relations between institutions in developed and developing countries the rights of the latter to maintain the integrity of their collections should be respected. This can be seen as the institutional counterpart of the individual's right to life,

liberty and security (B). It implies that originals held by them should not be harmed and that the integrity of their collections should be respected. The critical mass of the generally more modest holdings of libraries in developing countries should not be diminished. Agreements should ensure that they retain the original material after the digitization has been completed. Originals should not be shipped to institutions in wealthy countries on the pretext that the holding institution in the developing country lacks the resources to take proper care of the material. If this is the case, the more ethical approach would be to provide development assistance aimed at capacity building (Britz & Lor 2004). As a quid pro quo for making their material available for digitization, libraries and archives in developing country should be granted unlimited free access to the digitized corpus, and a copy of the database to host locally (F).

A third aspect to be taken into account is sustainability: Deegan and Tanner (2002:99) point out that "[s]ustainable development and funding are needed to manage digital content effectively into the future. Without these two factors, relevant technology cannot be renewed, skilled staff cannot be retained and the intellectual, cultural and educational rewards cannot be reaped." Scanning equipment rapidly becomes obsolescent. Workstations needed to access the digital holdings need to replaced every 3-5 years. What will happen 3-5 years after the end of the sponsored project?

Persons depicted or described

Here I refer to third persons whose portraits appear in digitized material, or whose activities and attributes are described or discussed in the digitized material, particularly if they are recognizable and if the depiction can give offence to them or their family (C; D). This is sometimes restricted to living persons. The issue is especially sensitive if the individuals concerned were not aware and had no say in the release of the information (A) (Anderson 2005:29). Jordan (2006:33) refers to the need for "[p]reserving respect for individual and family privacy" and urges caution when including items in digital collections "which may violate individuals' privacy or the privacy of named or depicted people's living descendants". The principles referred to earlier, in the discussion of incidental creators who post content on blogs and social networking sites in a spontaneous manner, are largely applicable here as well.

Information professionals should be aware that images or video footage of demonstrations and other protest activities may be used to prosecute

participants and may have very severe consequences for persons who can be identified (B). They therefore have a responsibility to prevent such harm. Various measures, such as time-limited embargoes, restriction of access to accredited researchers and software-supported means of anonymizing data, can be considered for this purpose (cf. Rauber *et al.* 2008; Baker 2011).

Digitizing or acquiring institutions

When engaged in digital preservation activities, information professionals have a responsibility to their institutions to ensure ethical decision-making on all aspects of digitization and preservation, for example, applying professional, non-arbitrary and non-biased criteria to the selection of material to be digitized and preserved (Baker 2011), the adoption of standards, judicious and principled deployment of resources, and the selection of tenders for hardware, software, and services. This is of course true for any heritage project. Information professionals should take a long-term view and eschew flash-in-the-pan prestige projects which constitute a big temptation especially in national and international projects with a high profile, where politicians or donors want quick, spectacular, results.

Ethical procedures should be followed in appointment of staff. Bearing in mind that much of the work in a digitization project is of a routine nature, project managers have a responsibility to ensure fair conditions of service for students, interns, and persons in assisted or sheltered employment. If digitization work is carried out in the country of origin or outsourced to developing country an ethical stance should be adopted to ensure that "sweatshop" labor is not used.

One way to finance digital projects and ensure longer-term sustainability is by marketing the digitized material. This may imply charging fees, or selling portions of the material to commercial publishers while making it available free of charge to users affiliated to the institution. This raises ethical issues, particularly when the institutions are supported by public funds (Limb 2004).

Users

If no access is intended or possible, ever, preservation has no point. The ultimate purpose of preservation is access, now or in the future. This introduces a time dimension, as information professionals have responsibilities in respect of users immediately, in the short term and in the

distant future. In the context of heritage we refer to the latter as 'posterity'. A question to be considered is whether access is to be provided to users in a specific institution, locally, nationally or world-wide. We say that we preserve digital heritage for posterity and humanity, but ultimately usage is through the eyes, ears and fingertips of individuals.

Information professionals are responsible to users to provide them with authentic content (F). *Authenticity* must be distinguished from *authentication*. Authenticity is concerned with "reliability over time" (McKenzie, cited in Deegan and Tanner 2002:185), while authentication is mostly used in the context of identifying persons who have permission to access digital data. To ensure authenticity we should digitize only the best version of any given work. Anderson (2005) cites the example of an eminent US university which, to avoid copyright issues, digitized an inferior edition of a book because it was out of copyright. The problem is that users of the digitized collection will assume that this is a good edition because of the high reputation of the university which scanned it. This could lead to errors in user's work (D). We also have a responsibility to users (as we have to creators) to ensure the *integrity* of the content. As noted earlier, digital content is 'fragile'. It is therefore important to ensure that care is taken in the digitization process, for example, not to omit pages, figures, or appendices when digitizing printed books, and to provide metadata detailing the state and completeness of the original. Creators, too have a responsibility for the integrity for their work, for example by providing good quality metadata.

Note that this emphasis on authenticity and integrity is in conflict with the idea of "anything is better than nothing", which is motivates some digitization projects. As in the case of inappropriate book donations to developing countries, such an attitude is ultimately a form of disrespect – vis-à-vis creators, users, and the originating community.

The right of access to information (F) implies a responsibility on the part of information professionals to provide free or affordable access to the digital content. Limb (2004) states that librarians have a professional responsibility to defend freedom of expression (E) and access (F) by combating censorship and promoting open government. The right of access to information implies equal access to users regardless of demographic and economic characteristics and regardless of where they reside – after all, this is what networked digital resources are for. This right also implies that measures be taken to provide access to users with special needs, e.g. the

visually impaired (cf. Anderson 2005:22). The Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C) provides important guidelines for Web accessibility (Web Accessibility Initiative 2011).

The Web contains huge amounts of inaccurate and misleading information. When the web is archived, such content is archived too – and should be, if the digitized content is to be an accurate reflection of the society which creates it. Thus web archiving, warts and all, is in the interest of (future) historians, sociologists, ethnologists etc. This does pose a challenge: should 'naive' users be protected from misinformation, disinformation and hoaxes? How can such 'protection' (B) be reconciled with freedom of access (F)? Part of the answer is to be found in information literacy, especially enabling users critically to evaluate the information they find on the Internet. Librarians can also construct web pages and portals to lead users to reliable sites (Limb 2004).

Thus far in this section responsibilities to individual users have been emphasized. We also need to consider the broader categories of 'humanity' and 'posterity', which are so often cited when digital projects are promoted. If the aim is to create a true 'global digital library', who will control it? Will it be a truly distributed resource or will it reside in a server in a wealthy country? What guarantee will users in the less affluent countries have that contents will not be censored or that access will not be denied for reasons of foreign policy or 'national security' if relations between their country and the country holding the digital content should deteriorate?

A similar question arises in respect of born-digital content, for example, e-journals and e-government publications. Increasingly, these are only distributed electronically (Anderson 2005). It is possible for publishers to cut off access to the data or withdraw items for commercial or political reasons. To what extent can we trust commercial entities, no matter how strong they are at present, to maintain the integrity and availability of their e-publications indefinitely into the future? Even major international corporations can fail or, less dramatically, simply dispose of businesses that are no longer profitable. Reliance on the private sector is unwise in the long term. If the task of long-term preservation and access is devolved to trusted repositories such as the Koninklijke Bibliotheek, the national library of the Netherlands, which does a fine job of preserving the e-journals published by Elseviers and other major journal publishers, we can expect greater

stability, but to what extent should the developing world remain dependent upon institutions in the developed world?

Finally, a critical responsibility is protecting the confidentiality of users' activities in using digital material (C). Confidentiality of users is closely tied up with freedom of expression (E) and access to information (F), since lack of privacy inhibits these freedoms.

Conclusion

In this paper I have tried to provide an overview of ethical and associated political-economic aspects of digital preservation. The sheer number and complexity of the issues may appear intimidating. Certainly, they do not allow of easy 'solutions'. Referring to web archiving, Rauber *et al.* (2008) suggest that complex ethical issues should not stop such archiving, but that along with technical research and development, some resources should also be devoted to research on ethical aspects. It is worth recalling here the observation of Hamelink (2000:6) that technological applications tend to be developed without prior consideration of ethical implications and that humanity has a long history of denying moral responsibility for technological innovations:

> The prevailing trend is to think that all possible problems can be fixed by technological means that do not require ethical reflection. All social problems are technological problems. In this frame of mind their solutions do not need any ethical reflection." (p.6)

Much of what is written about the rapid development of digital technology Web 2.0 and the like appears to have been written with breathless haste and is filled with dire warnings that information professionals and their institutions must adapt or be swept away by progress. This reflects a widespread technological determinism, which is defined by Hamelink (2000:8) as the "belief that social developments are determined by technological progress". From time to time information professionals who concern themselves with the preservation of the heritage of humankind, including the digital heritage, should take a step back for reflection, so as not to allow themselves to be stampeded into technology-driven solutions to problems that have not yet been understood. Such reflection should include consideration of the ethical and political-economic dimensions of their preservation efforts.

References

1. Anderson, Cokie G. (2005) Ethical decision making for digital libraries. Oxford: Chandos Publishing.
2. Assembly of Alaskan Native Educators (2000) Guidelines for respecting cultural knowledge. Fairbanks: Alaska Native Knowledge Network, University of Alaska Fairbanks. Available http://www.ankn.uaf.edu/publications/Knowledge.pdf, accessed 2011-09-26.
3. Association of Internet Researchers (2002) Ethical decision-making and Internet research; recommendations from the AoIR Ethics Working Committee (2002), compiled by Charles Ess. Available http://aoir.org/reports/ethics.pdf, accessed 2011-09-26.
4. Australian Library and Information Association (2005) Aboriginal and Torres Strait Islander Library and Information Resources Network Protocols. Available http://www1.aiatsis.gov.au/atsilirn/protocols.atsilirn.asn.au/ATSILIRNprotocols.pdf, accessed 2011-09-26.
5. Baker, Antoinette E. (2011) Ethical considerations in Web 2.0 archives. SLIS student research journal 1(1), article 4.
6. Brahmachari, Samir K. (2011) India launches Traditional Knowledge Digital Library (TKDL) to protect against bio piracy. Washington Bangla Radio, available at http://www.washingtonbanglaradio.com/content/83110711-india-launches-traditional-knowledge-digital-library-tkdl-protect-against-bio-pirac, accessed 2011-09-02.
7. Britz, Johannes J. & Lor, Peter J. (2004) A moral reflection on the digitization of Africa's documentary heritage. IFLA journal 30(3):216-223.
8. Burton, Adrian (2007) Avoiding a digital dark age. ScienceAlert Australia. ScienceAlert Australia & New Zealand; available http://www.sciencealert.com.au/opinions/20070504-14829.html, accessed 2011-09-26.
9. Center for Research Libraries (2004) Political communications web archiving: a report to the Andrew W. Mellon Foundation. Unpublished report by the Center for Research Libraries, Cornell University, the Internet Archive, the Latin American Information center, University of Texas at Austin, New York University and Stanford University, March 2004.

10. Creating a digital New Zealand: New Zealand's digital content strategy. (2007) Wellington: National Library of New Zealand.
11. Deegan, Marilyn & Tanner, Simon (2002) Digital futures: strategies for the information age. New York: Neal Schuman Publishers; London: Library Association Publishing.
12. Feather, John (2006) Managing the documentary heritage: issues for the present and future. In: Gorman, G.E. & Shep, Sydney J. (eds.) Preservation management for libraries, archives and museums. London: Facet Publishing: 1-18.
13. Hamelink, Cees J. (2000) The ethics of cyberspace. London, New Delhi: Sage Publications.
14. Holley, R. P. (2006). The ethics of scholarly research and the Internet: Issues of publication, privacy, and the right to speak. Journal of Information Ethics, 15(1), 27-34.
15. Jain, Priti & Nfila, Reason B. (2011) Open access to research in academic institutions: African perspectives. Critica bibliotecológica 4(1):31-48.
16. Jordan, Mark (2006) Putting content online: a practical guide for libraries. Oxford: Chandos Publishing.
17. Kuny, Terry (1997) A digital Dark Ages? Challenges in the preservation of electronic information. Paper presented at the 63rd IFLA Council and General Conference, Copenhagen, 31 August -5 September 1997. Available http://archive.ifla.org/IV/ifla63/63kuny1.pdf, accessed 2011-09-26.
18. Limb, Peter (2004) Digital dilemmas and solutions. Oxford: Chandos Publishing.
19. Lor, Peter J. & Britz, Johannes J. (2004a) Information imperialism: moral problems in information flows from South to North. In: Information ethics in the electronic age: current issues in Africa and the world; ed. by Tom Mendoza and Johannes J. Britz. Jefferson (North Carolina): McFarland: 15-21.
20. Lor, Peter J. & Britz, Johannes J. (2004b) A moral perspective on South-North web archiving. Journal of information science 30(6):540-549.
21. Mazover, Mark (2004) The strange triumph of human rights, 1933-1950. The historical journal 47(2):379-398.

22. Niggemann, Elizabeth; De Decker, Jacques & Lévy, Jacques (2011) The new Renaissance: report of the 'Comité des Sages' Reflection Group on Bringing Europe's Cultural Heritage Online. Brussels: European Commission. Available http://ec.europa.eu/information_society/activities/digital_libraries/doc/reflection_group/final-report-cdS3.pdf, accessed 2011-09-26.
23. Preserving the memory of the world in perpetuity (2002): a joint statement on the archiving and preserving of digital information. IFLA/ICA Steering Group, 27 June 2002. Available http://www.ifla.org/publications/preserving-the-memory-of-the-world-in-perpetuity-a-joint-statement-on-the-archiving-and, accessed 2011-09-26.
24. Rauber, Andreas; Kaiser, Max & Wachter, Bernhard (2008) Ethical issues in web archives creation and usage – towards a research agenda. Paper presented at the 8th International Web Archiving Workshop, Aarhus, Denmark, 18 & 19 September 2008. Available http://www.ifs.tuwien.ac.at/~andi/publications/pdf/rau_iwaw08.pdf, accessed 2011-09-26.
25. Rundle, Mary & Couley, Chris (2007) Ethical implications of emerging technologies: a survey. Paris: UNESCO
26. Troll Covey, Denise (2009) The ethics of open access to research: a call for civil disobedience and moral courage. Progressive Librarian 33 (2009): 26-42. Available at: http://works.bepress.com/denise_troll_covey/53, accessed 2011-09-23.
27. UNESCO (2003) Charter on the preservation of digital heritage. Available http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_PRINTPAGE&URL_SECTION=201.html, accessed 2011-09-26.
28. UNESCO (2008) World heritage information kit. Paris: UNESCO World Heritage Centre. Available http://whc.unesco.org/documents/publi_infokit_en.pdf, accessed 2011-09-26.
29. Web Accessibility Initiative (2011) WAI guidelines and techniques. Available http://www.w3.org/WAI/guid-tech.html, accessed 2011-09-27.
30. Woods, Ngaire (2000) The political economy of globalization. Basingstoke: Macmillan.

Oleg NAUMOV
Moscow, Russian Federation
Deputy Head
Federal Archival Agency

Archives and Digital Content: Problems of Coordination

With massive transition to computer technology underway, archivists face the challenge of preserving and using loads of digital content and scanned documents.

Today's imperative is that archival documents and historical heritage at large become accessible online from any type of computer or mobile device (including tablet computers and smartphones).

The Russian archivist community's top priority is to build a computer-based inventory for the National Archives (as of today, the country's federal, regional and municipal archives carry about 250 million entries overall), making it publicly accessible online.

We are planning to launch a pilot system of the kind at the Russian Archives' Web links page later this year and to gradually fill it up with information from federal and regional archives in the months ahead.

All the items for this new inventory will have to be digitized as they come along. The digitization of cultural heritage, including archival collections, and digital content creation are listed among the top priorities of the government programme *Information Society, intended for the period through 2020*.

Digitizing the National Archives will be no easy task, given the specific character and diversity of the documents stored. An R&D effort will have to be made to identify standard requirements for scanning equipment used in converting various content, including original video, audio and photographic documents datable between the 11th and the late 19th centuries, with the aim being to ensure their safe storage.

There is a need to devise uniform standards for the evaluation and quality control of digital copies and digital information carriers.

The introduction of digital television is another big challenge to cope with. We will have to ensure that the quality of digital copies of archival

films and videos matches that of digital TV standards. We will also have to make arrangements for film and video content (recorded in digital-TV-compatible formats and on compatible carriers) to be checked in for conservation so that they could be used in the future.

We will need to identify document digitization priorities and to devise and monitor a long-term stepwise scanning plan for archival documents.

Only with these tasks fulfilled will it be possible to get down to large-scale digitization and to arrange access to digital copies of particularly valuable and demanded paper documents as well as photographic, video and audio content.

Traditional archives are not intended for storing digital content, so archivists have to rely on modern technology for check-in, storage and use of such documents. The results of current digitization efforts will also require adequate storage.

With this in mind, we are now exploring the possibility of setting up a Digital Document Storage Centre.

Work with large digital documents takes sophisticated software and hardware, capable of storing millions of documents, providing access to them and ensuring their effective use, including via the Internet. The primary task, though, is to ensure permanent storage of digital content and scanned documents. Tens or even hundreds of millions of rubles are invested in digitization these days, and we can by no mean let that money goes to waste. Yet another problem is finding competent personnel who would be able to operate all that sophisticated equipment and software in compliance with technical requirements.

We hope our prospective Digital Document Storage Centre will help us resolve these problems, ensuring effective processing and storage of digital and scanned documents. We would like it to run programmes such as Principal Archival Catalogue and State Register of One-off Documents from the National Archives.

One such common centre will be much easier and cheaper to set up than having a relevant unit established for each of the federal archival collections. It could be based on the premises of the Scientific and Technical Documentation Archives, providing safe storage of federal archival documents' scans along with online access, both through in-house networks (for archivists and reading-room visitors) and in the public domain (for the general public).

If fitted with appropriate information security technology, the centre will be able to make documents accessible for viewing without any risk to their safety.

Instructed by the Board of the Russian Presidential Council for Advancing Information Society in Russia, the Federal Archival Agency is now busy preparing blueprints for the digital information storage centre. It is working in collaboration with other federal government agencies.

With conventional documents, the general practice was to accept them for storage 10-15 years after their release. Digital documents should be accepted much earlier than that to guarantee their safety.

This is why we have to make quick decisions as to the formats in which documents could be submitted and the most effective modes of their subsequent digitization and storage. This will let us set uniform standards for archivists and institutions involved in the creation of archival documents.

Yakov SHRAYBERG
Director,
Russian National Public Library
for Science and Technology
Moscow

The Preservation of Digital Information in Libraries: Problems and Prospects

'Tis easy to see and hard to foresee.
Benjamin Franklin

Librarianship has always been connected with the formation and development of documentary information carriers. Libraries have always directly felt technological improvements in the evolution of such media:

- wax and clay tablets;
- papyrus and parchment;
- paper (!);
- microfilms and microchips;
- audio and video magnetic media;
- DVD-ROM and other optoelectronic media;
- computer files, flash storages, LAN facilities; and
- web tools including Internet.

The preservation of originals has always been one of libraries' principal goals – it was their duty to preserve acquisitions for users, providing safe storage and easy access irrespective of internal and external factors.

The preservation of library information had never encountered so many problems as with paper-based materials in the 20th century: however convenient paper might be as document carrier, it is sensitive to whatever environmental changes in the library and outside, and so requires sophisticated storage. Experts warn that 35-40% of the global paper-based library collections are endangered, and 10% damaged irreparably.

Though digital media are also vulnerable, and make authentic reproduction of information impossible with time, this problem has attracted expert attention only quite recently with the need for reliable systems of preserving digital information. Meanwhile, information is being lost, and this matters most – the consequences of such losses are far worse than lost money: one can earn new money while some information can never be restored.

Whatever strategist of information content preservation should answer just two essential questions:

— what to store?

— and how?

The entire set of necessary measures also falls in two essential groups:

— preservation of documents as information carriers; and

— preservation of their content.

The preservation of digital information in libraries and its challenges are studied nowadays in a conscious effort due to the status of libraries in the current societal infrastructure. Libraries collect published documents and preserve them on whatever carriers that might be in use. A library is a social and academic institution with essential educational and other duties. It is universities' partner educationist and a platform to test and launch the latest information and communication technologies. Last but not least, it is a fully-fledged part of contemporary information society. Present-day libraries replenish their collections, as of old, by traditional printed editions while also collecting digital information and micromedia. Though digital content is coming into the foreground and ever new digital media appear, micromedia – mainly microfilms – remain the principal way of preserving library information due to numerous obstacles to the use of digital media for information storage.

In the digital age, library preservation retains two basic goals:

— preservation of original documents on conventional media;

— preservation of digital cultural heritage in compliance with the UNESCO Charter.

It is essential to realize that the current intensive development of technologies for the digitization of library collections makes library information more accessible without guaranteeing its preservation, which presents a major challenge that we owe to the following reasons:

- the evolution and regular replacement of digital information media;
- the development and improvement of system and application software, in which higher versions on many occasions do not emulate the earlier versions authentically, especially when elaborated within outdated operating system models (let us recollect the swarm of problems in working with programs developed in the MS DOS and WINDOWS environments, in spite of the declared complete emulation);
- the change of formats of data representation, and their formal and logical structure;
- the necessity to convert preserved data does not guarantee required authenticity for the same reasons – the different generations of programs, formats and data structures;
- the danger of electromagnetic field oscillations, blackouts, malicious programs, and other factors damaging computer files;
- and, last but not least, the term of digital document service life is vague, for the time being, as expert appraisals are extremely varied while the time to check them is too short due to the novelty of the technology.

Analysts forecast in 2007 that the volume of generated and preserved digital information would grow too rapidly for the available storage facilities to cope, and they would fail as early as 2010 (fig. 1). Practice proved that they were right. The world has not been flooded by information solely due to modernizing these facilities but the threat of their failure is still close at hand.

Reliable digital content preservation systems are of vital importance to libraries for the following reasons:

- Subscription for periodicals is being arranged mainly for digital versions,
- E-book and e-library scopes of application are expanding,
- Library collections are being digitized and digital collections are made accessible to users,
- Website functional sections are expanding, and
- Digital administrative content is expanding as well.

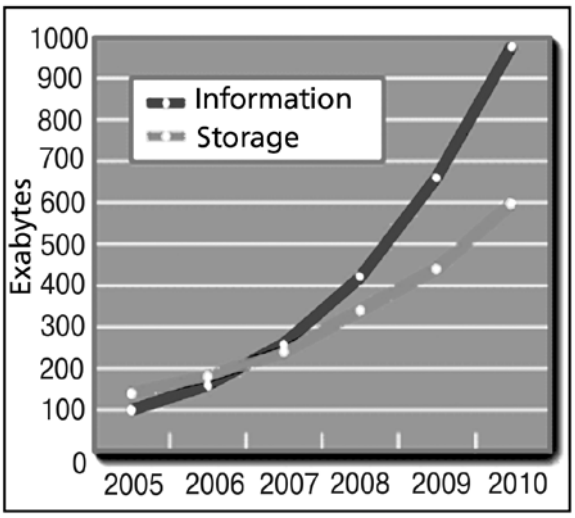


Fig. 1. *Information volume and storage capacity*

> Librarians were alarmed more than anyone else by the problems of preservation bred by the digital age. While the rest of us were gasping at the miracles of the information revolution, librarians haunted us even in the 1980s with the question how to store the new digitized materials for the generations to come.
>
> *Nathaniel Knight*
> Associate Professor and Chair, Department of History,
> Seton Hall University, New Jersey

Let us come back to what present-day libraries form and use in their daily routine in the field of digital content. In the foreground, we see:

- Electronic catalogues,
- Electronic libraries and fulltext systems,
- Databases,
- Computer-based reference information systems,
- Administrative and organizational/management documents in the digital format,

— Printed output,

— Web-technologies,

— and some others.

Evidently, the digital content dominates the library routine and deserves lasting storage. That is why regular computer technologies for information storage are available in the libraries that possess programs for the formation and development of such content or at least that employ advanced IT experts. These technologies envisage representation of several generations of data, on various media, through the use of network resources, and through the establishment of digital archives.

As libraries – mainly national libraries – are legally responsible for the preservation of printed and digital matter, they work on the application of technologies for the retrieval, implementation and archiving of online resources as part of the digital heritage for long-term storage and use.

Almost all developed and many developing economies have national preservation programs for documents on traditional media while large-scale preservation projects for digital documents are still scanty. Russia, for its part, has a national library collection preservation program functioning for many years under the aegis of the Ministry of Culture and recently prolonged to 2020 with an account for the latest technologies – but there is no special program of any significance for digital documents.

Noteworthy examples of the world practice include several digital library content preservation projects: EPPP of the National Library of Canada, launched in 1991, PANDORA of the National Library of Australia, launched in 1991, web-archive of the British websites – the UK Web Archiving Consortium, KULTURARW3 (1996) of the National (Royal) Library of Sweden and many others.

The US Library of Congress is running several digital content preservation programs; of special notice is its web archive kept since 2000.

Though Russia and the other CIS countries have no projects on a comparable scale, they are already aware of the challenge and are drawing such projects. Ukraine's Vernadsky National Library has made major progress in this respect.

Closely connected with the preservation of digital information is the establishment and development of the Open Access Systems and Open

Information Archives mainly because the latter not merely provide access to digital documents but also systematize them and guarantee them reliable storage. 41 countries have announced the establishment of Open Archives for today. According to the DOAR (Directory of Open Access Repositories), they possess a total exceeding 20 million documents.

Russia presently possesses 35 Open Information Archives though DOAR has registered only ten of them. The progress of Russian open archives is gaining pace, and Dspace, Eprint and other software systems for open archives widely used in the world all have Russian-language versions.

Web-archiving is an essential part of digital information preservation. Special agencies are engaged in it in many countries; in many instances, these are national libraries. Web-archiving is a process of collecting part of the World Wide Web and ensuring its preservation in an electronic archive. The archives of websites are highly demanded by historians, journalists and researchers (to say nothing of the value of the preservation function itself). Indicatively, the major world web-archiving operator, Internet Archive (www.archive.org), has served as a basis of establishing the International Internet Preservation Consortium.

Many researchers and managers consider that libraries should be responsible for the preservation of online knowledge, and deserve every support – material, technical and other – in these activities. The preservation of digital resources is not a critical problem yet but certainly an acute one, and demands urgent consideration and decision-making. It is our duty to do everything necessary for the survival of information before it is destroyed beyond retrieval and understanding. The longer we put off relevant decisions the greater the danger of information lost quite soon. All who create digital information must be warned that it might be short-lived. They must take care of its preservation even during the creation of information clusters.

The choice of software demands discriminate attitudes, strict compliance with the acting standards, and guarantees of support by the personnel. These should be duties shared by the entire information community.

Especially noteworthy among the crucial program documents published for today and pertaining to guarantees of preserving digital information are the following:

- The report of the European Commission on the preservation and accessibility, which has been prepared for UNESCO, and CASPAR, DPE, PRESTOSPACE, PLANETS and other EU projects,
- The TEL-ME-MORE project and the eContentplus and i2010: Digital Libraries EU initiatives,
- The UNESCO Charter on the Preservation of the Digital Heritage and documents adopted by the UNESCO summits on information society,
- The resolutions and special events of IFLA-PAC and IFLA-eIFL,
- The World Digital Library and EUROPEANA, and, the last but least,
- The priority directions of the UNESCO Information for All Program for 2008–2013.

Libraries cannot cope on their own with the creation, management and preservation of the present-day digital content. This goal demands active and efficient teamwork and mutual assistance of all information society members.

Plenary Meeting. Information Preservation: Digital Landscape

Alfredo RONCHI
Milan, Italy
Secretary General
MEDICI Framework of Cooperation

Digital Preservation: Cyber Ark, Rosetta Stone or Print?

Preface

In the last few decades we have witnessed two related processes: the increasingly visible inclusion of electronic devices in our everyday lives, and the rush to digital formats. Institutions, organisations and private companies have recently[1] begun to convert their own archives into digital formats. Moreover the general public has also started to convert personal data into digital formats: documents, music, movies, drawings and photos have been converted from their original formats into bit-streams in digital media.

People used to believe (and many still do) that digital formats were the ultimate formats for storing information indefinitely. The idea that texts, images and more in general data can be perpetuated by converting them into digital form is popular and widely supported.

As a result, a significant amount of our documents and data relies on digital technology. But is digital technology really suitable for long-term preservation? And are electronic devices, which are required in order to access information stored in digital formats, durable enough to guarantee future access to this information? If not, what can we do to overcome this problem?

[1] We mean: in the last few decades

"like tears in rain[2]"

The past contains much lost heritage that we can now only imagine: the Battaglia di Anghiari fresco by Leonardo da Vinci, the Alexandria Library, the lost plays of Aeschylus, or the original contents of the Palimpsests and the Egyptian Papyri. Could we lose much of our recent heritage too, after it has been sucked into a "digital black hole"? Will future generations, even those of the near future, be able to access the content produced by this generation? Will they be able to play our electronic music, experience computer art, browse our multimedia archives or play our archaic video games?

Will future generations have the opportunity to enjoy today's exhibits at Ars Electronica? Or the impressive recent products from MIT MediaLab? We rarely stop to consider this fragility until malfunctions, viruses, blackouts or "millennium bugs" occur. Similar considerations can impact on our perception of the "value", even the monetary value, of objects and their everyday use. Of course, we cannot preserve everything; there are products and content that will not necessarily reach future generations. Perhaps we should devote some time to choosing what should be destroyed/ recycled.

However, we do need to leave some objects to future generations; we cannot leave a "black hole" as our legacy. Consider the family photo album. We still have photos from the end of the nineteenth century, but we risk passing on nothing from the early twenty-first century to future generations: leaving them a photo album with lots of blank pages[3].

The ability of a culture to survive into the future depends on the richness and acuity of its members' sense of history.

Problems and issues

The general problem can be structured and analysed. On the top of a theoretical taxonomy we may identify two main branches:

a) Hardware / firmware and software aging and disappearing; such problems are mainly but not only related with preservation of devices and computer controlled functions. This refers to car components, electronic devices such as cameras, video, Hi Fi Stereo, major part of artistic installations and even

[2] Blade Runner directed by Ridley Scott (Warner Bros 1982) - The line "like tears in rain" was never scripted. Rutger Hauer just said that while filming. It became one of the most famous scenes in movie history.

[3] Such a result was shown by an old Kodak advertisement videoclip.

machinery usually designed for long lasting life (i.e. the ones controlled by relays in the past now incorporating micro chips).

A significant overview on such a problem was provided by a 1995 report from the American National Research Council in the following statement: "The fact that most electronic hardware is expected to function for no more than 10 to 20 years raises very serious problems for long-term (more than 20 years) archival preservation. Even if the operating systems and documentation problems somehow are dealt with, what is the archivist to do when the machine manufacturer declares the hardware obsolete or simply goes out of business? There will be an IBM or Sony in the year 2200? If they still exist, will they maintain a 1980-1990 vintage machine? Moreover, it must be realized that no archival organization can hope realistically to maintain such hardware itself. Integrated circuits, thin film heads, and laser diodes cannot be repaired today, nor can they be readily fabricated, except in multimillion-dollar factories."

b) Now leaving hardware preservation apart we can focus on the other side of the coin, "simply" content preservation. An even increasing number of Institutions, organizations and companies store more and more of their organizational information and content digitally. Furthermore, according with Moore's law computer technology makes available new generations of systems, which are more powerful and more cost effective, and thus, are quickly adopted usually losing some backward compatibility. As soon as old technologies are replaced by new generations of hardware and software, new data formats PICT, GIF, TIFF, BMP, JPG are introduced. Over time, we are losing our knowledge of how the previous generations of systems stored and accessed documents, simply consider vector graphics file formats over the last two decades: IGESs, DXFs, DWGs, Open Inventor, VRML, etc.

Within many organizations today, world-wide, digital documents that are official records must be categorized and managed in accordance with approved records schedules that should differ from country to country but basically they force to retain, by law, data for at least a minimum period of time. These records must be retained and accessible throughout their life cycle in accordance with the same laws and standards that govern paper records.

This means usually more than 90 years for medical records, ten years for basic technical documentation and accounting, 75 years for copyrights a "potentially unlimited" amount of time for historical records and state archives.

In the United States, for instance, in the case of government documents the law dictates that an official Government record must be classified into one of 26 retention periods set forth by the Archivist of the United States. These retention periods range from 30 days to Permanent storage and include time periods of 30 years, 50 years, and 75 years. In addition there are special requirements such as archives of copyrighted information created in order to preserve content till the expiration of copyrights and public domain fruition, this means in Europe more than 75 years storage.

Let's take a break and consider another point of view like in the story of optimist looking at the half full glass, the pessimist looking at the half empty glass and the engineer looking at the oversized container. Storing content in digital format we do not only face problems and potential danger, one of the positive aspects of digital object is "cloning", this solves the eternal conflict between preservation and access. In order to add some more positive aspects, Internet technology thanks to de facto standards and almost platform independency has improved the potential life span of its own content.

In order to approach long term preservation of pure content we can list the different levels of complexity and function to be preserved. Content and functionalities to be preserved could range from:

31. Plain text (for a long time the essence of interoperability and long term format);
32. Text with specific formats and functions (such as camera ready pages, spreadsheets, database import/export formats)
33. Multi/hypermedia content (images, movies, sounds in different formats, links, refs);
34. Technical Sketches, 3D models (vector graphics, interactive scenarios);
35. Content deeply merged with specific applications (archives, data base, , video games, custom applications);
36. Interactive installations, enhanced virtual reality applications, etc
37. Future applications involving a higher interaction with users and heterogeneous distributed data structure.

Of course each of such different levels of content complexity requires different efforts in order to be preserved or restored. Is evident that plain text is easy

to maintain whilst videogames requires usually emulation or preservation of original platforms, where emulation means even speed reduction of emulators in order to provide the original performance to the game.

The rationale behind digital preservation

The idea that once you have managed to convert your original into a digital format the hard work is over does not reflect reality; once the digital data have been obtained, it is necessary to consider a different conservation strategy.

One of the first things to consider is that content has its own life cycle. Even though the "paperless" office uses more paper than ever before, and low-cost storage technology has created terabytes of "digital garbage", we must take into account what is relevant and what it is not; not only because preservation will cost some resources but also because irrelevant data will simply increase the "entropy".

Another relevant aspect of preservation is related to the data refresh rate to understand fixed information and dynamic information.

The life cycle of the data will influence its own creation and will generate an accounting record for the resources to be preserved. Since prevention is better than cure, if we define preservation strategies we are halfway to the solution.

Accordingly with the nature of the problem and related potential strategies some interventions are:

- Refreshing – in order to avoid physical decay of data on magnetic storage media, information is simple refreshed every time a predefined time span is elapsed. There are specific software applications in order to look after such a process. If the storage media is write only or a significant portion of the life time of magnetic media has elapsed the system will move data on "fresh" storage media.
- Printing / microfilming – sometimes the last resource, keep data in safe between one generation and the next one. Unfortunately there is some digital stuff that does not fit on paper or microfilm format. Under such a circumstance, technology does not help because delivers everyday new generations of "digital objects" increasingly different from traditional ones. How can we revert a

digital signature in paper format or a cooperative document created in on the fly? How can we store permanently wikis or blogs? How can we manage interactive electronic art installations?

- Multiple Instances (copies) – this should be considered as a human philogenetic approach. In order to increase the opportunity to survive they simply replicate as much as possible (refer to the Meme'eye Web[4]).

Now starts the hard stuff, we are not merely dealing with plain text or simple documents, the three approaches here under are often to be considered as alternatives. The choice amongst them is to be carefully considered both at the moment and long term.

- System Preservation – this is the hard way to keep alive vintage technologies such as FORTRAN II or PDP11. To let old computer machinery work and perform properly for a reasonably long period of time is either very expensive or difficult. Sometimes this approach suggests buying used equipments in order to cannibalize them in case of need.
- Emulation – this approach looks like one of the most attractive even because of the chimera of the "universal emulator" enabling every vintage application to run on up to date computers without requiring the original system to be preserved. Referring to very common applications and single computer system emulation, video games like PAC MAN are usually preserved in such a way. Of course emulators could be shared amongst groups of applications originally running on the same platform. Problems arise if the application needs to interact with some specific software configuration of hardware device (i.e. Sega Goggles or Nintendo PowerGlove).
- Migration – is an alternative both to preservation and emulation of vintage systems, the idea is to migrate/move the digital information and/or application to a new system. Migration technique is very useful and widely used in software industry in order to move applications and database management systems from one system generation to the next one. One of the key points in such a process is to start migration before knowledge related to the application

[4] "The Meme's Eye Web". Susan J. Blackmore argues that memes have shaped human nature. By driving the evolution of the machinery that copies them, memes created the enormous human brain with its specialised capacity for language.

is lost. Dealing with personal information managed by desktop applications, internal data formats and functionalities are not known by the owner of information so he must rely on third parties for migration having no direct evidence of how fast knowledge is lost. In such a case it could help to have an additional copy of content coded in different standards (i.e. Internet technology standards). Open source software should help thanks to availability of the source code but, this is relevant only for users very skilled in software engineering, for major part of content producers is not relevant at all. Forward migration of digital objects to a new system generation is time consuming and costly, in addition if we consider one of the basic philosophical rule of conservation[5], minimize harm, migration should instead maximize harm. Migration process involves data modification; such a modification will degrade the information if the new format will not incorporate some feature of the original format. Some test were carried out in the past, for instance, trying multiple migration in circle coming back to the original format, sometimes the two originals did not coincide.

- Standardisation - means migration from original proprietary formats to standard formats (i.e. Internet technology formats). If we are going to create long lasting digital content we must start from definition of the standards we are going to use taking into account how much they are stable and does support forward migration. The idea to move from proprietary formats to de facto standards means that knowledge of data formats should always be available even if the standard we choose will no longer be supported. Standardisation addresses two of the challenges outlined speaking about migration: loss of information due to loss of knowledge about it, degradation of the content due to multiple migrations. The question of course is how to choose the right standard in order to support the information necessary for organisational preservation and how will a future user determine what standard was actually used to preserve the information. A possible and reasonable solution for a similar problem is the use of metadata; such are XML, enabling the rich description of the standard used by original owners. This process is usually called encapsulation.
- Encapsulation – this resource points to ensure future use of data wrapping the digital object to be preserved and accessed later

[5] Conservation as a countinous process of taking care of things.

on within a human readable skin. The wrapper could contain plain information that both supports organisational preservation and documents the preserved object in order to allow future users to decode and access it. Because there is not a unique and optimal solution encapsulation has two main challenges. The first considering actual applications are the need to generate encapsulated digital data directly from applications. Since current applications are not able to produce such data, a plug in or post processor must be developed and added to existing applications. The second challenge is the potential storage overhead due to the addition of encapsulated information; each record could be enriched with descriptive data. As a positive effect of mixing together preservation methods if we refer to data standard with our encapsulated records we can save a significant amount of storage but how long published standards will be available for?

Migration and encapsulation are in some sense related and dual, but if migration tries to solve the problem in advance when both detailed information about data and systems are available, encapsulation represents potential self-sufficient package of information and data that we deliver to future generation hoping to have included as much information as needed in order to let them be able to manage and read digital objects they received across the time.

Trying to summarize this topic: there is not an unique or optimal solution to preservation problems, the basic approach is to refer to experts and consultants in order to identify the proper strategy and related physical implementation; as ironically stated by Jeff Rothenberg in "Avoiding technological quicksand", "*Digital documents last forever, or five years, whichever comes first*".

It is not only a matter of technology

The management of the whole problem involves several different aspects in addition to bare technological ones, there are administrative, procedural, organisational, legal, IPR and policy issues surrounding long term preservation of digital content. This increased complexity is even due to the different nature of digital versus physical traditional documents.

At least one aspect should be investigated before settling on a particular preservation approach: the overall cost of preservation. This involves considering the best way to ensure future access to information during

the design phase of the long-term data set. This approach may involve some feedbacks on the way to choose technology and standards and even the way to shape data sets. Once the data set is created, in addition to infrastructure costs, running costs may include: additional room on storage devices to archive copies and/or documentation and metadata, software applications that manage data refreshing, and costs related to porting or emulation. Of course the massive use of the Internet, its technology, standards, multiple instances of the same content and different "time machines[6]" in some way helped very much the solve the problem.

A number of global studies[7] and projects have been and are being carried out into digital preservation; for instance the work carried out by the Taskforce on Archiving of Digital Information (94–96) on the mandate of The Commission on Preservation and Access and The Research Libraries Group Inc., as well as the OASIS Open Archival Information System project, CAMiLEON emulation and the VERS Victorian Electronic Record Strategy. Along with the ERA initiative launched by NARA, Interpares I, II, and III are some of the most well known projects in this field. The Ludwig Boltzmann Institute Media.Art.Research & Ars Electronica Futurelab – Linz, Austria launched a relevant research on the preservation of digital artefacts in 2006, Cyber Ark software from Israel developed almost in the same period of time a sophisticated architecture in order to solve this issue. In addition, a comprehensive vision of electronic record management is provided by the US Department of Defence standard entitled the Design Criteria Standard for Electronic Records Management Software Applications (dod 5015.2 STD)[8] .

Closing remarks

The long-term preservation of digital content is one of the big challenges of the digital age; important digital information is in danger of being lost forever. The technologies required by particular types of digital content become obsolete; application versions and files formats frequently change, making data inaccessible. Even when content is coded in the simplest format, such as ASCII code, the degradation and obsolescence of storage

[6] E.g. The Internet Archive – www.archive.org

[7] E.g. MEDICI Framework "Report on long term preservation of digital content", New York City May 2004 // The International Expert Meeting "Conservare il digitale", held in Asolo on 29 September 2006. The report, entitled Long-Term Digital Preservation: An International Focus (see http://www.ndk.cz/dokumenty/asolo_memorandum.pdf/download), was created in order to provide some guidelines and suggestions on this topic.

[8] Available at url: http://jitc.fhu.disa.mil/recmgt/standards.html

media can result in its disappearance. Online information sources such as web pages and databases become more difficult to find as the web(s) in which they are located become more complex and diverse (due to an increasing number of hyperlinks, cross-references as well as types of web application and even webs themselves).

A systematic approach to this relevant topic has to be adopted. First of all we must consider the economic approach in a broad sense; this means some more archiving space, some additional care and work. In addition a background analysis of our documents[9], a tight classification accordingly with basic archival principles and specific archival tagging[10] is required. The expectation is that such basic tagging and the enrolment in the ad hoc preservation "pipeline" will be included within the operating system services. This will significantly help the end user inserting a temporary annotation text file in a different preservation pipeline compared with a contract or a mission critical data base. Starting from this point many different solutions may be applied: stand alone preservation architecture, corporate architecture and service centre architecture.

Recently many IT Centres or Server Farms and "Clouds" started to offer a long term preservation service. Customers store their data on those systems and the host takes the charge to ensure long term access to such data. The increasing use of clouds as storage solution, mainly due to the opportunity to access from any device, anywhere and anytime your information, may support the implementation of this additional feature.

Finally, it is very important that research into digital preservation is carried out by strong interdisciplinary groups, since this should guarantee that an effective approach to a problem that concerns the foundations of the digital era is defined.

Perhaps the only chance of avoiding the "technological quicksand[11]" is to find the "digital Rosetta Stone[12]" in the digital desert.

[9] Text, music files, photos, video clips etc

[10] We may use parallel set of tags to identify relevance, quality, time span, etc. E.g. a mission critical document (to be preserved whatever it happens), highest quality/fidelity of the copy, 75 years, etc

[11] Jeff Rothemberg, Avoiding Technological Quicksand.

[12] Alan R. Heminger, Steven B. Robertson, "The digital rosetta stone: a model for maintaining - long-term access to static digital documents", Communications of the Association for Information Systems, volume 3 article 2 Jan 2000.

References

1. Preserving Digital Information: Report of the Task Force on Archiving of Digital Information - Commissioned by The Commission on Preservation and Access and The Research Libraries Group, Inc. May 1, 1996 -http://www.rlg.org/ArchTF/tfadi.index.htm
2. The DigiCULT Report "Technological landscapes for tomorrow 's cultural economy: Unlocking the value of cultural heritage", Luxembourg: Office for Official Publications of the European Communities, January 2002 - ISBN 92-828-6265-8 – http://www. salzburgresearch.at/fbi/digicult
3. Alan R. Heminger, Steven B. Robertson, "The digital rosetta stone: a model for maintaining - long-term access to static digital documents", Communications of the Association for Information Systems, volume 3 article 2 Jan 2000
4. Duranti L., et al.,Long-Term Digital Preservation: An International Focus proceedings of The International Expert Meeting "Conservare il digitale", held in Asolo on 29 September 2006. The report, entitled http://www.ndk.cz/dokumenty/asolo_memorandum.pdf/download), was created in order to provide some guidelines and suggestions on this topic.
5. National Research Council (1995) Study on the Long-term Retention of Selected Scientific and Technical Records of the Federal Government Working Papers. Washington, DC: National Academy Press.
6. Maria Guercio, "La conservazione a lungo termine dei documenti elettronici:normativaitalianaeprogettiinternazionali",proceedings http://www.unipd.it/ammi/archivio/3conferenza/3%20Conf%20-%20Mariella%20Guercio.pdf
7. Rothenberg, J., Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation, January 1999, Council on Library and Information Resources, ISBN 1-887334-63-7, http://www.clir.org/pubs/reports
8. Rothenberg J. Ensuring the longevity of digital documents. Scientific American. 1995; 272(1):24--29

9. Granger S. Emulation as a digital preservation strategy. D-Lib Magazine, October 2000. http://www.dlib.org/dlib/october00/granger/10granger.html.

10. Wheatley P. Migration - a CAMiLEON discussion paper. 2001. http://www.ariadne.ac.uk/issue29/camileon/. Accessed April 19, 2002

11. Raymond A. Lorie, Long term preservation of digital information, Proceedings of the first ACM/IEEE-CS joint conference on Digital libraries, p.346-352, January 2001, Roanoke, Virginia, United States

12. Lynch C. Canonicalization: A fundamental tool to facilitate preservation and management of digital information. D-Lib Magazine, September 1999. http://www.dlib.org/dlib/september00/lynch/lynch.html.

13. Reich V, Rosenthal DSH. LOCKSS: A permanent web publishing and access system. D-Lib Magazine, June 2001. http://www.dlib.org/dlib/june01/06reich/reich.html.

14. Consultative Committee for Space Data Systems- Reference Model for an Open Archival Information System (OAIS) July 2001. http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html.

15. Lavoie B. Meeting the challenges of digital preservation: The OAIS reference model. OCLC Newsletter January/February 2000; 26--30

16. Attributes of a trusted digital repository: Meeting the needs of research resources. An RLG-OCLC report. Draft for public comment. August 2001. http://www.rlg.org/longterm/attributes01.pdf.

17. OCLC/RLG working group on preservation metadata: A recommendation for content information. October 2001. http://www.oclc.org/research/pmwg/contentinformation.pdf.

18. Research Library Group (RLG) - RLG REACH element set for shared description of museum objects. 1998. http://www.rlg.org/reach.elements.html.

19. National Library of Australia. Preservation metadata for digital collections. 1999. http://www.nla.gov.au/preserve/pmeta.html.

20. Networked European Deposit Library (NEDLIB). Metadata for long term preservation. July 2000. http://www.kb.nl/coop/nedlib/results/preservationmetadata.pdf.

21. Attributes of a trusted digital repository: Meeting the needs of research resources. An RLG-OCLC report. Draft for public comment. August 2001. http://www.rlg.org/longterm/attributes01.pdf

22. Preserving Digital Information, Report of the Task Force on Archiving of Digital Information, May 1996, ftp://ftp.rlg.org/pub/archtf/final-report.pdf

23. The Australian Government Locator Service (AGLS) Manual for Users, Version 1.1, National Archives of Australia and Office for Government Online, August 1999, http://www.naa.gov.au/govserv/agls/AGLS_User_Manual_1.pdf

24. Bearman, D., Sochats, K., Metadata Requirements for Evidence, University of Pittsburgh, http://www.lis.pitt.edu/~nhprc/BACartic.html

25. Bearman, D., Reality and Chimeras in the Preservation of Electronic Records, D-Lib Magazine, Vol 5 No 4,April 1999, http://www.dlib.org/dlib/april99/bearman/04bearman.html

26. Reference Model for an Open Archival Information System (OAIS), Consultative Committee for Space Data Systems, CCSDS 650.0-W-4.0, White Book, September 17, 1998, http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

27. MEDICI Framework "Report on long term preservation of digital content", New York City May 2004

28. Waugh, A., Wilkinson, R., Hills, B., & Dell'oro, J., Preserving Digital Information Forever, CSIRO CMIS Technical Report (forthcoming).

29. Weibel, S., Kunze, J., Lagoze, C., Wolfe, M., Dublin Core Metadata for Resource Discovery, RFC 2413, September 1998, ftp://ftp.isi.edu/in-notes/rfc2413.txt

30. Duranti, L., Eastwood, K., The preservation of the Integrity of Electronic Records, http://slais.ubc.ca/users/duranti/into.html

31. Hedstrom, M., Migration Strategies (Draft), Prepared for Experts Committee on Software Obsolescence and Migration (1996), May 1997, http://www.sis.pitt.edu/~cerar/ftp-docs/Mig-Stra.doc

32. Keeping Electronic Records (Policy for Electronic Recordkeeping in the Commonwealth Government), National Archives of Australia, http://www.naa.gov/govserv/techpub/elecrecd/KeepingER.html

33. Recordkeeping Metadata Standard for Commonwealth Agencies, National Archives of Australia, Version 1.0, May 1999, http://www.naa.gov.au/govserv/TECHPUB/rkms/intro.htm

34. Functional Requirements for Evidence in Recordkeeping, University of Pittsburgh, School of Information Sciences, http://www.lis.pitt.edu/~nhprc/

35. Keeping Electronic Records Forever, Records Management Vision Development, prepared by Ernst & Young Public for Record Office Victoria, 1996, http://home.vicnet.net.au/~provic/vers/kerf.htm

36. Victorian Electronic Record Strategy, Final Report, Public Record Office Victoria, 1998, ISBN 0-7311 5520-3, http://home.vicnet.net.au/~provic/vers/final.htm

37. Management of Electronic Records, Public Record Office Standard (PROS) 99/007, http://www.prov.vic.gov.au/vers

38. Yergeau F., UTF-8, a transformation format of ISO 10646, RFC 2279, January 1998, ftp://ftp.isi.edu/innotes/rfc2279.txt

39. Documenting the Future (Policy and Strategies for Electronic Recordkeeping in the New South Wales Public Sector), State Records New South Wales, 1995, ISBN 07310 5038 X, http://www.records.nsw.gov.au/publicsector/erk/dtf/tofcont.htm

40. Extensible Markup Language (XML) 1.0, W3C, 1998, http://www.w3.org/TR/REC-xml

41. Resource Description Framework (RDF) Model and Syntax Specification, W3C, 1999, http://www.w3.org/TR/REC-rdf-syntax/

42. Levy, D., Heroic Measures: Reflections on the Possibility and Purpose of Digital Preservation, Proceedings of the third ACM conference on Digital Libraries, Pittsburgh, 1998, p152-161

43. The Universal Preservation Format: Background and Fundanmentals, Sixth DELOS Workshop: Preservation of Digital Information, Tomar, 1998, http:/www.ercim.org/publication/ws.proceedings/DELOS6/upf.pdf

44. Australian Standard on Records Management, AS4390-1996, Standards Australia, ISBN 0-7337-0306-2

On Line References

http://www.interpares.org (InterPARES project)

shttp://www.ispo.cec.be/ida (Moreq project)

shttp://jitc.fhu.disa.mil/recmgt/ (Standard 5015.2 "Design Criteria Standard For Electronic Records Management SoftwareApplications" US Department of Defense)

shttp://ccsds.org/RP9905/RP9905.html (OAIS standard)

shttp://sdsc.edu/NARA (National Archives and Record Administration)

http://www.archives.gov/electronic_records_archives/ (18th Annual Preservation Conference Preservation Reformatting: Digital Technology vs. Analog Technology)

http://www.npaci.edu.DICE/Pubs (National Archives Washington & S.Diego Supercomputer Center)

Grethe JACOBSEN
Copenhagen, Denmark
Head of Department
The Royal Library

Webarchiving in Denmark, 1998–2011

Webarchiving in Denmark is done within the framework of legal deposit, which means that those parts of the Internet, considered part of the cultural heritage of Denmark, are collected regardless of any qualitative criteria.

Legislation for legal deposit in Denmark dates back to 1697 and during the first 300 years, the object of deposit was printed works, such as books, newspapers, periodicals, maps, musical notes and printed pictures. Two institutions, The Royal Library[13] (since 1697) and the State and University Library[14] (since 1902), share responsibility for administering the law of legal deposit and for preserving the collections that have been collected through legal deposit. The Royal Library has the primary responsibility for printed materials (with the exception of newspapers) and photographs, while the State and University Library has the primary responsibility for audiovisual media, radio, TV and newspapers.

The internet contains all these media and it was therefore an obvious solution to let the two institutions share the task of webarchiving. The libraries have established a virtual institution, "Netarkivet.dk" (Netarchive.dk)[15] to perform the task of webarchiving. Netarkivet.dk is governed by a Steering Committee of 6 department heads (3 from each library) representing expertise in digital preservation, IT, legal deposit and national collection building. Netarkivet.dk has a daily manager who reports to the Steering Committee.

Webarchiving in Denmark began on a modest scale in 1998, when a new legal deposit law was passed, the Act on Copyright Deposit of Published Works, which was in force from January 1st, 1998, to June 30th, 2005. The law was the first modern legal deposit legislation in Denmark and it redefined the object of deposit as being published works, regardless of medium, rather than printed items. The key here is the term 'published', indicating not a break with the past but a recognition that printing no longer is the only means of publishing.

[13] http://www.kb.dk/en/index.html
[14] http://en.statsbiblioteket.dk/
[15] http://netarkivet.dk/index-en.php.

One paragraph of the law stated that "When works are published in the form of databases ... the person under a copyright deposit obligation must inform the copyright deposit institution of the publication and simultaneously enclose passwords and any other information necessary for the institution to gain access to the works. The copyright deposit obligation is fulfilled by the copyright deposit institution having access to request or make copies of the works." The aim of the paragraph was originally to secure deposit of those publications, that were printed on demand only, but the two libraries convinced the Ministry of Culture to interpret this sentence to let static works, published on the (then emerging) Internet, be subject to legal deposit as well. By "static works" was meant works only periodically updated (typically monographic works and issues of online periodicals) in contrast to the dynamic works (typically databases and homepages) which were not subject to legal deposit.

During the years the law was in force, the libraries collected a total of 42.403 works of which 32.943 (= 78 %) were periodical issues from 1.497 periodical titles and 9.460 (= 22 %) were monographs. The total storage space required for these works is 83 GB. The producers of works, subject to legal deposit, notified the libraries of a new publication using an online form giving the URL, where the publication was to be found. The staff at the Royal Library would then collect the web pages that together made up the work.

During this period (1998-2005) the Internet developed rapidly, and it became increasingly clear that lots of published materials, that would be considered part of the cultural heritage of the country, was not being collected, and that the law and the collection strategies needed to be revised. During this first step, the libraries did, however, gain valuable experience in collecting web-based published material, and the staff involved (librarians, library technicians, IT engineers and computer scientists) became familiar with handling materials published on the web.

The current Legal Deposit law is Act on Legal Deposit of Published Material No. 1439 of 22. December 2004, which went into force on June 1st, 2005. In line with the previous laws the object for deposit legislation is something that is published and therefore available to all whether free or for a fee. According to the current law it can be works in physical form, online material, radio and tv broadcasts and movies.

Part 3 of the law concerns legal deposit of "Material published in electronic communication networks" and section 8 states that "Danish material

published in electronic communication networks is subject to legal deposit" and that such material "is considered to be Danish when (1) it is published from Internet domains etc. which are specifically assigned to Denmark, or (2) it is published from other Internet domains etc. and is directed at a public in Denmark."[16] An accompanying governmental order states that, although the law also covers material published in other electronic communications networks, such as text messages sent to a general audience (for example bus schedules, information from public authorities, commercial announcements) such material is not collected for the time being. This was formulated before the advent of smart phones. The two libraries are currently considering how material published via smart phones may technically be collected.

Strategies for harvesting

The aim of web archiving in Denmark is to get a complete picture of Danish material published in the Internet. This, however, is not possible due to the changing nature of the dynamic and transient nature of this medium that moreover is continually growing and expanding. The Netarchive.dk has, therefore, three strategies for capturing as much material as is deemed necessary to document this part of our cultural heritage.

The first strategy is *Bulk Harvesting* (or cross-sectional harvesting), in which all published material on Top Level Domain DK and Danish materials published on other domains are harvested automatically with staff interfering only when problems arise. Ideally, this should be done four times a year. For technical reasons this has not been possible. During the first six years (July 2005 - July 2011) the Netarchive.dk has completed eleven bulk harvests. A harvest is begun by loading a list of URLs of all domains on Top Level Domain DK, supplied by the administrator of the DK domain, into the harvester program to which is added a list of URLs on other domains with Danish content.

The second strategy is *Selective Harvesting* which is done regularly at (more) frequent intervals to gather materials that are updated in such a pace that it would not be captured by the bulk harvesting. Web pages focused on in this strategy are (1) news sites (national and regional media), (2) "typical" dynamic and heavily used sites representing the civic society, the commercial

[16] English Translation of the entire law text may be found here: http://www.kb.dk/en/kb/service/pligtaflevering-ISSN/lov.html. The original text here: http://www.pligtaflevering.dk/loven/index.htm

sector and public authorities, and (3) experimental and/or unique sites, documenting new ways of using the web (e.g. net art). At present (September 2011) 90 sites are collected daily, weekly or monthly. The list of websites to be harvested is reviewed regularly, and for each website the staff manually sets the parameters for harvesting (how often, how deep).

The third strategy is *Event Harvesting* which is done at frequent intervals for a limited period to collect web pages from sites, dedicated to one event and which is expected to disappear when the event is over. An event is defined as something that (1) creates a debate among the population and is expected to be of importance to Danish history or have an impact on the development of Danish society, (2) causes the appearance of new websites devoted to the event, and (3) is dealt with extensively on existing websites. Typical events are elections, political events involving Denmark (COP15 meeting in December 2009) or sporting events with Danish participants (Olympic Games). This type of harvest also require manual handling as websites have to be found, evaluated as to their relevance, and the parameters for harvesting (how often, which period) set.

The above strategies are strategies for collecting the materials, not for building special collections; therefore, all harvested material merge into one archive, regardless of how it was harvested. However, each harvest is documented, including the list of URLs to be harvested, so each harvest may be reconstructed. The Netarchive.dk takes care to document the harvesting activities in order to provide future users with details on the web archive's content.

Statistics

The first bulk harvest, which took place during the Fall of 2005, gathered 8 TB of data. The eleventh bulk harvest (Spring 2011) gathered 24 TB, that is a tripling in just 6 years. As of September 20th, 2011, the archive contained 236 Terabytes (236.000 GB) with 7,2 billion objects. There are literally thousands of file types but the most common file types comprising more than 90% of all file types are HTML (text) and JPEG, GIF, PNG (pictures) and this has not changed over the years.[17] The Netarchive.dk harvests from 1 million active <.dk> domains and from appr. 44.000 Danish sites on other domains (<.com>, <.org>, <.nu> etc.)

[17] As there are technical problems in harvesting video and streamed picture and sound these media types are underrepresented in the statistics, but the impression is that even if all videos and streamed media was included the picture would not change that much.

Access to Web Archives[18]

Access to the Danish web archive is very restrictive when compared to the general access to the national collections that the two libraries otherwise offer their users. The obstacles are primarily the EU Directive on Data Protection (Directive 95/46/EC on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of such Data) and the national data protection legislation based on this directive as well as copyright law. Danish copyright law is a minor obstacle, as librarians and rights holders by now have much experience in creating agreements of mutual benefit concerning conditions for access and copying.

The major obstacle is the Danish Personal Data Act which limits access to the collected online material to scholars, and for purposes of research and statistics only. The argument is that the harvested material may contain sensitive personal data and should not, therefore, be generally accessible according to the EU Data Protection Directive, as interpreted by the Danish Data Protection Agency. This restriction goes against the aim and purpose of a national library, which is to give access to the cultural heritage of the nation to its citizens, with restrictions determined by preservation considerations and copyright. It also goes against the important aim of any democracy of giving its citizens the possibility to be informed. An obvious example concerns resources related to elections. Any Danish citizen should be able to see the web pages generated during the election campaigns that have been archived and hold the elected politicians accountable if she or he so desires. The Netarchive.dk has collected web pages concerning national elections since 2005 but only researchers are allowed to see this material.

The basic issues behind this legal problem are both formal and political. The formal issue is whether Directive 95/46/EC should be applied to web archiving. The Danish Data Protection Agency has decided that is should. The political issue is the balance between protecting the individual and protecting the public's right to know. At the moment the scale is heavily weighted on the side of the former (protecting the individual), and it requires political action to change that balance. UNESCO could play a most welcome role in promoting this debate as to where the line between protection of privacy and the right to information should be drawn. Netarchive.dk has tried to find technical solutions for the problems of

[18] For a more detailed discussion see Grethe Jacobsen, "Web Archiving: Issues and Problems in Collection Building and Access", Liber Quarterly, vol. 18 (2008), no.3-4 (http://liber.library.uu.nl/publish/issues/2008-3_4/index.html?000273)

access but it has not been possible by automatic means to guarantee 100 % that no sensitive personal data will be accessible to the general public, and so, access remains an issue to be solved by the politicians, not jurists, computer scientists or librarians.

International cooperation

International cooperation is an absolute necessity for collecting, preserving and giving access to digital material. The Netarchive.dk is a founding member of the IIPC (International Internet Preservation Consortium, http://www.netpreserve.org/about/index.php), and this organization has developed into a major player when it comes to preserving online material. In addition, the Netarchive.dk has developed a special program, NetarchiveSuite,19 whose primary function is to plan, schedule and run web harvests of parts of the Internet with a built-in bit preservation functionality and it is built around the Heritrix web crawler, which it uses to harvest the web. It was developed in open source as a way to combine resources and join forces on an international level. It is now used and developed also by the French and Austrian national libraries.

It has been the Danish experience that no library or heritage institution alone can ensure the collecting, long-term preservation and providing access to digital materials. It is imperative that such institutions work together to preserve the Internet,[20] just as the Internet itself is crossing boundaries and connecting people in a manner, that was no one had imagined less than two decades ago. UNESCO would, therefore, have an important role in facilitating such international cooperation.

[19] http://netarchive.dk/suite/

[20] Grethe Jacobsen, "Webarchiving Internationally: Interoperability in the Future?" paper at the 73rd World Library and Information Congress, Durban 2008, http://archive.ifla.org/IV/ifla73/papers/073-Jacobsen-en.pdf

Tatyana BOGOMAZOVA
St. Petersburg, Russian Federation
Head of IT Department
Museum of Anthropology and Ethnography (Kunstkamera)
Russian Academy of Sciences

Meshed in Social Networks: Traditional Memory Institutions in the New Information Environment

> "The internet's main knock is not mainly into households but into the brains, and it is a habitual structure of civilization that is bulging at the seams rather then just structures of everyday life"
>
> *Ivan Davydov, Expert Journal, January 16, 2012*
>
> Wikipedia: I know everything.
> Google: I'll find everything.
> Vkontakte: I know everyone.
>
> Internet: Without me, you are nothing.
> Electricity: Now, now, guys...
>
> *Internet tale*

Pre-conceptualization

Since 1997, the Museum of Anthropology and Ethnography (MAE) of the Russian Academy of Sciences, has been pursuing sociological studies of visitors' perception and specific demands, headed by sociologist Valentina Uzunova. The early stages of the studies proved that lack of backup between the Museum and its audience was a vital problem.

As a result of the studies, the conclusion was made that the sociological portrait of a typical visitor envisaged the following specific characteristics: "The need for dialog, discontent with the passive educational concept of tours presentation, the need to include visitor's personal interests into exposition context" *(Uzunova, 2000)*. We recognized that the multitude of exhibits at display dissected the wholeness of perception and makes it

discreet and intermittent. Even then, long ago, in the scale of information technologies, we set ourselves the task to restore the continuity of perception through information technology introduction and to compel it to the logic of interrelated elements of environment and human statuses within it. We hoped to make up for the lack of dialog between the museum and its visitors taking into account the specific character of visitors' perception.

The Museum directed it further vector of IT development consistently with the revealed visitors' demands – we built out a web-site, collections databases, digitized hundred thousands of archival documents and photos, developed electronic exhibitions of museum rooms, designed game software. Despite indisputable benefits of these developments, one should admit, though, that they got a major disadvantage – they lacked, or nearly lacked the user content.

To be fair, we should admit that not only museums held this conservative didactic position. Most of information owners behaved the same way – they were measuring out and deciding on what visitors had or did not have to see following the logic of limitation and non-admission, exclusion and isolation.

Let us admit – our visitors were too patient. They waited long enough to tell their word. Even a worm will turn. The great McLuhan was right when he warned that the medium is the message. The revolution is now the accomplished fact – being tired of waiting for the mountain, Mohammed has gone to the mountain himself – lacking a spot where to look for the answers to their questions, users rushed where it was possible just to ask them. Incredible pace of technologies, liberation from any hierarchy or social status, limits-free self-expression through universal media, has given users the unique chance to turn to where the personal and individual characteristics, favors, demands were the determining ones to establish communication.

Subjective definitions and objective statistics

The World Wide Web today shows long definitions of social media. I prefer the definition given by Andreas M. Kaplan and Michael Haenlein who define social media as "a group of Internet-based applications, built on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content" (Andreas M. Kaplan, Michael Haenlein, 2010).

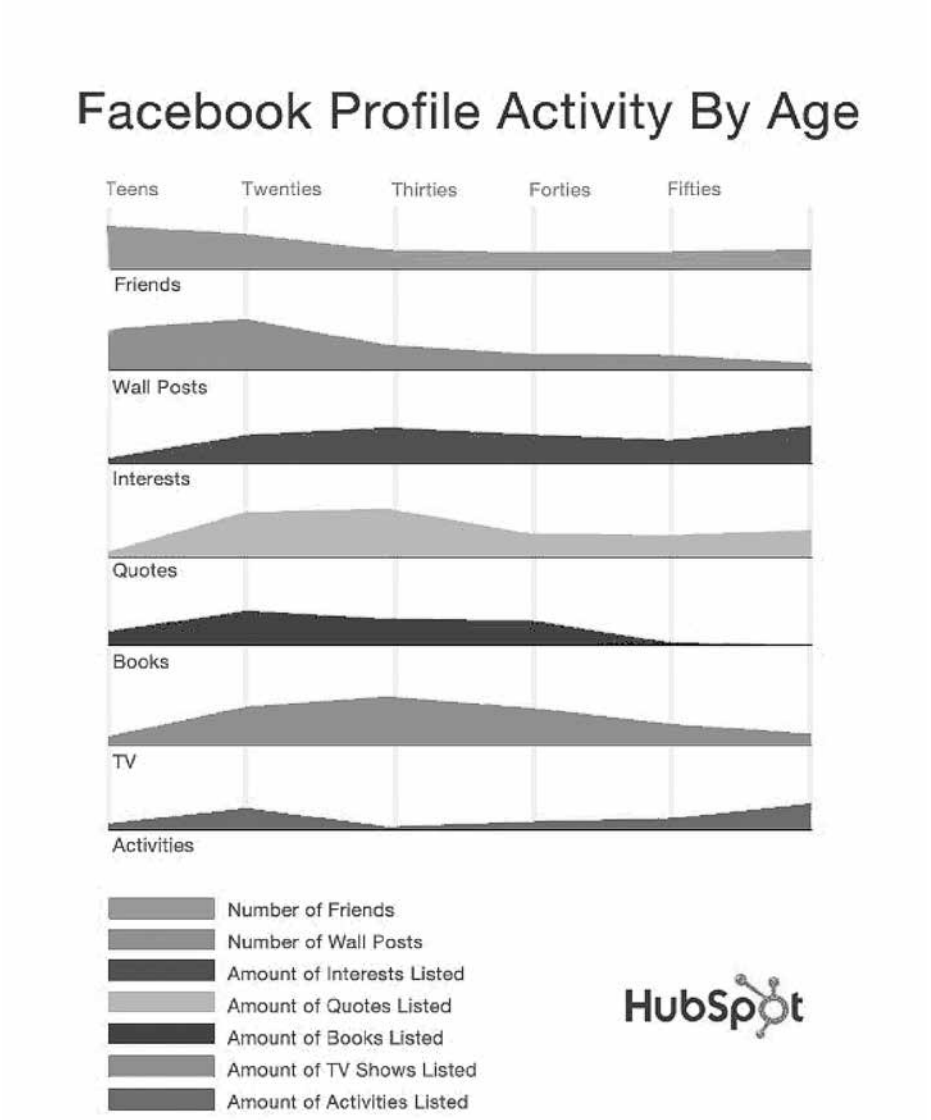
As a result of the global study undertaken by Pew Research Center in the period from March 21, 2011 to May 15, 2011, social media use reached the following values: 53% of population use them in Israel, 50% - in the USA, 43% - in Russia and the UK, 42% - in Spain, 39% in Lithuania and Poland. Out of 21, in 15 states at least 25% of respondents use social networks. Among the group, Russia is the only country where almost every Internet user uses social media sites.

It is extremely symptomatic that just 6% of Russian Internet-users claim they never visit those sites, while in Germany – 35% of users claim the same, 35% in France and 32% in China.

Since 2010, insignificant statistical changes in social media usage have taken place in most of the reviewed states, excluding Egypt and Russia – the countries where the role of social networks, in particular, was in the public focus during recent political events. In both countries, in one year their use increased by 10% - from 18% in 2010 up to 28% in 2011 in Egypt, and from 33% up to 43% in Russia, correspondingly *(PewGlobal.org)*. The fact that the young users have been increasingly using social media via cell phones is seen as extremely significant.

Among social networks, Facebook has been still leading, with its 800 m accounts, 200 m of which were created in 2011. The Royal Pingdom estimates the mobile Facebook (FB) auditorium at 350 million users. The number of Twitter registered accounts makes 250 million, with 250 million twits posted daily. According to Google's estimates, as of June 2011, the number of (Russian-language) Vkontakte (*'InContact'*) social network got 38 m unique users with their number rapidly increasing. Meanwhile, as of March 1, 2012, 118,731,729 personal profiles were registered on Vkontakte (the author has got the value through using the network's advertising tool and setting the widest range of target audience).

Let us apply statistical tools to answer the question *what does communication mean within social media*? The statistical control units are: number of friends, posts on the wall, claimed interests, quotes, or reposts, books, TV programs and activities marked by users. HubSpot reports that Facebook activities prevailing in the United States, Europe, Australia, and the Middle East are distributed in the following way:



Thus, definite age priorities in information activities are revealed to be considered when building social media.

Let us point out other impressive values. In 2011, every twelfth earthman used FB: 48% of users age 18-34 check FB when they wake up, 28% of users check FB on their cell phone before going to bed every night. It is important that over 30% of FB users are over 35, while the increase of number of FB users aged 18-24 (mainly students) makes 74% a year.

To proceed with statistical values, let us try to "measure" the intensity of media use through the types of information activities. Thus, GO-Globe.com estimates that the following events occur on the World Wide Web every 60 seconds:

- 694,445 retrieval requests in Google;
- 6,600 images loaded into FLICKR;
- 320 new users in Twitter with over 98,000 twits;
- 320,000 talks via Skype;
- 13,000 audio records played in Pandora;
- 13,000 IPhone apps downloaded;
- 600 new videos uploaded on YouTube;
- 168,000,000 e-mails sent;
- 1,700 loads of Firefox browser;
- 695,000 FB «statuses» updates;
- 79,364 posts on FB;
- 510,000 comments to FB events.

Unfortunately, we failed to find studies on the World Wide Web to dissect these activities into the segments of application sphere, including those related to museums. Nevertheless, we analyzed the representation of Russian museums on *Vkontakte* social network, primarily, in the format of official communities and pages. The list of largest (over 500 members) official museum communities:

1. The State Hermitage Museum – 15,386
2. Erarta (St. Petersburg) – 13,048
3. Mikhail Bulgakov Museum in Moscow ("The Odd Flat") – 3,940
4. Freud's Dream Museum (St. Petersburg) – 3,022

5. Moscow Museum of Modern Art – 2,970
6. Novgorod State United Museum Reserve –2,637
7. The State Museum of the History of Religion (St. Petersburg) – 2,446
8. Russian Museum (St. Petersburg) – 2,300
9. The State Tretyakov Gallery (Moscow) – 2,171
10. PERMM (Perm) –1,983
11. Museum of the History of Ulan-Ude City – 1,722
12. Kunstkamera. Peter the Great Museum of Anthropology and Ethnography (St. Petersburg) – 1,386
13. Nizhny Novgorod State Art Museum – 1,314
14. New Museum (St. Petersburg) – 1089
15. State Historical Museum (Moscow) – 1020
16. The Russian Academy of Fine Arts Museum (St. Petersburg) – 958
17. Museum of the History of Photography - 886
18. The Museum of Fine Arts of the Republic of Karelia - 819
19. All-Russia Decorative Art Museum (Moscow) - 769 участников
20. Ekaterinburg Museum of Fine Arts - 766
21. The Cathedral and Immanuel Kant Museum (Kaliningrad) - 759
22. Museum of City Sculpture (St. Petersburg) – 746
23. Semenkovo Museum (Vologda Oblast) – 688
24. Samara Literary Museum – 647
25. P.V. Alabin Museum (Samara) – 617
26. Yaroslavl Art Museum – 616
27. Togliatti Art Museum – 562
28. Ekaterinburg History Museum – 531.

The similar rating of museums' official pages on *Vkontakte*:

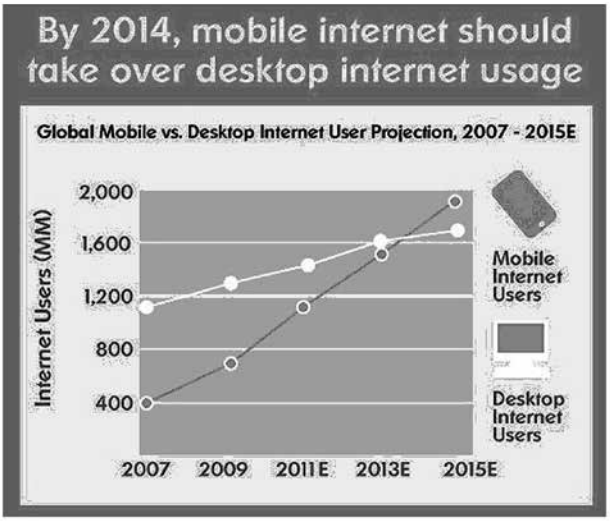
1. Kunstkamera. Peter the Great Museum of Anthropology and Ethnography (St. Petersburg) – 4,221
2. Erarta (St. Petersburg) – 2,008
3. Dostoevsky Memorial Museum (St. Petersburg) – 1,204
4. Russian Museum (St. Petersburg) – 716
5. Samara Literary Museum – 601
6. State Museum of Political History of Russia (St. Petersburg) – 600
7. Ryazan Art Museum – 551.

To be fair, the search for pages and communities in V*kontakte* network was accomplished mainly by the word "museum" in the communities name, nevertheless the findings (data as of March 1, 2012) are seen as very lamentable, especially as compared to the Metropolitan Museum of Art (612,198 friends) and Louvre (489,376 friends) on Facebook. It is also symptomatic that the recognized museum leaders top the lists followed by much more popular and, what is more significant, newer museums, which proved to be more dynamic than many 'heavyweights'.

About forty more Russian museums of various departmental affiliations have got their groups though the membership does not exceed 500, or mainly 200 members. It should be said, though, that there are examples of the more successful 'Kulturtrager" experience on *Vkontakte* social network – for example, the community of Petersburg *ETAGI* Loft Project numbers 46,377 members, while *Vinzavod* (Center for Contemporary Art) community numbers - 35,915.

The situation with museums representation in other social networks, excluding VK, and in cloud services is even worse, and I believe, it is too early to analyze them seriously. Nevertheless, it is worth mentioning top indicators related to Facebook – The State Hermitage page has got 11,726 followers, and the Tretyakov Gallery – 5,157 followers.

In conclusion, let us turn our attention to the forecasts for Internet access products. By some estimates, by the year 2015 mobile Internet will leave broadband desktop access behind by 15%.



The analysis findings prove that the most significant factors to evaluate the prospects for social media usage by museums visitors belong to three main planes, i.e.

1. Degree of visitors involvement in social media communities moderated by the museums;
2. Types of information interaction of visitors with museum content both quoted from original museum sources and that created and interpreted independently by users;
3. Localization of users relative to a museum at the moment of information interaction with social media.

Information policy

Having stated the synchronic nature of the problem, let us look for the origins of such non-dynamic and inefficient usage of social networks by museums, with users that make today 43% of population, i.e. all Russia's active population.

It is well-known that this is the result of museum communication hierarchical pattern. I will cite a long passage from the recent publication by a media-expert Ivan Davydov who wrote: "In any communication sphere, the possibility to communicate was achieved through passing

through a complex filtration net... Indeed, filters - depending on applied efforts – differed, but not everyone was able to get into a newspaper, moreover, on TV, university rostrum, or parliamentary tribune. That was the standard. And here the Internet appeared on the stage... It came to remind the public that communication is inherent to human nature rather than it is an art, and it does not require any additional instruments rather than a mouth. The public believed and started to speak... The world is not divided into groups of people who have the right to speak and those who have to listen. Everyone talks today... It is needless to say that no one listens. Hierarchies that cemented culture are being washed away before our eyes... The only hierarchy that keeps relevant is the hierarchy of request answers. The first ten links retrieved by a search engine are important. We would not browse them further... The Internet has made the culture absolutely accessible — and absolutely redundant. In the world where people are equal, laws have retroactive effect, as the Internet equalizes not only those who are present here at the moment but also those who were creating culture. I would say classics and contemporaries. Appeal to authorities does not work because there are no more authorities" *(Ivan Davydov, 2012)*.

I agree, the quote is rather provocative, though in many ways the author has captured the current trends antagonistic to the very letter of the law the museums and museum specialists had followed for centuries. It does not mean, though, that museums have to give up their expertise and, to some degree, their tutorial role. However, gradually they will have to change their relations with visitors in the information realm.

Museums in North America and Western Europe recognized long ago the need to re-orient museums information flows from internal circuit to outer directions. The alternative to restricting information is maximum information access, to licensed media – completely open public resources, to subordination and authoritarianism – partnership, to acquisition process – provision for creation, to organizational structure – self-organizing chaos, and to management tasks coming top-down – manageable initiatives taken by staff vested with ample powers *(Dana Allen-Greil, 2011)*. As applied to museum practice, this means into the transition from didactic approach to assimilation into user communities, from tutorial paradigm to the paradigm of partnership and influence, from exceptionally institutional content to user-generated one.

Thus, sooner or later modern museums will inevitably transform from being ***content supplier***s to ***information platforms*** for users to create and exchange information, and to unite into communities. Then museums tasks will significantly change, as they will have to provide visitors with the possibility to:

- record their personal responses
- communicate various opinions to communities
- involve visitors
- inspire dialog
- support collective creative practices.

Taking the above into account, a museum information platform can be defined as a system of data exchange, knowledge and opinions on a museum being transmitted online.

The most significant step on the way towards a museum information platform is building its internal information system available to visitors via mobile Internet when they first step in a museum building, or adjacent territory which, in its turn, will enable younger visitors to make information-integrated visits.

Measuring social media

And finally, let us discuss measuring social media, which means monitoring their efficiency through analyzing number of blogs, wikis, micro blogs, the above mentioned user-generated contents in social networks, traffic on web-sites of video/photo sharing, forums, message boards, etc.

Today, [Russian] museums scarcely make use of available free instruments and corporate platforms to monitor social media. When choosing such instrument, museums should take into consideration the presence of several ideologies of collection and analysis of quantitative information on social media. The major players in the market, such as *Klout* and *PeerIndex*, measure agent's influence taking into account agent's voice volume and competence, while *SplashMedia* monitors efficiency in social media according to a set number of social media measuring correlations between activities (entries) and results (outcomes).

Both approaches have their strong and weak points. Errors of impact measurement with *Klout's*, or *PeerIndex'* tools are rather significant – their estimations would suggest that M. Gandhi or F. Roosevelt are not so

much influential while Brad Pitt is one of the most influential persons on the Internet. On the other hand, when measuring social media, efficiency might mean that a local hairdresser can become very "efficient" while his or her media coverage makes just one thousandth of his or her colleagues' one.

There are more applications for quantitative instruments of social media measurement; thus, *Gephi*, *Weka* and *RapidMiner* that enable to build network visualization, as well as to evaluate aggregate mood of a community, or to design a chain of computer algorithms to classify online content by its attribute.

In our opinion, implementation of the above mentioned tools into museum practice is very important as this will contribute to the solution of measurement tasks, but also will enable to accumulate and save at least analytical information on the role and significance of social media in the life of a given museum.

However, it is topic of other studies which undoubtedly will be undertaken by IT experts in museums and other cultural institutions in the nearest future.

References

1. V. G. Uzunova, Sociology of Museum Studies. Studying specific demands and perception of MAE visitors// 285th Anniversary of St. Petersburg Kunstkamera (Digest of the Museum of Anthropology and Ethnography, 48). – [In Russian]. // St. Petersburg, 2000. P.325.
2. Andreas M. Kaplan, Michael Haenlein. Users of the world, unite! The challenges and opportunities of Social Media // Business Horizons, 53 (1), 2010. P. 59–68. http://www.pewglobal.org/2011/12/20/global-digital-communication-texting-social-networking-popular-worldwide/
3. Ivan Davydov. Hierarchy Busters // Expert Journal, №2 (785).January 16, 2012. – [In Russian]. http://expert.ru/expert/2012/02/razrushiteli-ierarhij/
4. Dana Allen-Greil, National Museum of American History, Smithsonian Institution; Susan Edwards and Jack Ludden, J. Paul Getty Trust; and Eric Johnson, Monticello, USA. Social Media and Organizational Change // Museums and the Web, Philadelphia, April 6-9, 2011.
5. http://www.museumsandtheweb.com/mw2011/papers/social_media_and_organizational_change

Andris VILKS
Riga, Latvia
Director
National Library of Latvia

Access to Digital Content – Strategies, Opportunities and Threats

Latvian National Digital Library Letonica

Latvian National Digital Library Letonica (NDLL) is a project of the National Library of Latvia (NLL). The project aims at the digitization of the collections of the National Library of Latvia and other cultural heritage institutions and to make them accessible on the Internet.

The project aims at ensuring:

- Preservation and storage of cultural heritage in the digital format to prevent valuable content from being lost and ensure that future generations have access to digital materials;
- Digitization of analogue collections and accession of digitally born materials for their wider use in the information society;
- Wide online accessibility of digital content as a precondition for maximising the benefits for citizens, researchers and companies;
- Adoption of international standards, protocols and data formats that foster research in the digital library sphere; development of a common digitization methodology for all Latvian cultural heritage institutions and dissemination of best practices;
- Cooperation at the national and international levels, including serving as the national aggregator for Europeana library and other international projects.

The main long-term objective of the project is to digitise and make as widely accessible as possible the entire Latvian cultural heritage and to develop an IT infrastructure and services as well as common standards and methodology to support the above.

History

First efforts towards digitization in Latvia were made in mid-1990's by the Academic Library of Latvia who digitised a 10-volume collection of the historian, chronicler and law drafter Johann Christoph Brotze entitled 'Sammlung verschiedener Lieflandischer Monumente' that contained thousands of drawings and descriptions as well as documents on the history, architecture, and everyday life of Riga, Latvia and Estonia. The Archives of Latvian Folklore have digitized the Cabinet of Folksongs included in UNESCO's 'Memory of the World' programme. The Institute of Mathematics and Computer Science of the University of Latvia has digitized ancient specimens of the Holy Bible, dictionaries, grammar books and sermon books from collections of the National Library of Latvia and the Academic Library of Latvia and is currently working on the national language corpus based on these works. The National Library of Latvia started its work on digitization in 1999 with the collection of 19th century newspapers.

The idea to join digitization efforts of Latvian cultural heritage institutions took shape by the beginning of 2001. An informal working group was established with the aim to harmonise standards and develop pilot projects. The initiative was backed up by the Ministry of Culture and acquired the status of a pan-Baltic co-operation project. In 2002, the first Baltic workshop on digitization brought together representatives of libraries, archives and museums from all the three Baltic states, as well as guest speakers from Finland, Sweden, Norway, Denmark, Czech Republic, Scotland and Great Britain, Netherlands and other countries. The Ministry of Culture established a standards committee for all the three heritage sectors under the auspices of the library standards committee. The next step was the transformation of the Consortium of library information networks into the National Agency of Cultural Information Systems and the development of requirements for cooperation between archives, libraries and museums.

To foster digitalization co-operation, a national target programme for 2002 – 2005 was launched by the State Culture Capital Foundation. In the framework of this programme, national and regional memory institutions could implement joint interagency digitalization projects. About twenty joint projects were thus implemented on the digitalization of texts, visual and audio materials, audiovisual and 3D objects. In addition to the digitalization of collections, the greatest achievement of

the programme was best practices in digitization identified in the process. The implementation of the programme led to the conclusion that much work was to be done in the field of metadata and standards harmonisation. The success of the programme convinced the Ministry of Culture of the necessity of a long-term support for digitization projects.

As a result, in 2005, the government of Latvia decided to support the idea of establishing the National Digital Library. In 2006, the Latvian National Digital Library Letonica project was launched under the management of the NLL with the aim to establish a unified basis for all digitization efforts in Latvia.

The project's primary task was to create an integrated technical infrastructure for digital objects management. In 2007, the National Library of Latvia, Cultural Information Systems and Microsoft signed a three-party agreement on the development of a communication platform for a digital library, which was put into production in 2009.

Furthermore, much work was done to develop standards and methodology. International standards and best practices were adopted which became a basis for „The Digitizer's Handbook" for those involved in digitization processes. Best practices in the digitization of various types of materials formed the basis for two thematic pilot projects, one of which was dedicated to the famous Latvian composer and music master Jazeps Vitols (2006) and the other comprised materials of Latvian national song and dance festivals (2008). The most prominent pilot project was the new collection of periodicals, which supported full text search. In order to complete the project the existing OCR technologies had to be revised to adapt to the peculiarities of the Latvian Fraktur script.

Collections

Today, the Digital Library comprises digitized collections of newspapers, graphics, maps, books, printed music and audio records.

- Newspapers
  - ▷ *Heritage 1.* The first attempt in newspaper digitization. More than 100 historical titles scanned and available as PDF files.

- ▷ *Periodika.lv*. A more advanced approach to newspaper digitization. 350 000 pages scanned, OCR-ed and segmented into articles.

- Graphic documents
  - ▷ *Poster in Latvia 1899 – 2000*: ca. 3300 posters of social and commercial activities.
  - ▷ *Latvian History of Civilization in Images. Part A: Portraits*. 3700 portraits (drawings, daguerreotypes, silhouette portraits, and photographs) created prior to 1914 from the collections of the NLL and the State Archive of Latvia.
  - ▷ *Bookplates and Prints*. Ca. 3000 bookplates and prints created by many renowned Latvian artists.
  - ▷ *Postcards*. Ca 2000 postcards – portraits, art postcards, greeting cards.
  - ▷ *Collection of the Baltic Central Library*. 50 photographs of Riga.
  - ▷ *In Search of Lost Latvia*. The latest accession to NLL digital collections. Contains ca. 17 000 postcards, drawings and photographs of Latvian historical and contemporary views.
- Maps
  - ▷ *Latvia in the 16th – 18th Centuries*. 130 maps of Latvia and its historical regions.
- Sound Records and Scores
  - ▷ *Scores of symphonic music by Latvian authors*. 7 digital transcriptions of scores of popular Latvian tunes which had not been published earlier and available only in the written form.
  - ▷ *Wax cylinders*. 5 records from wax cylinders stored at the Department of Rare Books and Manuscripts.
- Born digital resources
  - ▷ *Academia*. Open repository of scientific publications issued both at NLL and partner institutions.

- ▷ *Web harvesting* has been used since 2005; in 2011, 3000 sites of the .lv domain will be available to public.

- Special collections
  - ▷ *Latvian Song and Dance Festivals*. Latvia has a unique and ancient tradition of holding Song and Dance Festivals (organized every four years). Historical materials on these events from the first Song Festival in 1864 to the Latgalian Song Festival of 1940 are open to examination within this digital collection created in cooperation with the Folk Art Centre.
  - ▷ *Jazeps Vitols*. This collection is devoted to the renowned Latvian composer, conductor and music master Jazeps Vitols. The collection was created in cooperation with the Jazeps Vitols Latvian Academy of Music, the Centre of Early Music, The Latvian Academic Library, The Latvian State Archive of Film, Photo and Audiovisual documents.

Ongoing and future projects

The National Digital Library of Latvia is currently participating in numerous projects, including two projects funded by the European Regional Development Fund. These projects are expected to lead to the following results by the end of 2012.

- *Increased digital content arrays*. 2 million newspaper pages and 1.5 million book pages will be digitized and made available in the new interface that provides the display of printed materials.
- *Resource aggregator*. A resource aggregator will be put in place, which will allow users to search in all digital collections, data bases and catalogues of NLL and its partners simultaneously.
- *IT infrastructure*: enhancing storage and processing capacities; procedures for lifelong digital preservation; centralised user management; improved integration of the DOM system and other IT digital library systems.
- *New user interfaces*. Specialized user interfaces that enable displaying maps, receiving and reading content on mobile devices, providing services to children and handicapped people.

- *New data infrastructure*. The system will be developed towards implementing principles of Linked Open Data, establishing national mechanisms for indentification.

Interagency cooperation

One of the strengths of the National Digital Library is emphasis on cooperation with organizations from various sectors.

- *Library sector*. The NLL is the centre of the library sector in Latvia and serves in this capacity also as a methodological centre for digitization efforts in the library sector. The NDLL involves libraries in its digitization projects and assists them with its technical knowledge and capacities;
- *Cultural heritage institutions*. The NDLL strives to involve also museums and archives in the digitization process, though co-operation here is still underdeveloped. The State Agency of Cultural Information Systems is the main driving force of computerization in these sectors, but efforts and resources at the national level lack harmonization. The NDLL has concluded separate partnership agreements with many museums and is acting as a project coordinator and technical and methodological centre;
- *Publishers*. NLL has always had strong ties with the publishing sector and strives to involve them as well in the development of a digital library. The NDLL currently concludes agreements with publishers on the permanent preservation of the digital copies of their manuscripts and cooperates on the development of the Latvian modern language corpus.
- *Scientific and research institutions*. The NDLL has launched its version of an open access repository and is invites to cooperation scientific and research institutions interested in a wider dissemination of their publications. Besides, the NDLL participates in numerous scientific projects in the area of social sciences, for instance, in national programmes or projects within the Network of National Identity, National Scientific and Research Centres.

International co-operation

The NDLL also takes part in international projects, among these:

- *Europeana*. The NDLL is the national aggregator of digital content from Latvia. It also participates in pan-European projects aimed at the development of content for Europeana, such as Europeana Local and Europana Travel;
- *The European Library*. The NLL as a Conference of European National Librarians (CENL) member participates in the latter's project 'The European Library' (TEL), which serves as a prototype for Europeana. TEL is currently largely overshadowed by Europeana, but remains a dynamic partner for Europeana and one of the largest and most stable content providers, acting as a content aggregator for all national libraries in Europe. The most significant project developed by TEL in close cooperation with the NLL is the project 'Reading Europe', which includes major literary gems from all over the Europe.
- *The World Digital Library*. The World Digital Library (WDL) is a joint project of UNESCO and the Library of Congress, which focuses on the world heritage treasures rather than on the simple quantity of content. The NDLL offers content included in both the world register and the Latvian national register of the 'Memory of the World' programme.

Today, digitization is certainly the most urgent issue on the agenda of cultural heritage institutions, since it implies numerous risks alongside its advantages. Some of these risks, especially those associated with preservation aspects, have been extensively examined. Digitization is a relatively new discipline, and we do not have answers to all the questions of ensuring the sustainability of our efforts. We do not know what will happen to paper in hundreds of years; we still have a very vague idea of how to ensure that data digitized today be readable in twenty years' time. Shall we be able to cope with increasing costs of digitalization in the light of globally shrinking resources for the cultural sector? Research is carried out on all these issues, and we are approaching the stage at which we can find solution to some of the challenges we face.

There are other risks that have not yet been sufficiently examined. There is, for instance, the issue of the policy for digital collection acquisition. More and more information exists in digital form only, but we cannot be sure which should

be retained for future generations. We harvest already the World Wide Web, but is it enough? There are hidden treasures in myriads of forms of personal electronic communication – social networks, e-mail, instant messaging, SMS. If we fail to safeguard these types of materials now, they will soon be lost forever – and quite a lot already are. I am afraid that for this reason we have little chance to reconstruct the vibrant thought of today's cultural elite.

Another risk is the quality of digitization. It seemed to us that the electronic form of knowledge organization would allow for an easy creation of information context, but this has not happened yet. While mass digitization is economically feasible, its results in terms of quality are often but a shadow of knowledge. If digitized materials are published as they are now, without any contextual information about events, persons, places described in the documents, without annotations targeted at specific audiences, they are of little value to the reader. The main aim of digitization should be to help the reader to reconstruct the reality described in documents. It is certainly a very important task which cannot be achieved by one organization; it can only be done through joint efforts. Cultural heritage institutions must be on the frontline of scientific research aimed at developing semantic web technologies. They should apply the linked open data principles to address the emerging challenges and do their best to ensure that the quality of digitization meets expectations.

However the biggest risk of all is our competence. The key to success will not be content proper; it will be the availability and inclusion of digital content in the user's workflows, the quality and ease of the use of services. In the hyper-competitive information space, the services provided by cultural heritage institutions can easily become useless and fall off the map of information landscape if they do not meet the requirements of the moment. To tackle this risk, cultural heritage institutions should develop totally new competences, including IT and social media.

The creation of digital objects without their preservation and the creation of metadata is a way to nowhere. Content creation without a planned and reliable long-term solution for long-term preservation implies the risk to be denied access in the long-term prospect and requires adequate equipment. There are no organizations in the world that could develop recommendations for preservation standards. Objectives and policies for long-term preservation and safety must be developed by every organization on its own, as well as by each memory institution and at the national level.

Sergey KLESHCHAR
Tula, Russian Federation
Head of Department
Research Institute of Reprography

Long-Term Preservation of Digital Information by Means of Modern Electronic Micrographic Technologies

The Federal State Unitary Enterprise "Research Institute of Reprography" conducts studies into the creation of high- quality digital copies of records on various types of media.

It has developed and introduced advanced electronic micrographic technologies to be used in the creation of high-quality information resources to ensure long-term preservation of and efficient access to documented information.

Following the dictates of the time, this technology ensures:

- Production of backup copies of records with authentic content both in analog and digital formats on compact media,
- Long-term preservation of documented information when exposed to various extreme conditions,
- Possibility to confirm the legal validity of backup copies of records,
- High-level protection from unauthorized access, distortion and deliberate damage of documented information,
- High reliability and functional stability of the system in wartime and emergency situations,
- Long-term stability of media, technologies, equipment and software,
- Efficient search and issue of records from the backup collection upon inquiries of different types,
- High quality of the information provided and its comfortable perception,
- Telecommunication access to the stored backup information,

- High level automation of information monitoring and management based on digital data processing.

According to certain estimates, in 5-10 years, over 50 percent of the records preserved by Russia's enterprises will be created and circulated in the digital format (though impressive, this number is still falling behind the economically developed countries).

The electronic micrographic technology is based on the possibility opened up recently to qualitatively transform analog information into the digital format, which is recorded on a digital medium, and to make a reverse transformation of the digital information from the digital medium to the analog format, which is recorded on a microfilm by means of electronic microfilming.

Nobody knows yet what should be done to ensure long-term preservation of machine-readable information. The point at issue is to enable not only reliable copying of digital files but also their transfer to new media and new equipment. An important dictate of the time is a possibility of transferring the most valuable pieces of information to other types of media in order to ensure their long-term guaranteed preservation for a period of no less than 500 years.

The electronic micrographic technology, which uses an optical filming procedure, allows for recording paper-based information on a micrographic medium.

A micrographic medium may be at any time digitized to provide access to the recordings. Another option is to digitize an archive copy and record it as a machine code on a micrographic medium in order to ensure its long-term preservation.

The electronic micrographic technology allows for transferring to micrographic media, at minimal costs, the already digitized collections of documents for the purpose of their long-term preservation. Today, we propose to transfer to micrographic media the most valuable and unique information which is stored today in the digital format and whose loss may be irrecoverable. To ensure the transfer of digital copies of records to micrographic media for their long-term preservation digitization should meet stricter requirements.

Since the role of digitization in Russia is becoming increasingly significant in record circulation and creation of arrays of electronic records, from operating collections for use to backup collections for long-term

preservation, it is critically important to conduct a thorough analysis of, as well as develop and introduce the national normative documents regulating the processes of creating and preserving the collections of digital records.

The lack of standardized approaches to the evaluation of scanning quality results in spontaneous and unjustified projects accepted for local digitization. This impacts the output and poses technological and organizational problems.

Among the already elaborated and adopted documents are "Procedure of Preparation and Delivery of Digital Records for Backup Microfilming" and "Microfilms of the Backup Collection of Records Produced within COM-Systems". However, these documents have been introduced as recommendations only. The Institute of Reprography strives to improve the applicable normative documents. By the end of the year, we are planning to develop the Scanning Quality Control Methodology which is to be transformed into a national standard.

Creation of high quality digital collections is not the only challenge. Another challenge is to ensure their long-term preservation.

According to the global practice, the possibilities of long-term preservation of digital records are limited by frequently changing generations of digital media and the accompanying hardware and software platforms which rapidly turn obsolete and extinct. This is an established fact. In search of a way out of this situation various methods were proposed to extend the life of digital records in the digital environment. The most popular solutions are migration of records to novel software environments and formats, their regular and repeated copying on new media, and emulation, i.e. imitation of an older software environment in new operating systems and on new equipment.

However, the methods of migration and emulation suffer from the intrinsic limitations of the digital environment, which is dynamic, changing and unstable. Continuous migration and emulation require high financial and labor resources. Moreover, it has been proved experimentally that these processes do not ensure protection of information from loss when frequently rerecorded or reformatted. Thus, they do not guarantee that information will be preserved in its unchanged and original form.

For this reason, scientists and experts are trying to develop more reliable and cost-effective archiving strategies for the most important digital information by using the long-term preservation technologies that do not require

continuous updating and maintenance. At this moment they turn their attention to microfilms - time-proved analog media with a huge potential.

The primary goal is to ensure a long life of digital information in an acceptable form and guarantee its integrity. An archive microfilm is the most suitable platform here since it is a technologically autonomous medium which provides a guarantee of information preservation for up to 500 years, as well as its integrity and stability due to a minimal intrusion into the preservation process.

However, the question is how to combine an analog medium – a microfilm – and the digital content of electronic records? To this end, an integration of digital and analog technologies has become necessary in micrography. An opportunity for such integration appeared in the early 1970s after the invention of COM-devices which allow for exhibiting text and graphic information on microfilms. From this time on, these devices have gradually developed and improved to become rather popular in the late 20th and early 21st centuries. Today, there are around 20 models of COM-systems offered by the leading producers. These systems differ in recording methods, microform types, source file formats and other technical parameters, but all of them are capable of recording digital information from computers to film media. The latest achievement in the COM-systems is the development of a laser color system capable of high quality and fast recording of digital information on a color film.

Together with scanners, COM-systems can be legitimately called a key element of modern electronic micrographic technologies, a certain bridge between the digital and analog worlds. Several years ago, the imperfections and drawbacks of certain models and a general enthusiasm for the rapid evolution of digitization technologies gave some scientists a cause to believe that microfilm as a medium was getting out of date and that COM-systems could be used locally only for preserving specific types of digital records. However, the failure of many long-term preservation strategies forced researchers to reconsider their views and return to the traditional microfilm already as a digital information preservation medium whose long-term and stable potential can be reinforced by the capacities of modern COM-systems.

COM-devices have changed drastically the procedure of archive microfilm creation. Instead of using it for creating a traditional optical image, this technology is used to read binary data of a digitized image and record on a

film the position of each pixel by means of a laser (directly) or similar devices. A variation of this technology is a recording device capable of transferring screen images to microfilms. This became possible due to the development of new graphic maps and special ultra high resolution screens. Modern COM-devices can read most of the popular text and graphic formats, as well as new devices allow for enhancing the output quality regardless of the quality of the incoming original images.

The role of COM-systems in the preservation of digital materials is confirmed by the official adoption of the ISO 11506 standard "Archiving of electronic data - Computer output microform (COM)/computer output laser disk (COLD)" which entered into force in 2009. This standard became the first in the global practice to codify the strategy of long-term backup preservation of digital information by means of computer recording on microfilms for long-term preservation and on laser optical disk for expeditious use.

As far as Russia is concerned, there are currently around 50 COM-systems of various types and producers in service. The main users of these devices are organizations and institutions involved in creation and provision of content to the Russian single backup collection of records, and other organizations which recognize the necessity of long-term backup preservation of their own information assets. Russian market of such equipment seems to be sufficiently well-developed. Virtually all major producers of COM-devices are presented on it. However, in spite of all their advantages, electronic micrographic technologies are unable to ensure preservation on microfilms of audio and video records, three-dimensional records, software documentation and some other types of digital information. That said, recent innovations in the COM-systems have significantly enlarged their scope of application. For example, according to the latest studies a theoretical and experimental base has been provided for a new approach to the preservation of digital information, such as audio and video files, on microfilms.

In the light of the latest scientific achievements, the technological potential of the electronic and micrographic technology for the long-term preservation of digital information seems to be very encouraging. It is obvious that new technologies still need to be improved; additional research and tests are needed in the selection of recording parameters, mode refining, equipment compatibility, system settings optimization, technical and economic analysis, etc. However, the first steps have already been taken and further research into the application of this promising technology will undoubtedly be undertaken.

Plenary Meeting. Policies and Education for Digital Information Preservation

Susana FINQUELIEVICH
Buenos Aires, Argentina
Senior Researcher
National Council for Scientific and Technical Research
Director of the Research Program on Information Society
Instituto de Investigaciones Gino Germani
University of Buenos Aires

Live Memory. Digital Information Preservation: Proposal for Public Policies for South America, With Focus on Argentina

The main objective of this paper is to identify the current policies for digital material preservation in South America, with a particular focus on Argentina, and to make proposal for the development of national and macro-regional strategies and policies[21].

This document is based on a collaborative research, which was carried on by way of a survey of relevant publications, both locally and internationally, complemented with interviews with key informers. The research has implied reviewing relevant existing documents in the field of Information Society Planning, legislation, policies and declarations; diverse countries' expertise in the field of Information Society planning and legislation regarding the preservation of digital information (explicit national digital agendas; national, regional and local Information Society policies; national and regional legislative measures; etc.); and relevant national and international documents in the field of preservation of digital information planning, legislation, policies and declarations.

Introduction: Defining the subject to be discussed

The massive growth in the construction and dissemination of digital objects by governments, authors, publishers, corporations, academicians,

[21] Paper is coauthored with Mr Carlos Brys (Computer Science Department, National University of Misiones, Argentina) and Ms Elida Rodríguez (National Institute of Social Services for the Retired, PAMI, Argentina)

and others, has emphasized the speed and ease of short-term dissemination. Nevertheless, little concern has been paid yet to long-term preservation. Digital information is characteristically more delicate and insubstantial than traditional technologies such as paper, photographs, or microfilm. It is more easily corrupted or altered, without detection**.** According to Russon (1999), digital storage media have shorter life spans and require access technologies that are changing at an ever increasing pace. The time frame between the creation of the object and the need for its preservation becomes shorter. The scientific and technical community risks the loss of valuable information without an adequate infrastructure for digital archiving and preservation

It is first of all necessary to define the subject that is being discussed. According to a definition provided by the National Archives of Australia (2001, quoted by the UNESCO Digital Preservation Project, South African Research Team, 2006) the term 'preservation' includes all actions that can be taken with the aim of ensuring the current and long-term survival and accessibility of the physical form, informational content and relevant metadata of archival records, including actions taken to influence records creators prior to acquisition or selection.

According to the mentioned document (UNESCO, 2006) "'Digital preservation' can be defined as the process and activities which stabilize and protect reformatted and digital authentic electronic records in forms which are retrievable, readable and usable over time (NARS, April 2004). 'Digital preservation' involves a number of organized tasks associated with a variety of technical approaches or strategies for ensuring that digital resources are not only stored appropriately, but also adequately maintained and thus consistently usable over time (UKOLN, 2004). It involves the processes of maintaining accessibility of digital heritage materials over time, for as long as they are needed (UNESCO, 2003: 34, 157)".

Quoting IFAP – UNESCO: "In a world increasingly being shaped by digital technologies, the traditional guardian institutions (libraries, archives and museums) are challenged to keep pace with the rapid growth in information. They also face a new challenge – as technology advances the stability and lifespan of documents is considerably decreasing. If nothing is done, many important documents in electronic format will not survive or will become completely inaccessible within a very short time. The result will be a permanent loss to the collective memory of humankind. This challenge needs to be tackled urgently and the costs of preserving digital

information should not be underestimated – these far exceed the preservation costs experienced to date with five millennia of traditional documents. Digital preservation also contributes to at least two other IFAP priorities – information for development and open & multilingual access to information. Digital technologies open up access to information and knowledge in democratic dimensions that has never been experienced before".

For the authors of this paper, digital preservation is a way of preserving living memories, customs, cultures, and identities.

A policy can be defined as a development or general plan of action embraced by a governmental organization, multisectorial body, party or person. A National Policy for Information Society (NISP) can be defined as a roadmap, a national, regional, or local plan for the inclusion and appropriation, by Governments, institutions, communities and individuals, of the benefits derived from the construction of an Information Society. *The NISP is a highway, not a harbour. It is a process, a collaborative, open, and permanent building task. In order to travel this highway, it is necessary to envision it, to plan and build it, to make it travelable for all the citizens* (Finquelievich, Rozengardt, Davidziuk, and Finquelievich, UNESCO, 2010).

A digital preservation policy would state the principles and long-term direction that would guide preservation strategies and actions. A policy statement would set clear priorities. In terms of the South American context, a national and macro-regional digital preservation policy would reinforce mutual efforts of governments to preserve materials which document our laws, public administration actions, science and technology efforts, and cultural heritage. It would define organizational roles and (funding) responsibilities. The roles of established memory institutions, such as libraries, archives and museums, would also be addressed.

The document has been written as a series of questions and answers that deal with the subject of policies for digital material preservation in South America, with a particular focus on Argentina, the social agents that have developed actions in this area, and proposals towards the development of national and macro-regional preservation plan strategies and policies.

1. *Do South American Digital Agendas contemplate the subject of policies for digital material preservation?*

Most original NISPs in South American countries do not plan policies or strategies for digital material preservation, although it is expected that the

processes of updating these NISPs will. However, some countries, such as Brazil and Colombia, participate in Inter Pares, an international project supported by the British University of Columbia and The Social Sciences and Humanities Research Council of Canada's Community-University Research Alliances (SSHRC-CURA). Inter Pares member countries are Africa, Brazil, Canada, US, Cataluña, Colombia, Korea, China, Italia, Malaysia, Mexico, Norway, Holland, Belgium, UK, Singapore, and Turkey.

Policies for the preservation of digital material are undertaken, often jointly, by South American libraries, museums, and University, but governments have not yet arrived to formulate joint policies. However, the issue is currently being discussed at the Union de Naciones Suramericanas, Unasur (Union of South American Nations), the Regional organization that includes twelve South American countries, due to an Argentine initiative.

Another macro-regional agenda is eLAC 2007, the Regional Action Plan for the information society, which was officially approved at the Regional Latin American and Caribbean Preparatory Ministerial Conference for the World Summit on Information Society on June 10, 2005 in Rio de Janeiro, Brazil. It is based on constant dialogue and cooperation among all the Latin American and Caribbean countries, and leading to the adoption of a common policy agenda.

Regarding the preservation of electronic information, Elac 2007 proposed:

- Article 15.4: Contribute to the use of electronic/digital signatures in governmental procedures, both by public officials and civil servants and by citizens
- Article 15.5: Promote the adoption of information security and storage models at all levels of government with a view to engendering trust in the digital information managed or provided by the State.

2. *Which antecedents exist in South America regarding the preservation of digital information?*

One of the first events to raise the problem of preservation of the digital heritage was on the agenda of an expert consultation for Latin America and the Caribbean that UNESCO organized in November 2002 in Managua, Nicaragua. The results of the consultation would be the regional input to the Charter on Digital Heritage that UNESCO was organizing. In that opportunity, Dr. Isidro Fernandez Aballi, from UNESCO, spoke about

the context of the digital era in the XXI Century, as well as about the huge quantity of information which exists in digital format, and which disappears in cyberspace, without being transferred to an updated, more durable support. Dr. Fernandez Aballi encouraged the professionals which were present to respond to UNESCO´s call, contributing their criteria to save the digital patrimonies.

In 2007, the Ibero American Chat of Electronic Government[22] (CIGE) was approved by the Ibero American Conference of Ministers of Public Administration and State Reform in Pucón, Chile, on June 1st, and adopted by the XVII Ibero American Summit of Government and State Chiefs, organized in Santiago de Chile on November 10th, 2007.

Among the Electronic Government Principles stated in paragraph 24, the CIGE enunciates:

> *"g. Principle of technological adequacy: the administrations will choose the most adequate technologies to satisfy their needs. It is recommended the use of open standards and free software based on reasons as safety, security, long term sustainability, and to prevent the privatization of public knowledge. In no case this principle implies any limitation to the citizens' right to use the technology of their choice in the Access to public Administrations".*

Moreover, there are a number of initiatives undertaken by "memory guardian institutions": museums, libraries, and universities. One of them is the Latin American and the Caribbean Network of Social Sciences Virtual Libraries (Red de Bibliotecas Virtuales de Ciencias Sociales de Am rica Latina y el Caribe[23]), created by the Latin American Council of Social Sciences (CLACSO). Its Virtual Reading Room features more than 4000 complete texts of books, papers, documents, and journals. 168 research centres (of which 54% are universities) in 21 countries contribute their publications. Copyrights rights remain in the hands of the authors and the original editors, with a Creative Commons license for their dissemination with academic goals. This license certifies documents´copies, dissemination, and public communication, with the condition of recognizing the credits in the ways specified by the author/editor, for non-commercial uses, and of not altering the contents without previous authorization. CLACSO uses Greenstone fir its virtual library. It is a set of software programs designed

[22] http://www.segib.org/documentos/28/esp/Carta%20Iberoamericana%20de%20Gobierno%20Electronico%20Esp.pdf
[23] http://www.clacso.org.ar/biblioteca

to create and distribute digital collections, providing a new way to organize and disseminate information though the Internet, or CD-ROM.

3. *Which are the advances in Argentina regarding protection of critical infrastructures and the preservation of digital information?*

From 2009 onwards, Mr. Eduardo Thill, Under Secretary of Management Technologies in the National Cabinet of Ministers has reinforced the argentine participation in the Meridian Process[24], with the goal to establish the necessary mechanisms to guarantee the protection of the information´s critical infrastructures, as well as long term preservation of digital information. The goal is to establish the necessary mechanisms to ensure the protection of security In information critical infrastructures. This is particularly relevant within the context of the governmental actions on digital inclusion, such as Argentina Conectada and Conectar Igualdad, which will provide the whole country with optic fiber connections.

As a consequence of this work, the Ministers' Cabinet created the National Program of Protection of Critical Infrastructure of Information and Cyber Security, by the Resolution JGM N° 580/2011[25], within the National Bureau of Information Technologies - ONTI.

Argentina has developed the ArCERT Program (Coordination of Emergencies in Argentina's Teleinformatic Networks), while supporting Latin American countries (México, Belize, Costa Rica, El Salvador, Guatemala, Nicaragua, Panamá, República Dominicana) which are developing policies for information protection and preservation. Argentina is currently encouraging the creation of a Macro Regional network for to transfer and share information and knowledge on information preservation policies, with the goal to strengthen the public administrations' information systems.

[24] The Meridian process aims to provide Governments worldwide with a means by which they can discuss how to work together at the policy level on critical information infrastructure protection (CIIP). An annual conference and interim activities is held each year to help build trust and establish international relations within the membership to facilitate sharing of experiences and good practices on CIIP from around the world. Participation in the Meridian process is open to all countries and aimed at senior government policy-makers. The Meridian process is founded on the G8 principles that provide a basic framework for understanding and implementing CIIP measures. As new challenges of connectivity and dependencies arise beyond national borders, Meridian enables Governments to explore the benefits and opportunities of cooperation with the private sector, and exchange of information and good practices in CIIP between governments internationally. Tools to raise awareness and share information include the CIIP Directory to facilitate intergovernmental contacts and the Traffic Light Protocol to facilitate distribution of information.

[25] http://www.agendadigital.ar/docs/res_580_2011_inf_criticas.pdf

The First International Congress for the Protection of Critical Infrastructures[26] took place in March 14th, 2011, in Buenos Aires, Argentina, at the Argentine Enterprise University, UADE. The Congress, which gathered argentine and international experts, showed the Argentine government and the corporative sector´s interest in working jointly to generate a conscience for the protection if the critical and Cyber security structures, as well as to consolidate future actions, strategies, and policies.

In this Congress, Mr. Thill stated the importance of achieving the users' social engagement in critical infrastructures. The proposals for a policy on the preservation of digital information state that, even is such policy should be coordinated by the national governments, its design should include representatives from the corporative sector, Universities and "memory guardian public equipments", and it should not be the work of isolated countries, but that public strategies on these issue should have a macro regional scope.

4. *Which are the Argentine initiatives towards the preservation of the digital information?*

The limits of time and space allow us only to mention a few examples in a vast universe of digital repositories and libraries. In the Registry of Open Access Repositories (ROAR, 2008), Argentina is represented by three reservoirs: SciELO Argentina (registered in year 2000), la Revista Cartapacio de Derecho Published by the Universidad Nacional del Centro (registered in 2004), and the CLACSO Network of Social Sciences Virtual Libraries of Latin America and the Caribbean (registered in 2008).

In the Directory of Open Access Repositories (Open DOAR, 2008), Argentina is represented by other three repositories: The Digital Library for Identity, the Academic Memoir of the Faculty of Humanities and Educational Siences of the National University of La Plata, and the Service for the Dissemination of Intellectual Creation, SeDiCI (registered in 2008), also from the National University of La Plata.

As a whole, 15 repositories have been identified in Argentina.

It is worth mentioning the Project ECO-PORTAL MERCOSUR (Institutional Repositories for the Academic Intellectual patrimonies of the Faculty of Economic Sciences, University of Cordoba). The project

[26] http://www.agendadigital.ar/index.php/component/content/article/6-principal/107-avances-de-argentina-en-materia-de-proteccion-de-infraestructuras-criticas

was started in 2008. Repositorio Institucional del Patrimonio Intelectual Académico de la Facultad de Ciencias Económicas, Universidad Nacional de Córdoba). At present, it gathers the research works that were disseminated in the Internet, and offers the academic community a tool to publish their academic work.

Moreover, in the National Congress, in September 2011 took parliamentary state a law project proposed by deputy Giannetasio, regarding the creation of institutional, open access digital repositories, in which the public institutions and organizations integrating the National Science, Technology & Innovation (SNCTI), and funded by the National State, must deposit the scientific and technological production resulting from the work of their researchers, technologist, post doctorate fellows, and postgraduate students. This S&T production includes all the documents (journal papers, S&T projects, academic thesis, etc., which result from research work funded by the State, published and unpublished, and that undergo a process of quality evaluation.

It is worth remarking that the Province of San Luis is the first state in Latin America that has not only digitalized all the public administrative information, but it also currently generates exclusively digital documentation in the public administration. Public administration has moved from working with paper to work exclusively in digital support. It works solely with digital documents and digital signature.

5. *Preserving digital information means to decide which information is to be preserved, and which has to be discarded. Which are the criteria that will guide these decisions?*

"Funes the Memorious" (original Spanish title: "Funes el memorioso") is a fantasy short story by Argentine writer Jorge Luis Borges, published in 1942. "Funes the Memorious" tells the story of a fictional version of Borges himself as he meets Ireneo Funes, a poor, ignorant teenage boy who lives in Fray Bentos, Uruguay, in 1884.Funes reveals to Borges that, since a fall from a horse that had left him crippled, he perceives everything in full detail and remembers it all. He remembers, for example, the shape of clouds at all given moments, as well as the associated perceptions (muscular, thermal, etc.) of each moment. In order to pass the time, Funes has engaged in projects such as reconstructing a full day's worth of past memories (an effort which, he finds, takes him another full day), and constructing a "system of enumeration" that gives each number a different, arbitrary name. Funes is incapable of

Platonic ideas, of generalities, of abstraction; his world is one of intolerably uncountable details. He finds it very difficult to sleep, since he recalls "every crevice and every moulding of the various houses which [surround] him".

The fact that Funes remembers absolutely everything, that his memory occupies his brain, makes him incapable of conceptualizing and generalizing. One of the issues of the preservation of digital information is that if we lose our memories, we lose our identities. But if we preserve all the information, we will be invaded by it. The question, then, is to decide which information is to be preserved at each technological turning point, and which will be discarded.

Digital preservation is a complex problem. We´re talking about the set of processes and activities that ensure continued access to information and all kinds of records, scientific and cultural heritage existing in digital formats, and for long periods of time. This includes the preservation of materials resulting from digital reformatting, but particularly information that is born-digital and has no analog counterpart. In the language of digital imaging and electronic resources, preservation is no longer just the product of a program but an ongoing process. In this regard the way digital information is stored is important in ensuring its longevity. The long-term storage of digital information is assisted by the inclusion of preservation metadata.

Ken Thibodeau states that probably, the only valid prediction on the future of information technology is that it will continue to change constantly. Every preservation system conceived as a final solution, even if it seems to solve all the problems of fragility and obsolescence of the tools known until then, will become inevitably obsolete in a relatively short period. This is why any decision about digital preservation should include the capacity to accommodate to fast technological changes, and to incorporate the new products generated by information technology. The only relatively durable factor in a digital preservation solution is the adopted conceptual scheme: the criteria and strategy that will be used.

The Argentine digital preservation strategy conceives preservation as an institutional responsibility, with multi stakeholder participation, political support, and a firm engagement from all the participant agents. According to Serra y Serra (2002), the definition of a preservation plan has to find answers to the following questions: a) which information is to be preserved, and why?. b) Where is it going to be preserved?. c) Until

when is this information going to be kept?. d) How will it be possible to find it later?. e) What actions are necessary to keep these materials unaltered?. f) Which measures have to be taken to avoid obsolescence?

The Argentine National government´s strategy has chosen to create and select digital collection with duration according to its administrative and cultural importance. Work is being developed to establish a well-defined preservation policy, establishing which information is to be preserved, and though which norms and procedures. This policy has to be periodically revised and updated, in order to define the new technological supports, to improve the preservation methods, and to redefine the sets of object to be preserved. Since some information packages or objects are more durable than others, these periods of preservation should also be periodically assessed and updated.

According to Bia and Sanchez, the evaluation of electronic documents and objects provides two main advantages:

- It avoids costly migration methods for documents meant for short term preservation. The identification of such documents previously to taking decisions about technical migration and recopying allows reducing costs and technical difficulties.
- It allows advancing in the transference of the electronic documents destined to long-term preservation to historical archives, when the office or department that produces them cannot assume the costs of implementing a policy for digital preservation.

6. *Which agents should take the decisions about digital preservation?*

Governments should generate, implement and coordinate the policies and strategies regarding digital preservation, but it should be an open process that includes multistakeholder participation. Other agents, such as enterprises specialized in information preservation, Science and Technology institutions, Universities, libraries, museums, technicians, and NGOs, data storing centres, among others, should be able to contribute their experiences, know-how, and opinions.

a. *Which issues should be considered by a South American policy of preservation of digital information?*

Some recommendations could be:

- To conform multistakeholder agencies to make de fundamental decisions about short, medium and long term preservation of digital materials.
- To inform and sensitize the diverse social agents about the importance of the preservation of digital materials.
- To engage the compromise and involvement of the enterprises that deal with information preservation.
- To ensure that the digital materials kept as file documents are stable and fix, in content as well as in their form.
- To ensure that the digital materials are adequately identified.
- To guarantee that digital materials are classified in logical aggregations.
- To use authentication techniques which promote the maintenance and preservation of digital materials.
- To create and periodically update norms, legal frameworks, and procedures for long term preservation of digital materials.
- To protect digital materials from non authorized interventions.
- To assess documents and digital materials to determine the periods of their preservation.
- To make campaigns to sensitize the population about the value of digital preservation.
- The criteria used when it is decided to transfer documents form paper or film to digital materials are valid only within the context in which these criteria were adopted. Prospective studies could help when deciding which digital materials will be most useful for future generations.
- The choice of the stocking technologies should be careful and constantly updated, since there is no physical device that can guarantee the perpetual life of the digital contents.
- The intellectual property costs (copyright) in the digitalization processes are relevant when deciding the policies and projects, and

should be considered when making the selection of the materials to be preserved.

- It is necessary to elaborate legislative actions to protect information resources; an option could be the open knowledge licenses, such as Copyleft, Creative Commons, among others, as well as to clearly explicit the limitation for the use of digital contents.
- Quality is a significant element. It is necessary to capture the best intellectual and visual contents, and then to choose the forms to present these contents to the users in the way that responds better to their needs.
- The selected digital resources should be apt to work in diverse types of informatics platforms.
- The responsible agent is the State through its cultural institutions such as: Secretaries of State, Ministry of culture, national libraries, etc. The second responsible agents should be the institution that sponsors the development of universities, research centers, libraries, etc. These organizations have to emphasize that all digital development project is accompanied by a digital preservation policy to ensure the preservation of the institutional developments.

Conclusions

In summary it can be said that a piecemeal groundwork is currently and gradually being built in South America and in Argentina. It appears that South American countries have advanced in digitalizing their vital information, but public policies and strategies for the long-term preservation of information are still to be designed and implemented.

Moreover, the process of digitalizing information does not imply per se the preservation of knowledge. In order to ensure this preservation it is necessary to implement clear and unambiguous policies of long term digital preservation.

There is a basis of existing partial legislation and policies, which can be used to draft a national long-term preservation policy for South American digital materials. The South American Government's e-government policy can be used as a point of departure, but needs to be urgently updated with reference to the establishment of e-depositories. Some South American

academic institutions, particularly in Brazil, Colombia, and Argentina, have well-established e-depositories. Much debate is still required. There is a pressing need to work together in regional institutions such as Mercosur and Unasur. A national long-term digital preservation policy is needed for South America with the purpose of enabling cooperation between major stakeholders and to be discussed on the political agendas of Government. The following elements need to be debated:

- Ensure political security against loss, access denial or falsification.
- Provide open, easy digital access (with the exception of externally sensitive, legally protected data).
- Use open, standard and non-proprietary formats, make regular migration to current formats.
- Retain formal and semantic content of documents.
- Context retention: references and cross links to other documents.
- Ongoing translation into appropriate specialist language of the time.
- Rules for selection of sources (based on content, expected impact, future users) regardless of type of object, distribution or publication method, publishing house, author category or document type.

When analyzing the main issues related to the preservation of digital information, it is possible to conclude that technological issues are not the most important. Many experts believe that the main defy faced by preservation is located in the institutional area (the design and implementation of public policies), the economic area (the adequate funding to develop such policies), and the legal framework (clarifying intellectual property issues, and defining who have the legitimate right to exploit and access these digital contents).

One of the main difficulties regarding South American strategies and policies for the digital preservation of information is consists in making South America governments aware of the urgent need to include the preservation of digital information in their National Digital Agendas, as well as in regional policies, among other goals, to register and preserve the diverse stages if the public policies and strategies for building feasible Information Societies.. A relevant problem is that the responsibility of the preservation of digital information is distributed among diverse social agents.

Diverse experts usually propose campaigns of sensitization and information of civil servants as a driving factor that will determine a positive change of attitudes towards digital preservation. This implies assuming *a priori* that individuals and groups which are taught to identify and acknowledge the relevance of digital preservation will change their attitudes towards this issue, and even will generate actions towards creative solutions and strategies. It is generally believed that individuals an public administrations do not generate actions regarding the preservation of digital material because they are not provided with the adequate information, and that when they receive this information, they will change their attitudes and become proactive agents of the cause of long-term digital preservations. Nevertheless, while the dissemination of information remains a valuable tool, it is not enough, by itself, to generate behavioral changes.

Which factors do engender such necessary changes in attitudes, policies, and strategies? We believe that it is not enough to "sensitize" the public administration using reasonable, politically correct arguments informing about an obvious problem, that entrails present and future dangers about collective memory and identities. The fact that the loss of collective memory through the successive losses due to the lack of policies, and of consequent and consecutive losses at each technological turning point, can be perceived by public administrations as a relevant problem is based on the ethic foundations and principles that the diverse types of social actors use to assess and decide action priorities, in individual as well as in collective ways.

This is a complex issue, in which individual and collective behavioral patterns are clearly determinate by institutional cultures and bureaucratic rules that frame, and may encourage or inhibit each course of action. This is the main reason why it becomes absolutely necessary to integrate the issue of preservation of digital information to South American National Digital Agendas, as well as to generate an active debate at Unasur and Latin American level.

These policies and strategies could consider the following issues, among others:

- Providing ample information, seminars, conferences, virtual courses, and publications on international best practices on the subject.
- Generate multistakeholder spaces for participation and actions.

- Provide the individuals and groups that have generated positive changes in strategies and actions regarding the preservation of digital materials with appreciation, recognition, and empowerment.
- Facilitate the international mobility and learning exchanges of public administration members concerned with the preservation of digital materials.

References

1. Manuel Sánchez Quero and Alejandro Bia Platas. (2002). "Desarrollo de una política de preservación digital: tecnología, planificación y perseverancia", Biblioteca Virtual Miguel de Cervantes, Universidad de Alicante, Spain. ISBN 84-688-0205-0, págs. 41-50.
2. International Federation of Library Associations and Institutions (IFLA) (2002). "Guidelines for digitization projects for collections and holdings in the public domain, particularly those held by libraries and archives". Available at: http://www.ifla.org/VII/s19/pubs/digit-guide.pdf
3. Finquelievich, Susana, Adrian Rozengardt, Alejandra Davidziuk and Daniel Finquelievich. (2010). "Public Policies for Information Society. A Template", IFAP – UNESCO, 2010. Available at: http://works.bepress.com/susana_finquelievich/3
4. Fernandez Aballi, Isidro, Editor (2007): "Building National Information policies: Experiences in Latin America, Unesco", Information Society Division, Kingston, pp. 20–27 Available at: http://unesdoc.unesco.org/images/0015/001528/152806m.pdf
5. Fyffe, Richard et al. (2004). "Preservation planning for digital information. Final report of the HVC2 Digital Preservation Task Force". University of Kansas.
6. Hedstrom, Margaret et al. (2003). "Invest to save. Report and recommendations of the NSF-DELOS working group on digital archiving and preservation". National Science Foundation & European Union.

7. Jiménez León Alejandro and Maria Graciela Gutiérrez Vallej. (2007). "La preservación digital, un valor agregado en el desarrollo de contenidos digitales", en 2º Congreso Internacional de Innovación Educativa. Ciudad de México, November, 2007

8. Russon, David. (2004). "Access to Information: Now and in the Future", in: Science for the XXI Century, FORUM I , I.6 Sharing Scientific Knowledge, UNESCO, Budapest, Hungary, 2004.

9. Serra Serra Jordi. (2002): "Estrategias de Preservación de Documentos Electrónicos: El Nacional Archives and Records Administration and El Public Record Office", Facultat de Biblioteconomia i Documentació, Universitat de Barcelona, (Actas de las V Jornadas de Archivos Electrónicos. Priego de Córdoba: Archivo Municipal, 2002)

10. Thibodeau, Kenneth. (2001). "Building the archives of the future: advances in preserving electronic records at the National Archives and Records Administration". En: D-Lib Magazine, 2001, February, vol. 7 num. 2.

11. United Nations Educational, Scientific and Cultural Organisation (UNESCO) (Prepared by National Library of Australia). (2003). "Guidelines for the preservation of digital heritage". Available at: http://unesdoc.unesco.org/images/0013/001300/130071e.pdf

Mikhail LARIN
Moscow, Russian Federation
Director
All-Russian Research Institute of Archival Science
and Records Management

Problems of Archival Storage of Digital Records for the Good of Citizen, Society, State

The development of computer technologies and communication information networks opened up opportunities for the exchange of paperless, i.e. digital, recorded information. A new type of documents appeared – electronic records, which must be preserved both in the process of record management and in archives.

A certain portion of electronic records (e-records) is for long-term (permanent) preservation. Thus, records management and archives need a solid legal and methodological framework. The one we have now is inadequate. The methodology for studying this issue should build on conceptual documents, among which *The E-Government Concept of the Russian Federation* and *The Strategy of the Information Society Development in the Russian Federation* should be singled out. These conceptual documents define as priorities of national policy the following:

- Creation of an integrated infrastructure capable of ensuring legally valid electronic interaction
- Development of a protected inter-agency records circulation system
- Establishment of a necessary legal and regulatory framework for an e-government

This means that the share of electronic records in the bulk of national documentary resources will be growing steadily. This adoption of Federal Law (No. 210-FZ of July 27, 2010) 'On Organizing the Delivery of Public and Municipal Services', which is explicit about the government's position on the introduction of the so-called electronic government.

Alongside with this one another law was adopted, Federal Law No. N 227-FZ 'On Amending Certain Legislative Acts of the Russian Federation in Connection with the Adoption of Federal Law 'On Organizing the Delivery

of Public and Municipal Services'". This law has considerably extended the scope of electronic records use. 26 laws all in all were amended.

Decree of the Government of the Russian Federation No. 477 of June 15, 2009, stipulates the rules of records management for federal executive bodies. These rules establish a unified records management procedure for federal executive bodies. In other words, the decree sets unified records management principles and rules for all federal executive bodies.

These rules define electronic record as equivalent to a hard-copy record on paper, provided that the established order is observed in the process of its preparation. Thus, for the first time in the history of Russian records management and archival science, the electronic record has acquired a formal legal status.

With a view of putting in place a system of inter-agency electronic records circulation, on September 22, 2009, the Government of the Russian Federation adopted the Regulation on the System of Interdepartmental Electronic Record Circulation.

This Regulation provides a legal framework for an integrated infrastructure that ensures legally valid electronic interaction between the central government bodies and local authorities.

Decree of the Government of the Russian Federation 'On the Unified System of Information and Reference Support for Citizens and Organizations on Interaction with Executive Bodies and Local Authorities by Means of Internet Information and Telecommunication Network' was adopted on June 15, 2009. On September 8, 2010, the Government of the Russian Federation adopted Decree No. 697 'On the Unified System of Inter-agency Information Interaction'; under the Decree, the Ministry of Communications will ensure introduction of a unified interagency interaction system. The implementation of the above decrees will help substantially increase the electronic records exchange within the "citizen-state" system.

Government Decree No. 176-p of February 12, 2011, approved an action plan for transferring the documentation process within federal executive bodies in their internal procedures to the paperless records circulation.

International agreements and standards on information exchange, including ISO 15489 'Information and Documentation - Records Management' provisions, should be taken into account as well.

Thus, we can see a line of statutory acts on a wider use of electronic records in management practice gradually taking shape. Among these instruments should be first mentioned: Law No. 149-FZ 'On Information, Information Technologies and Information Protection' of July 27, 2006; Law No. 153-FZ 'On Personal Data' of July 27, 2006; Federal Law 'On Archiving in the Russian Federation' (2004); and Federal Law No. 63-FZ 'On Electronic Signature' of April 6, 2011.

However, legislative and other legal acts adopted in the Russian Federation only provide general regulation electronic records use in management activities. The sporadic nature of statutory acts on electronic records hinders their use in all the areas of the life of society and state and prevents us from taking advantage of electronic records over hard-copy documents.

Let us look into Federal Law 'On Archiving in the Russian Federation'. Article 5 of this Law entitled 'Structure of the Archive Fund of the Russian Federation' specifies that the archive collection comprises archive documents irrespective of media type, including electronic records.

The Law contains neither a definition of electronic record nor distinct legal norms for their processing. This means that recognizing the fact of the existence of electronic records and their equality to traditional documents, Russia's law on archiving does not contain guidance to archivists in their work with e-records. A similar situation can be observed in legislation in general.

Even the new law on electronic signature does not provide answers to all questions about the practical use of this signature, leaving those to bylaws which are still under development.

There are other problems which hamper the use of electronic records. Many of those are caused by unsettled theoretical issues, among which are:

- Terminology and conceptual framework of electronic records
- Specific features of electronic records
- Authenticity of electronic records
- Validation of electronic records
- Studying the experience of work of various organizations with electronic records
- Development and legal consolidation of requirements and guidelines for electronic records processing

We believe that the main obstacle on the way to universal effective introduction of electronic records is the terminological ambiguity of the notion "electronic record". Unfortunately, this term has not been fixed in any of the adopted statutory acts.

In our opinion, as an electronic record should be understood:

- An electronic record proper, with its lifespan entirely in electronic environment
- Electronic (digital) copies of documents on paper and other media
- Databases in the form of a complicated record (registers, lists, inventories, catalogues, etc.)
- Web-records

Today, the first two groups of electronic records attract greater interest. Since we are talking of the practical usage of electronic records, it is only thinkable today to replace traditional hard-copy documents with their electronic analogues. These records should be included in relevant file registers and subsequently passed over to archives. Decree of the Government of the Russian Federation No. 176-p of February 12, 2011, approved an action plan for transferring federal executive bodies in their internal work to paperless records circulation, which is being implemented.

According to the plan, the Rosarkhiv archival depository and the All-Russian Research Institute of Archival Science and Records Management have developed guidelines for federal executive bodies on making lists of records which are to be created, stored and used as electronic records. These guidelines are available on Rosarkhiv's web-site. More than 50 per cent of federal executive bodies have already made up such lists.

We believe that the methodological principle of splitting organizations' records arrays into three types, in accordance with their archival value, should be applied here. Array 1 should include records for permanent preservation; Array 2 should include records for long-term preservation, and Array 3 should include mass production documents with the retention period of up to 5 years. These three types of documents constitute the major (no less than 60%) portion of every organization's records. These very documents should be first and foremost considered for replacement with electronic records. This would minimize the risk of losing valuable information.

Within the framework of the 'Electronic Russia' Federal Target Program, the All-Russian Research Institute of Archival Science and Records Management has developed a number of projects:

- guidelines on handling electronic records in records management processes and archives of public institutions;
- guidelines to public archives on organizing permanent (long-term) preservation of electronic records.

These guidelines take into consideration and further develop the provisions of two basic documents: 'Rules of Preservation, Accession, Registration and Use of Documents of the Archive Fund of the Russian Federation and Other Archive Records of Public and Local Archives, Museums and Libraries, and Institutions of the Russian Academy of Sciences' adopted by Order No. 19 of the Ministry of Culture of the Russian Federation of January 18, 2007, and 'Basic Working Rules for Archives within Organizations» (2002).

Both sets of guidelines were to become primary handbooks of the sort for archives within organizations and public archives in their work with electronic records, in particular, in the processes of electronic records appraisal, accession, registration, preservation, and use.

They can be used as an additional tool when working with electronic records. However, they still remain projects and need a thorough follow-up revision.

The future of the emerging modern information resources, i.e. databases, electronic registries, web-records, etc., still remains uncertain. Their characteristics need a more careful examination in terms of their suitability, as archival objects, to the theory and practice of archival science.

Only a thorough scientific consideration will provide the basis for recommendations concerning the archival storage of these modern information resources.

The issue of electronic signature is largely disputed among archivists. Our legislation unfortunately fails to provide solutions to all the issues of practical application of the law. It is obvious that the accession of electronically signed records to state archives may complicate their work. Upon termination of the signature's validity certificate it would be necessary to confirm the authenticity of the respective e-record in other way that is not specified by current rules.

Therefore, in order to avoid charging archivists with additional responsibilities of confirming the validity of e-records electronic signatures, an instrument should be provided for removing the electronic signature from the record when the latter is passed over for permanent (long-term) preservation, and the collection-builder should be held responsible for the authenticity of electronic records passed over to state archives for permanent (long-term) preservation. That is the way it is done, for instance, in the Federal Republic of Germany.

Another problem up in the air is that of electronic records formats that should be used:

— for permanent preservation of archive electronic records;

— for keeping the backup set of archive electronic records;

— for building collections of the user of archive electronic records.

Today, it is recommended that electronic records be passed to state archives in the format specified by the archive, preferably in PDF-A following the recommendation of the International Council on Archives. However, the global experience knows other options. For example, it is recommended to consider TIF as an alternative. This task is clearly important and urgent for Rosarkhiv and for the Electronic Record Preservation Center established within the Russian State Archive of Scientific and Technical Documentation.

The existing rules remain unambiguous about the type determination of electronic records, correlation of registration and preservation items, the necessity of streamlining new registration forms for electronic records, and other scientific and methodological issues of archival science. To settle them, we must build on international experience and the practices of the most advanced archives in their digitalization efforts.

We hope the Archival Activity Computerization Programme adopted in 2011 will help address these issues.

Barbara J. FORD
Urbana-Champaign, USA
Member
U.S. National Commission for UNESCO
Director
Mortenson Center for International Library Programs
Distinguished Professor
University of Illinois at Urbana-Champaign

Education about Preservation of Digital Information

The UNESCO Charter on the Preservation of Digital Heritage from 2003 includes principles, responsibilities, partnerships and cooperation, and the role of UNESCO in this essential activity to guard against the loss of digital heritage. A series of actions including developing strategies and policies, selecting what should be preserved, protecting the digital heritage, and preserving cultural heritage are outlined. With issues of varied access to digital resources often called the digital divide, international cooperation to enable all countries to ensure creation, dissemination, preservation, and continued accessibility of digital heritage is essential.

There is a need for education about the preservation of digital information for the general public as well as for professionals who oversee and guide digital preservation efforts. This presentation will provide an overview of some efforts to provide education about digital preservation.

U.S. National Commission for UNESCO

In order to highlight the effect of digital information across UNESCO-associated sectors, the U.S. National Commission for UNESCO (http://www.state.gov/p/io/unesco/) has developed a project plan for the production of a series of videos. These videos will feature interviews with experts in the field of information to address questions like:

How will digitization affect concepts of education and learning?

How will digitization affect access to information?

How will digitization affect publishing and the media creation?

One of the UNESCO Commission's current interests relates to the effects (commercial, educational, cultural, scientific) that increased digitization

will have on books moving into the future. As a way to promote thinking on this topic and highlight the knowledge U.S. experts have in this area, we are engaged in the production of a short video project.

We have asked experts to respond to the following questions on video:

> *What do you think books will look like in 15 years?*
>
> *What do you think the biggest changes, surprises, challenges and opportunities associated with books, libraries, bookstores, and/or newsstands in 15 years?*

We can take the video from the experts in whatever format is easiest then edit it into a product we hope will be useful in educating the public. If the videos are well received, we will then think about other projects we can work on relating to this topic. Providing more education for the general public about the importance of digital information and how it may affect their work and lives is essential as more and more information is available only in digital format.

U.S. Library of Congress

As part of the *National Digital Information Infrastructure and Preservation Program*, authorized by the U.S. Congress in December 2000, the Library of Congress is helping develop a national strategy to collect, archive and preserve the growing amounts of digital content for current and future generations. The Library's commitment to providing digital access to materials of historical importance extends to content that exists only in digital form, including Web sites.

As part of this effort, the Library of Congress has a very useful and comprehensive website on digital preservation. It includes information useful to everyone such as personal archiving to preserve our digital memories. A web site (http://www.digitalpreservation.gov/you/) provides information on preserving digital photographs, audio, video, electronic mail, personal records and websites. This site provides some tips to help make personal "born digital" information last. The guidance is basic and is meant to be a place to get started on personal preservation efforts.

Through a web archiving program which is a collaboration between the Library of Congress and the Internet Archive, students including children and teenagers archive websites that represent their lives and interests. The students use a web-based web archiving service from the Internet Archive,

to capture sites and manage, describe and browse their collections. Through this activity they not only develop critical-thinking skills and learn group problem-solving, they also develop an awareness of the transitory nature of web content. The Library of Congress archives the sites the students collect, and those collections become primary sources of information for future researchers. The students' experience should help them to evaluate the authenticity and value of other primary sources. You can find more information about this activity at http://www.digitalpreservation.gov/videos/k-12_web_archiving/k-12_wa_program.html.

As part of the Library of Congress efforts, the Digital Preservation Outreach and Education initiative fosters national outreach and education to encourage individuals and organizations to actively preserve their digital content, building on a collaborative network of instructors, contributors, and institutional partners. This effort defines educational needs, reviews existing curricula, defines core principles, builds an instructor base, evaluates delivery options, and develops outreach materials. The efforts that are underway are described at http://www.digitalpreservation.gov/education/index.html.

The Library of Congress provides a calendar as a public service to help people access training in the practices of digital preservation. Providers have been asked to designate their educational offerings by level (beginner, intermediate, or advanced) and intended audience (executive, managerial, or practical). You can sort by date, course, format, location, and cost at http://www.digitalpreservation.gov/education/courses/index.html.

Here is one example of a training program that the Library of Congress has taken the lead in developing. A train-the-trainer workshop for digital preservation consists of six modules:

- Identify - what digital content do you have?
- Select - what portion of that content is your responsibility to preserve?
- Store - how should digital content be stored for the long term?
- Protect - what steps need to be taken to protect your digital content?
- Provide - how should digital content be made available?
- Manage - what provisions should be made for long term management?

The workshop provides attendees with not only the basic digital preservation curriculum but also tips and techniques for conducting successful workshops. The first workshop was held at the Library of Congress in September 2011. Twenty-four attendees from around the country were instructed by a group of digital preservation educators.

The National Digital Stewardship Alliance (NDSA) was launched in July 2010 as an initiative of the National Digital Information Infrastructure and Preservation Program (NDIIPP), which Congress established in 2000 and the Library of Congress administers. Information is at http://www.digitalpreservation.gov/ndsa/index.html. In creating NDIIPP, Congress charged the Library of Congress with building the capacity for public and private organizations across the United States to save digital content of current and future value to the nation.

The NDSA's initial members were drawn from the existing NDIIPP partnership network. The Library of Congress acts as the secretariat of NDSA and provides program support. Much of the work of the National Digital Stewardship Alliance takes place in five working groups focusing on the following areas of digital preservation: content; standards and practices; infrastructure; innovation; and outreach.

The Outreach Working Group of the NDSA is focused on building relationships with stakeholder communities and preparing and sharing digital preservation information resources. These activities include: identifying and assessing tools the NDSA community needs for communication; identifying and promoting key digital preservation information resources; and defining opportunities for outreach within the NDSA and to the community beyond.

International Federation of Library Associations and Institutions (IFLA)

IFLA focuses mostly on international standards and guidelines and working with national governments and intergovernmental organizations to preserve digital information. The annual IFLA conferences feature education and training on digital preservation and curation and related topics. Librarians and information professionals work together in IFLA to develop manifestos, guidelines, and standards that can be used for educational purposes around the world.

The IFLA Manifesto for Digital Libraries (http://www.ifla.org/en/publications/ifla-manifesto-for-digital-libraries) includes preservation. IFLA has provided a variety of educational sessions on the topic and worked closely with national libraries on preservation efforts. In the Manifesto, IFLA encourages national governments, intergovernmental organizations and sponsors to recognize the strategic importance of digital libraries and to actively support their development. Contributions to large-scale digitization programs serve to make cultural and scientific information resources more widely available, and advance national and international digital library initiatives that will be sustainable over time. The Manifesto states that interoperability and sustainability are essential to the vision of digital libraries able to communicate with each other. Digital libraries that conform to commonly agreed upon open standards and protocols improve world-wide knowledge dissemination and access.

The IFLA–CDNL Alliance for Digital Strategies (ICADS) is a joint alliance of IFLA and the Conference of Directors of National Libraries that was established in 2008. Safeguarding digital heritage is a major issue, especially for national libraries, because of their legal task for preserving the national heritage of a country. The focus of the ICADS alliance is strategic and state-of-the-art digital library developments at national libraries. Information is at http://www.ifla.org/en/about-the-ifla-cdnl-alliance-for-digital-strategies. The alliance provides the international library community with current information, documentation and links to a wide variety of information about innovative digital projects in which partners are involved using a web directory of projects hosted by the National Library of Australia.

The National Library of Australia's Preserving Access to Digital Information (PADI) initiative aims to provide mechanisms that will help to ensure that information in digital form is managed with appropriate consideration for preservation and future access. The PADI website is at http://www.nla.gov.au/padi/topics/712.html. Its objectives are: facilitate the development of strategies and guidelines for the preservation of access to digital information; develop and maintain a web site for information and promotion purposes; actively identify and promote relevant activities; and provide a forum for cross-sector cooperation on activities promoting the preservation of access to digital information.

The website is a gateway to digital preservation resources and has a discussion list for the exchange of ideas about digital preservation. An international advisory group provides guidance for this important initiative. Unfortunately not all countries are represented and issues of the digital divide are evident as many countries seem to not be part of this alliance.

American Library Association (ALA)

The American Library Association (http://www.ala.org) has been active in developing national practices and educating members about the issues relating to preservation of digital information. Librarians and other professionals working on digital preservation have opportunities to learn about the topics through formal education and workshops. The ALA has developed a definition of digital preservation. It states that digital preservation combines policies, strategies and action to ensure access to reformatted and born digital content regardless of the challenges of media failure and technological change.

A Digital Preservation Interest Group serves as a venue for discussing the preservation management of digital assets whether commercial, born-digital or converted e-resources. Information about the group's activities is at http://www.ala.org/ala/mgrps/divs/alcts/resources/preserv/defdigpres0408.cfm. A variety of learning opportunities are available including free webcasts on topics such as "Preserving your Personal Digital Memories".

ALA also sponsors Preservation Week each year. ALA encourages libraries and other institutions to use Preservation Week to connect communities through events, activities, and resources that highlight what can be done, individually and together, to preserve our personal and shared collections. Information about the week can be found at http://www.ala.org/ala/mgrps/divs/alcts/confevents/preswk/index.cfm. There is a toolkit to plan and promote events and preservation resources including videos, statistics, links, and preservation awareness stories. These are the kinds of resources we can all use to help educate the general public and ourselves about preservation issues.

Education for Digital Preservation Professionals

There are a variety of educational opportunities for professionals involved in the work of digital preservation. Library and other professional associations

offer courses and certificates. Schools of library and information science provide degrees, courses, continuing education and opportunities for those working in the area to connect and share their expertise.

Conservation Online – Resources for Conservation Professionals (http://cool.conservation-us.org) is supported by the Foundation of the American Institute for Conservation of Historic and Artistic Works. It provides a helpful source for course and meeting announcements, as well as a forum to exchange information and a place to ask questions about conservation materials and techniques, and specialized email lists. Information is available about digital imaging and electronic materials. There is a fairly comprehensive list of educational opportunities in museums, libraries, and archives conservation/preservation that includes many international listings. Conservation and preservation information for the general public is also available on the site.

There are a number of organizations that provide educational opportunities. These include graduate schools of library and information science which offer programs for digital preservation experts. The program at the Graduate School of Library and Information Science at the University of Illinois at Urbana-Champaign is one example of the kinds of educational programs available and information is at http://www.lis.illinois.edu/.

The University of Illinois has a long connection with digital materials. University alumnus Michael Hart was known for his invention of ebooks and as the founder of Project Gutenberg. More recently, the University was awarded a National Leadership grant from the U.S. Institute of Museum and Library Services. Preserving Virtual Worlds II: Methods for Evaluating and Preserving Significant Properties of Educational Games and Complex Interactive Environments is the topic of the grant and is an example of the kind of research that helps support teaching and learning at the graduate level in digital preservation.

The Graduate School of Library and Information Science offers a master degree with a specialization in data curation which includes courses on digital preservation and archiving, data standards, and policy. Courses on digital preservation are available to all graduate students in library and information science regardless of their specialization. The master degree is also available online to those who are not able to relocate to the campus.

A certificate of advanced studies in digital libraries offer opportunities to learn more about issues related to the topic for those who have already received the master degree. During their studies, students also have the opportunity to work with the community on a number of levels to educate and share information about digital preservation.

Future Opportunities

Information preservation is one of the priorities of the UNESCO Information for All Program and this conference is a step toward working out the many challenges of the preservation of digital information. Efforts are needed on many levels to ensure that everyone - from citizens to professionals - has the information they need to understand and support digital preservation efforts. National UNESCO commissions can assist in these efforts as can other community and professional organizations. Meetings like this one and the UNESCO global event planned for 2012 will help us move forward in our efforts to preserve digital information.

Natalia GENDINA
Kemerovo, Russian Federation
Director
Research Institute of Information Technologies in Social Sphere
Kemerovo State University of Culture and Arts

Information Literacy and Personal Information Culture: A Sine Qua Non for Solving the Problem of Digital Information Preservation. Dialectics of Humanitarian and Technocratic Approaches

Integrated nature of digital heritage preservation: humanitarian and technocratic aspects

Article 1 of the UNESCO Charter on the Preservation of Digital Heritage defines digital heritage as unique resources of human knowledge and expression (5). The complexity and magnitude of the task of preservation is determined by the diversity of digital materials which incorporate texts, databases, still and moving images, audio and graphic files, software and web pages. Article 10 of the above Charter reads that *"measures should be taken to urge hardware and software developers, creators, publishers, producers and distributors of digital materials as well as other private sector partners to cooperate with national libraries, archives, museums and other public heritage organizations in preserving the digital heritage"* (5). The Charter names the major factors that contribute to the loss of digital heritage: obsolescence of the hardware and software, which are required for accessing digital materials; uncertainties about the resources; responsibility and methods for maintenance and their preservation; lack of supportive legislation. At that, Article 3 of the Charter stresses that *"attitudinal change has fallen behind technological change.[...]The threat to the economic, social, intellectual and cultural potential of the heritage — the building blocks of the future — has not been fully grasped"* (5). Thus, along with legal, technical, technological and organizational aspects, a humanitarian aspect, which is linked to the people's mindset, is moving to the forefront. To recognize its role, we have to provide information training to the citizens.

Information literacy as a component of a systemic response to the digital heritage preservation challenge

With respect to the issue under consideration, information training of citizens becomes especially acute since special information knowledge and skills should be obtained not only by hardware and software developers, as well as producers and distributors of digital materials, but also by a broader range of people who can be potential creators and users of the electronic content. Getting more involved in information interaction, people start acting not only as passive information consumers but as producers of electronic information resources and services. If they do not master information literacy, the task of digital heritage preservation can hardly be successfully accomplished.

The term of information literacy has been adopted by the international community to denote a wide range of knowledge, competences and skills connected with the ability of an individual to use information and communication technologies (ICT) in order to orient oneself in vast information flows, and to find, evaluate and use this information to resolve various everyday life problems. The leading international organizations, which have initiated the promotion of information literacy, are UNESCO and IFLA. The concept of information literacy has been shaped thanks to the efforts these organizations. The following steps, taken by IFLA and UNESCO, helped promote the information literacy objectives:

- Establishment of the IFLA Information Literacy Section (68th IFLA Council and General Conference in Glasgow, 2002),
- Creation of the IFLA - UNESCO strategic alliance to implement the decisions of the World Summit on the Information Society, including the solution of the information literacy issues (UNESCO Open Forum at the 72nd IFLA World Library and Information Congress in Seoul, 2006),
- Publication of *Guidelines on Information Literacy for Lifelong Learning* (2006) edited by Jesus Lau, Chair of the IFLA Information Literacy Section (9),
- Publication of the book *Towards Information Literacy Indicators* (2008) issued within the UNESCO Information for All Programme (12),

- Publication under IFLA's aegis of the book *Information Literacy: International Perspectives*, 2008 (10), which was translated into Russian and published on the initiative of the Russian Committee of the UNESCO Information for All Programme (2),
- Meeting of the international group of experts on the development of the media and information literacy indicators, held by the UNESCO Communication and Information Sector, UNESCO Institute for Statistics and UNESCO Bangkok Office on November 4-6, 2010, Bangkok (Thailand),
- Publication of the *Media and Information Literacy Curriculum for Teachers*, 2011 (11).

A special role in the promotion of the ideas of information literacy has been assigned to the UNESCO Information for All Programme. The concept of the Programme rests upon the idea of the world in which every individual can access personally significant information, and has the opportunity and skills to use this information for building a better society. The Information for All Programme raises the awareness of the importance of information literacy for all and supports the projects which contribute to the development of information literacy skills.

In the 2008-2013 Strategic Plan of the UNESCO Information for All Programme, information literacy is listed among the most important priorities along with information for development, information preservation, information ethics and information accessibility (4). The interpretation of information literacy, as given in this document, is based on the special proclamation adopted on November 9, 2005, at the Alexandria High Level Colloquium on Information Literacy and Lifelong Learning - *Beacons of the Information Society: The Alexandria Proclamation on Information Literacy and Lifelong Learning*. As is noted in this declaration, information literacy serves as a basis for life-long learning; it allows people in all walks of life to seek, evaluate, use and create information effectively to achieve their personal, social, occupational and educational goals. It is a basic human right in a digital world and promotes social inclusion of all nations. (6)

According to The Alexandria Proclamation, information literacy:

- Comprises the competencies to recognize information needs and to locate, evaluate, apply and create information within cultural and social context;

- Is crucial to the competitive advantage of individuals, enterprises (especially small and medium enterprises), regions and nations
- Provides the key to effective access, use and creation of content to support economic development, education, health and human services, and all other aspects of contemporary societies, and thereby provides the vital foundation for fulfilling the goals of the Millennium Declaration and the World Summit on the Information Society; and
- Extends beyond current technologies to encompass learning, critical thinking and interpretative skills across professional boundaries and empowers individuals and communities.

This interpretation reveals the role and importance of information literacy as one of the most important means of responding to technical, technological and organizational challenges of digital heritage preservation. At the same time, it should be stressed that the fulfillment of complicated ideological tasks within the digital heritage preservation problem requires development of personal information culture.

Personal information culture as an integrative notion

The notions of "work with information" or "personal information training" do not have a clear definition. They include a wide range of knowledge and skills which an individual needs to use when working with information and technologies that make this work easier. As noted above, following the IFLA and UNESCO's terminology, information training of an individual can be conducted through developing information literacy. The efforts of IFLA and UNESCO resulted in the development of the information literacy concept. Russian researchers believe that it will be possible to overcome the fragmentary nature of knowledge in the field of human/information (traditional and electronic) interaction and to reach comprehensive and complete understanding of the specific character of this interaction by developing a new systematic approach to information training of citizens based on a deeper and more general notion of personal information culture.

This concept is being implemented in Russia. It implies that mass enhancement of the information culture of society is possible only on the basis of special training of young citizens, i.e. on the basis of information training. The concept recognizes as equally important special information

training and adequate level of information culture, on the one hand, and computers and communication channels as indispensable attributes of the information society, on the other.

The concept offers the following definition of the term "personal information culture": this is a component of the personal general culture; a combination of the information style of thinking and a system of knowledge and skills which ensures purposeful and unrestricted activity aimed at fully meeting personal information needs through applying both traditional and new information technologies. Personal information culture is a major factor in successful professional and everyday activities, a factor of personal social protection in the information society (1, pp. 58-59).

Information style of thinking occupies a special place within the term of personal information culture. Its essence is in the value-based (meaningful, responsible) attitude to information, information products created and used, technical means and information technologies. Information style of thinking is a system of personal views on the universe of information and the place of an individual in this world; a system incorporating values, opinions, ideals, and principles of cognition and activity.

The connection between the style of thinking and personal practice is a very significant prerequisite for changing mere awareness into belief. It is impossible to instill information style of thinking or to knock it into somebody's head; it is necessary to create the necessary conditions so that an individual could arrive at a certain decision himself/herself. Information style of thinking is inseparably connected with motivation of information trainees to learn, for example, information literacy. It is motivation that determines the success of information training of citizens.

Information culture and information literacy: similarities and differences

The comparison of the two key notions allows us to reveal similarities and differences in the two approaches: the international approach (the concept of information literacy) and the Russian one (the concept of personal information culture). Both approaches focus on the development of the ability of every individual to receive, evaluate and use information which may be presented in any form and through various technical means (technologies). At the same time, the use of the term of information literacy in Russia is somewhat limited. These limitations are of a psycholinguistic

nature: in the Russian language, the word "literacy" means the ability to read and write, and indicates the lowest, elementary level of education. That is why, the term of information literacy, which denotes a complex phenomenon of human being/information interaction, is perceived to be somewhat elementary and primitive.

The notion of personal information culture is broader than information literacy. In addition to the skills of receiving, evaluating and using information, it embraces motivation and information style of thinking. It is inseparable from the sphere of culture and aims at removing in the information society confrontation of the two opposite cultures - technocratic and humanitarian. In general, the approaches to information training used in the Russian science and practice, are quite compatible with the international ones. The differences between the ideas of developing personal information culture studied in Russia and the international concept of information literacy are not fundamental. They merely reflect the intention of the Russian researchers and practitioners to combine the advantages of the international theory and practice with the national cultural and learning traditions, as well as the experience accumulated by Russian libraries and educational institutions (7-8).

Training course "Fundamentals of Personal Information Culture" as a means of organizing a comprehensive and coherent information training of citizens

We believe that purposeful and systematic information training of citizens will be practical if we introduce a special training course entitled "Fundamentals of Personal Information Culture". This course is designed to facilitate the lives of information consumers in the conditions of the modern "information bang"; to teach them how to raise the efficiency of information retrieval, analysis and synthesis; to equip them with the method of "information self-service". In (1) you will find a detailed description of the concept of personal information culture development that has been worked out by the Research Institute of Information Technologies of Kemerovo State University of Culture and Arts, as well as a set of curricula designed for virtually all groups of young people (1st-11th grade secondary school students, college and university students; persons seeking a scientific degree, including postgraduates).

A generic model of this course includes the following sections that reflect the content of information training:

> Section I. Information resources of the society and information culture
>
> Section II. Main types of information retrieval tasks and the algorithms of their solutions
>
> Section III. Analytic and synthetic processing of information sources
>
> Section IV. Information product preparation procedure

These sections are a compulsory and fixed part of the course which allows to solve the following tasks: give the vision of the world entering the information society; build the students' awareness of the complexity and diversity of the existing information resources; equip the students with the algorithms of search for, analytic and synthetic processing, retrieval and evaluation of information, transformation of the information retrieved and development of new data on its basis; teach the users how to document the results of independent learning, research and vocational activities.

The other part of the course, changing depending on the students' category, takes into consideration their age, type of activity (studies or work), profile and training record, professional profile, level of information culture, information needs, and some other factors.

Activities of the Russian Committee of the UNESCO Information for All Programme aimed at promoting the ideas of information literacy and information culture in Russia

The UNESCO IFAP Russian Committee coordinates the activities of Russian libraries, information and educational institutions, research teams and individual researchers engaged in information training of citizens, promotion of information literacy, which is a major priority of this Programme, by developing national traditions of personal information culture. The main spheres of its activities are organization, awareness raising and publishing, and education (3). The position of Chairman of the Committee is held by Evgeny Kuzmin, Candidate of pedagogic sciences,

Chairman of the Intergovernmental Council for the UNESCO Information for All Programme, Member of the Commission of the Russian Federation for UNESCO, President of the Interregional Library Cooperation Center, Honored Worker of Culture of the Russian Federation, and Winner of the Russian Federation State Prize.

The organization includes the initiative to create and provide a decade-long support to the Research Institute of Information Technologies in the Social Sphere at Kemerovo State University of Culture and Arts. With the assistance of the Committee, our Institute has turned into a research center that develops a theoretical basis for personal information culture in libraries and educational institutions, as well as carries out studies and provides training to information literacy coaches in Russian libraries, schools, colleges and universities.

Working within the framework of the UNESCO Information for All Programme, our Institute is applying a comprehensive approach to the development of personal information culture. The main areas of the Institute's activities are:

- Development of a theoretical basis for personal information culture with due regard of the Russian specifics,
- Launch of a wide-scale experimental initiative at Russian schools, colleges, universities and libraries,
- Promotion of the best world and national practices in the field of information training of citizens,
- Training of coaches (librarians and teachers) in information literacy and the basics of personal information culture,
- Training of various groups of users (school and university students, specialists, etc.) in information literacy and basics of personal information culture.

The results of the Institute's activities are published at the Institute's website: http://nii.kemguki.ru.

The Organizational activities of the UNESCO IFAP Russian Committee also include conferences, workshops, round tables, and discussions on information literacy and information culture. Among the most significant

conferences within which the Russian Committee organized the events dedicated specifically to information literacy and information culture are the following:

- International Conference *UNESCO Information for All Programme: Development of the National and International Information Policy* (Petropavlovsk-Kamchatsky, 2003);
- International Conference *UNESCO Information for All* Programme: *Universal Access to Information* (St. Petersburg, 2004);
- International Conference *UNESCO Between Two Phases of the World Summit on the Information Society* (St. Petersburg, 2005);
- International Conference *Development of Personal Information Culture: Challenges of the Global Information Society* (Moscow, 2006 r.);
- *The First Consultation Meeting of the UNESCO Information for All Programme National Committees* (Moscow region, Atlas Park-Hotel, December 7–8, 2009);
- *The All-Russian Scientific and Practical Workshop Development of Media and Information Literacy: an Imperative of the 21st Century* held within the All-Russian Library Congress: the 16th Annual Conference of the Russian Library Association (Tyumen, 2011).

It has become a tradition for the IFAP Russian Committee to hold special events dedicated to information literacy and information culture within the annual Crimea International Conference on *Libraries and Information Resources in the Modern World of Science, Culture, Education and Business* (Sudak, Autonomous Republic of Crimea, Ukraine).

The Russian Committee is actively engaged in publishing. In collaboration with its partners, it collects, organizes and examines relevant publications and documents published in Russia and by UNESCO, the UN and other international organizations on the humanitarian aspects of both global and Russian information societies. Part of these materials, which are most closely related to the topics of the Information for All Programme, are translated into Russian and published both in printed and electronic formats and are later distributed for free among Russia's federal and regional libraries, educational institutions, scientific and information centers. First of all, Russian translation is made for the materials which may play the

most significant role in shaping the scientific vision of the humanitarian issues of the global information society, and contribute to the development of a scientifically substantiated policy of Russia's information society.

All publications are issued in special series, including *Information Literacy and Information Culture* series. The following publications came out in 2002-2010 in the above series:

- *Information Literacy: International Perspectives* (2010);
- *ICT Competency Standards for Teachers: Policy Framework* (2009);
- *ICT Competency Standards for Teachers: Competency Standards Modules* (2009);
- *ICT Competency Standards for Teachers: Implementation Guidelines* (2009);
- *Quality Principles for Cultural Websites. Handbook* (2007);
- *Guidelines for Library-based Literacy Programmes* (2006).

During a decade-long UNESCO IFAP implementation process in Russia, the Russian Committee has been engaged in a large-scale outreach activities in the form of conferences, workshops and round tables, information and education initiative entitled "The Days of the Information for All Programme in Kuzbass", media publications. The best way to obtain data on information literacy issues is via the web-site http://www.ifapcom.ru/.

The need for large-scale awareness-raising activities is especially highlighted in Article 4 of the UNESCO Charter on the Preservation of Digital Heritage: *"Awareness-raising and advocacy is urgent, alerting policy-makers and sensitizing the general public to both the potential of the digital media and the practicalities of preservation"* (5).

Raising information literacy of the population to preserve digital heritage

In conclusion, we would like to note that at present information training of citizens is viewed as particularly topical and socially significant activity due to the development and consolidation of the information society and the global nature of ICTs. The analysis and understanding of the information training experience allow us to ascertain that it is unrealistic to fulfill this task only through occasional, non-systemic work of educational

institutions and libraries, the work which has been performed on a voluntary basis exclusively. In our opinion, the paramount prerequisites for raising information literacy of the population in order to preserve digital heritage are the following:

- Training within the system of educational institutions and libraries, irrespective of their type and kind;
- Training of personnel capable of professional teaching of information literacy to various groups of students;
- Creation and use of the provided information training materials and means: information resources, information publications and guides on electronic resources, conventional and electronic library catalogues, computers, facilities to access remote Russian and foreign information resources;
- Elaboration of teacher- and student-oriented educational materials on information literacy with a specific focus on digital heritage preservation.

References

1. N.I. Gendina, N.I. Kolkova, G.A. Starodubova, Yu.V. Ulenko - Development of Personal Information Culture: The Theoretical Substantiation and Modeling of the Training Subject. – Moscow, Interregional Library Cooperation Center, 2006; p. 512.
2. Information Literacy: International Perspectives/ edited by Jesus Lau; translated into English by E.V. Malyavskaya; scientific editor of the translation, N.I. Gendina – Moscow: Interregional Library Cooperation Center, 2010. – p. 237.
3. UNESCO Information for All Programme (2000-2010). To the 10th Anniversary of the Russian Committee of the UNESCO Information for All Programme: Collected papers/ Compiled by E.I. Kuzmin, A.V. Parshakova, T.A. Murovana, S.D. Bakeikin. – Moscow: Interregional Library Cooperation Center, 2010.- p. 208.
4. UNESCO Information for All Programme Strategic Plan (2008-2013)/ Russian Committee of the UNESCO Information for All Programme. – Moscow, 2009. – p. 47.

5. Charter on the Preservation of Digital Heritage // Cultural and Linguistic Diversity in the Information Society. – St-Petersbourg, 2004. – pp. 89 – 96.

6. Beacons of the Information Society: The Alexandria Proclamation on Information Literacy and Lifelong Learning [Electronic resources]. – URL: http://www.ifla.org/en/publications/beacons-of-the-information-society-the-alexandria-proclamation-on-information-literacy.

7. Gendina, N. I. Information Culture in the Information Society: the View from Russia [Text] / N.I. Gendina // UNESCO between Two Phases of the World Summit on the Information Society: proceedings of the international conference held in Saint Petersburg, Russian Federation, May 17-19, 2005.-Moscow, 2005. - pp. 97-105).

8. Gendina, N. Information Literacy or Information Culture: Separation for Unity: Russian Research Results [Electronic resources] / Gendina N. // World Library and Information Congress: 70th IFLA General Conference and Council, August 22-27 2004, Buenos Aires, Argentina: Conference Papers. CD-ROM

9. Guidelines on Information Literacy for Lifelong Learning. – Access mode: http://www.ifla.org/VII/s42/pub/IL-Guidelines2006.pdf.

10. Information Literacy: International Perspectives / Edited by Jesus Lau – Munich: K.G. Saur, 2008. – p. 160.

11. Media and Information Literacy Curriculum for Teachers (prepared by UNESCO) [Electronic resource] / Wilson, Carolyn; Grizzle, Alton; Tuazon, Ramon; Akyempong, Kwame; Cheung, Chi Kim. - Paris, UNESCO, 2011. - p. 192, illus. - URL: http://unesdoc.unesco.org/images/0019/001929/192971e.pdf).

12. UNESCO Information for All Programme. Towards Information Literacy Indicators. Conceptual framework paper prepared by Ralph Catts and Jesus Lau. Edited by the Information Society Division, Communication and Information Sector, UNESCO: Paris, 2008. – p. 44.

Plenary meeting. National Approaches, Solutions, Vision

Sabine SCHRIMPF
Berlin, Germany
Scientific Officer
Deutsche Nationalbibliothek

Digital Preservation Activities in Germany: German National Library and the Network of Expertise *NESTOR*

The library and information system in Germany and the system of memory institutions is decentralized, analogous to the political federal structure. Besides the Federal Archive and the National Library on federal level, each of the 16 states maintains state archives and state libraries. Each state government has the power to define its own collection mandate.

On federal level, the collection mandate of the German National Library was extended to cover online publications in 2006. It now includes text, image and sound-based works that are made available on public networks by authors, publishing houses or companies based in Germany. This can be, for example, e-books and e-journals, university dissertation and theses, music and audio files, websites of public interest and similar media works. Pure software, online data bases, radio and TV programmes are not collected.

The situation on state level is diverse. Some states cover online publications in their collection mandate, some cover only digital government publications and some haven't regulated collection mandates for digital publications at all.

Besides state libraries, several subject libraries have been confined with responsibility for special subject collections, in particular the Technische Informationsbibliothek (National Library for Science and Technology) in Hannover, the Deutsche Zentralbibliothek für Medizin (National Library of Medicine) in Cologne, and the Deutsche Zentralbibliothek für Wirtschaftswissenschaften (German National Library of Economics) in Hamburg and Kiel. Their collection missions naturally extend in the digital realm.

With this background, considerations concerning a digital preservation infrastructure for Germany can only be based on a distributed model. The National Library with its national responsibility takes a coordinating role thereby that has lately been confirmed in a report of the "Kommission Zukunft der Informationsinfrastruktur" (KII - "The future of the information infrastructure") commissioned by the Joint Science Conference (GWK - Gemeinsame Wissenschaftskonferenz) of the German federal and state government.[27]

German National Library

The German National Library has had a changeful history. In 1912, the German Booksellers Association founded it as "Deutsche B cherei" in Leipzig as a depot for all German publishing products (German and foreign-language literature published in Germany and German-language literature published abroad). After World War II and the division of Germany, the Deutsche B cherei remained in Leipzig in the Eastern part of Germany. A second site was opened in Frankfurt for the Western zones. The two sites were merged again to one institution when Germany united in 1990. Since 2006, the library officially carries the name "Deutsche Nationalbibliothek" ("German National Library"), when the "Law regarding the German National Library" came into force. It extended the collection brief of the library to include online publications. Its official task is now to collect, permanently archive, bibliographically classify and make available to the general public regardless of the format all German and German-language publications from 1913, foreign publications about Germany, translations of German works, and the works of German-speaking emigrants published abroad between 1933 and 1945.

The library had begun to collect digital publications on a voluntary basis already in 1998, starting with online university dissertations and theses. With the legal mandate for online publications, the collection was in a first step extended to e-journals, e-books and other static online publications. Due to the complex nature of dynamic online content, workflows for the collection of websites are still under development in 2011.

Because of the seemingly infinite wealth of online resources and the finite resources of the library, cooperation is a key factor in this area. The library actively seeks dialogue with content providers, for example represented

[27] http://www.leibniz-gemeinschaft.de/?nid=infrastr&nidap=&print=0

by the German Booksellers Association, the Marketing and Publishing Services of the German Book Trade (MVB), or the German University Libraries. It also partners with other German memory institutions that are involved with digital collections and digital preservation, and it has been engaged in several national initiatives striving for solutions in the digital preservation challenge, for example *NESTOR*, the German network of expertise in digital preservation, the *KOPAL* project (Cooperative Development of a Long-Term Digital Information Archive), *DP4Lib* (Digital Preservation for Libraries) and *LuKII* (*LOCKSS* and *KOPAL* Infrastructure Interoperability).

NESTOR

NESTOR is the German network of expertise in digital preservation, originally established in 2003 and led by the German National Library. It was inspired by the idea of a cooperative infrastructure for collecting and preserving digital information:

> "A cooperative structure for digital preservation, corresponding to the structure of the analogue realm, ought to be developed, which
>
> - ensures preservation and availability of all digital resources published in Germany, in German language or about Germany,
> - ensures preservation and availability of the most important objects in all scientific fields, not matter, if it is text, facts, images, or multimedia,
> - ensures preservation and availability of digital archival records".[28]

NESTOR unites stakeholders of the communities that are most affected by the digital preservation challenge: archives, libraries, museums, universities, research institutes, and multimedia centres. The network offers to its partners, but also to other interested stakeholders, a platform to discuss and align roles and responsibilities in the digital realm. *NESTOR* received funding from the Federal Ministry for Education and Research from 2003 until 2009. Since the termination of the project funding in 2009, the partners

[28] U. Schwens, H. Liegmann: Langzeitarchivierung digitaler Ressourcen, (2004). The paragraph quoted is originally in German.

(12, at this time)[29] are associated via a cooperation agreement and carry the network from their own resources. Institutionally and personally, the *NESTOR* partners bring together leading expertise in digital preservation. The German National Library, the Lower Saxony State and University Library and the Bavarian State Library have developed digital long term archives for their digital publications. The three large national subject libraries (Science & Technology, Medicine, Economics), unified in the Goportis – Leibniz Library Network for Research Information, are jointly implementing a digital long term archive. The Baden-Württemberg State Archive has implemented a long term preservation system in the archival sector. Other partners bring in digital preservation experiences from the museums perspective and the universities and research and development perspective.

To fulfill its mission to foster knowledge exchange, *NESTOR* hosts several working groups, publishes materials, and organizes workshops and events. Among others, it hosts working groups on multimedia preservation, copyright and related legal issues, preservation planning, and repository audit and certification. Not only *NESTOR* members can participate in the working groups. Quite the contrary, they are open to anyone interested in Germany. Altogether, around 80 institutions have been engaged in *NESTOR* working groups.

NESTOR also runs a cooperation with the German Institute for Standardisation (DIN), which is the German equivalent to the International Organization for Standardization ISO, to help crystallize standards in the relatively new field of digital preservation. A standard for audit and certification of trustworthy digital archives is almost finalized and publication as DIN 31644 is expected in autumn 2011. DIN standards for the transfer of information objects into digital long-term archives as well as for trustworthy Persistent Identifier systems are under development.

Another main activity of *NESTOR* is training and qualification. Together with several higher education partners from the German-speaking area, *NESTOR* develops initial and further training courses in the field of digital preservation. Meanwhile, the topic has been included in the teaching

[29] Bavarian State Library, German National Library, Hagen Open University, Lower Saxony State and University Library Göttingen, Humboldt University in Berlin, Baden-Württemberg State Archive, Prussian Cultural Heritage Foundation / SMB - Institute for Museum Research, Baden-Württemberg Library Services Centre, German Language Institute, Computer Games Museum, Goportis - Leibniz Library Network for Research Information, PDF/A Competence Center

curricula of 11 universities. Additionally, *NESTOR* hosts annual summer schools and seminars and workshops in irregular intervals. Such, it contributes to the condition that qualified staff is available to deal with the digital preservation challenge.

KOPAL

The *KOPAL* project, which ran from 2004 until 2007, and its successor *DP4Lib* (2009 until today) are important building blocks in the development of a co-operative infrastructure for digital preservation. In 2004, the German National Library partnered with the Lower Saxony State- and University Library (SUB), IBM and the IT centre of the G ttingen University (GWDG) to develop and test a long term preservation system for co-operative use. The kopal system is based on IBM DIAS, at that time a standard-setting implementation of the OAIS reference model. Because the DIAS system was designed as an in-house long term archive for the KB, it was extended in the *KOPAL* project to meet the needs of several independent institutional users. In principle, each institutional user can register its individual secure storage space, in which it has its data holdings under its own administrative control.

The open source "KOPAL Library for Retrieval and Ingest" (*koLibRI*), which was also developed in the project, connects individual institutional users with the archiving system. *koLibRI* can be configured according to the needs of each individual *KOPAL* user. Such, it allows users with various different selection profiles and with different types of digital object to share one and the same archiving system.

koLibRI validates the objects' file formats and packages the objects together with technical metadata in Submission Information Packages (SIPs). The SIPs are imported, and, in compliance with the OAIS Reference Model, stored as Archival Information Packages (AIPs) in the DIAS archival storage unit. Each *KOPAL* user can, via koLibRI, address and retrieve exclusively its own data. Technology watch and preservation planning can in principle be implemented as a joint service for all *KOPAL* users. As preservation action, migration was tested within the *KOPAL* project. After the end of the project, the *KOPAL* archiving system was maintained by the German National Library and the Lower Saxony State- and University Library and taken forward in the DP4Lib project.

DP4Lib

In the *DP4Lib* project (Digital Preservation for libraries), which is funded by the German Research Foundation, DNB and SUB open up the *KOPAL* system to several more users and extend its functionality. The overall goal is to establish and run a ready-to-operate service for long-term preservation. Besides DNB und SUB, six potential *KOPAL* users joined the extended project consortium.[30]

The partners bring varying use scenarios to the project. The German Institute for International Pedagogical Research, for example, is a research institute with large specialized holdings, digitized and born digital journals as well as databases. The Library Service Centre of Baden Württemberg offers long-term preservation as a service to its customers and seeks a safe harbor for the data for which it has assumed responsibility.

Based on their individual needs, the *DP4Lib* partners conjointly compiled a catalogue of requirements for long-term preservation as a service. As one consequence (to give just one example) the *KOPAL* principle that each customer administers its own secure storage space, was modified. The partners did not insist on their own, logically isolated, storage spaces. Instead, they favored a shared storage space, with only an administrative separation between the data holdings. This solution simplifies the creation of SIPs, AIPs, and DIPs, and it advantages joint preservation planning.

A second, important activity of the *DP4Lib* project is the development of business and a cost model for co-operative long-term preservation operations. Today, one knows the cost of hardware and software and can such estimate storage costs, but the cost of preservation planning and action is not so easy to quantify.

LuKII

In *LuKII* (LOCKSS and *KOPAL* Infrastructure Interoperability), the German National Library co-operates with the Humboldt University Berlin to test interoperability between the *LOCKSS* concept (Lots of Copies Keep Stuff Safe) and the *KOPAL* archiving system. The project is funded by the German Research Foundation.

[30] Baden-Württemberg Library Services Centre (BSZ), Leibniz Institute for Educational Research and Educational Information (DIPF), GBV – the Library Consortium of Nothern Germany, Saxon State and University Library in Dresden (SLUB), Technical Information Library (TIB), Thuringian University and State Library Jena (ThULB)

LOCKSS is a relatively cost-effective distributed preservation approach, originally developed at Stanford University in the United States.[31] The *LOCKSS* partners are connected via a network of computer servers (so called "*LOCKSS* boxes") and each partner stores their digital holdings in seven copies in the network. The strength of *LOCKSS* is bit stream preservation. Redundant storage (the said seven copies) in the network insures the digital materials against damage, corruption, or loss. Preservation planning and preservation action are not supported within the network. Each partner has to monitor its own holdings and, if needed, initiate preservation actions outside the *LOCKSS* infrastructure.

In the *LuKII* project, DNB and Humboldt University establish a *LOCKSS* network for distributed storage in Germany, including infrastructure to provide ongoing technical support and management for *LOCKSS*. Moreover, the project partners investigate interoperability between *LOCKSS* and *KOPAL* in order to combine cost-effective bitstream preservation with the well-developed preservation tools of koLibRI and *KOPAL*. In a last stage, the interoperability prototype is tested by archiving data from German institutional repositories.

To establish the German LOCKSS network, seven partner libraries[32] implemented and connected their LOCKSS boxes. The network is designed as Private LOCKSS Network (PLN), which can only be accessed by the partners, but not by the general public. A competence center at Humboldt University in Berlin offers technical assistance to the project partners and to other interested parties in the German-speaking area.

In order to enable interoperability between the two archiving systems, modifications have to be made to both systems, which have different structures for their Archival Information Packages and their exchange interfaces. Two different scenarios are conceivable. In one, the *KOPAL* information packages are ingested into the LOCKSS network. The other option is to ingest LOCKSS information packages into the *KOPAL* system.

To test the modified software, data from German open access repositories will be harvested and ingested into the LOCKSS network two times. In

[31] http://www.lockss.org/lockss/Home

[32] University Regensburg, Bavarian State Library, Berlin State Library, the Researh Centre in Jülich, Lower Saxony State and University Library Göttingen, North Rhine-Westphalian Library Service Centre and the Saxon State and University Library in Dresden (SLUB)

a first step, the unmodified LOCKSS software will be used. Towards the end of the project (beginning of 2012), data will be harvested from the same repositories using the new, interoperable LOCKSS software so that the results can be compared.

Summary and outlook

Besides national initiatives, the German National Library is involved in international collaborations such as the International Internet Preservation Consortium (IIPC), which promotes global exchange and international relations between institutions engaged in Web Archiving, or in the Unified Digital Format Registry (UDFR), which aims at building a single shared formats registry that can be jointly used for preservation planning purposes. It is common to the national and international collaborations of the DNB that they prepare and foster a distributed and co-operative approach to digital preservation. Thereby, the realities of local requirements and decentralized developments on the one hand and the need of leadership, central co-ordination and standard setting on the other hand have to be balanced carefully.

New topics that arise on the agenda are the long term preservation and re-use of digital research data and the distribution of responsibilities between archives, libraries, and specialized research data centres. The German Research Foundation and the "Kommission Zukunft der Informationsinfrastruktur" mentioned in the very beginning of this article have for the first time addressed it on a truly national level.

Another aspect that becomes increasingly clear and urgent is the integration of digital preservation considerations in the lifecycle of digital objects (no matter if they are digital publications or research data). Well described object with context information in open and standardized formats can more easily be preserved than objects created in proprietary formats with poor documentation. In this regard, much awareness-raising remains to be done.

Ramesh GAUR
New Delhi, India
Librarian & Head
Kala Nidhi Division
Indira Gandhi National Centre for the Arts

Development of the Digital Repository of Indian Cultural Heritage: A collaborate Project under the National Digital Preservation programme of India

The Indira Gandhi National Centre for the Arts (IGNCA) in New Delhi, India, established in 1987, was designed as an autonomous center encompassing the study and experience of all the arts. Kala Nidhi, a national information system and databank, consists of a reference library of print collections, a large microfilm/microfiche library, a collection of slides, and photographs covering many disciplines related to Greater India, South Asia, Southeast Asia, and West Asia. A cultural archives consisting of rare collections, a conservation laboratory, and a multimedia unit also form part of the organization. The basic mission of Kala Nidhi is to support the research of the different divisions of the center as well as researchers and scholars from other academic institutions. This paper describes the Indian cultural heritage resources at the IGNCA and provides some details about the digitization initiatives underway.

Introduction[33]

The scope of the arts covered by the Indira Gandhi National Centre for the Arts (IGNCA) (Figure 1) encompasses a wide area of studies, such as creative and critical literature, both written and oral; the visual arts, ranging from architecture, sculpture, painting, and graphics to general material culture; photography and film; the performing arts, including music, dance, and theater in their broadest connotation; festivals and fairs; and everything in life that has an artistic dimension. It is within the Centre's agenda to explore, study, and revive the dialogue in areas pertaining to the arts between India and her neighbors, especially in the south and Southeast

[33] Much of the information in this section is drawn from the IGNCA brochure, published in 2002

Asia. The IGNCA manifests its academic research work in its publications, international and national seminars, conferences, exhibitions, and lecture series.

Figure 1: Website of the Indira Gandhi National Centre for the Arts, http://www.ignca.gov.in

To fulfill the objectives and its principal aims, the IGNCA's functions are organized into five divisions that are autonomous in structure but interlinked in programming. The Kala Nidhi Division comprises a reference library which includes printed books, slides, microfilm, photographs, audio-visual material, a conservation laboratory, a multimedia unit, and the cultural archives. The Kala Kosha Division undertakes fundamental research and investigates the intellectual multidisciplinary traditions. As a research and publication division, it endeavors to place the arts within the integral framework of a cultural system, combining the textual with the oral, the visual with the aural, and theory with practice. The Janapada Sampada Division complements the programs of Kala Kosha. Its focus shifts from the text to the context of the rich variegated heritage of the rural and the small scale societies. Its activities focus on the lifestyle study programs comprising the Loka Parampara, which revolves around a community, and the Ksetra Sampada, which revolves round a region. It has developed a core collection of material and documentation of folk arts and crafts including those of tribes; several multimedia presentations; and multi-disciplinary lifestyle studies of tribal communities for evolving

alternative models for the study of the Indian cultural phenomenon in its totality and the interlinking of environmental, ecological, agricultural, socio-economic, cultural, and political parameters. The Kaladarsana Division provides a forum for interdisciplinary seminars, exhibitions, and performances on unified themes and concepts. The Sutradhara Division provides administrative, managerial, and organizational support and services to all the other divisions.

Kala Nidhi: A Cultural Knowledge Resource Center[34]

Kala Nidhi serves as a major cultural information/knowledge resource hub to support research objectives and the technical information needs of other units of IGNCA. The Kala Nidhi holds an extensive and unique collection of over 600,000 resources, reflecting the multi-dimensional and interdisciplinary nature of the arts. As a major repository of reference material relating to the humanities and the arts, it comprises an outstanding reference library and equally rich cultural archives with access to multimedia databases and information systems. It also houses a well organized audio-visual and photo-documentation facility and a conservation laboratory.

Kala Nidhi Reference Library

The reference library (Figure 2) has an array of primary and secondary material in the broad areas of humanities and the arts. These include about 150,000 books in Indian and foreign languages encompassing archaeology, philosophy, religion, and ritual studies; history and anthropology; art and literature; folk, pastoral and community studies; volumes on tribal culture and languages; encyclopedias, dictionaries, and bibliographies; biographies and autobiographies; monographs and atlases; catalogs of unpublished manuscripts of Indic and Asian origin; and translations and conference proceedings, among others. The library also subscribes to many general and specialized journals relating to the arts and houses back volumes of several periodicals.

[34] Much of the information in this section is drawn from the Kala Nidhi brochure, IGNGA, published in 1996

Figure 2: View of the Kala Nidhi Reference Library. Photograph by Ramesh C. Gaur, IGNCA.

The library is fully automated, providing easy access to its resources through its online catalog. Some of the important collections of the reference library are listed below:

Rare Books. Acquisition of rare books, some over 200 years old, is a special function of the library. The collection numbers over 3,500 volumes. Some of these books are illustrated with sketches, drawings, lithographs, etchings, and wood cuts, providing clues on the art, architecture, and culture of the bygone era, which has paled with time. A fascinating aspect is the collection of travel literature dating back to eighteenth and nineteenth centuries, offering firsthand accounts of Europe's understanding of the regions of India in terms of their commercial, colonial, and recreational interests. The contributors to the publications were chiefly the members of British army and navy, surveyors, engineers, medical staff, and travelers. Much of what they wrote was intended for amusement for their readers, as well as for information and instruction. The rich intellectual content of some of these books reveals to the readers the condition of the society when the country was passing through a transition phase as a result of cross- cultural currents—experiencing new influences from the West, adapting to new conditions, and at the same time retaining the age old values, traditions, and customs. Many of these books from eighteenth to the twentieth centuries contain fascinating illustrations.

Area Collections. Research and exchange programs with Southeast Asia, East Asia and Eurasia have brought in cultural material covering a wide range of disciplines and enriching the core collections on the regions both in print and microforms. The microfiche collection pertaining to these areas numbers around 150,000 items. The books and other publications have

been acquired from various countries in above areas through purchases, gifts, and exchanges. A collection of over 20,000 publications in microfiche received from INION (Institute for Scientific Information in Social Sciences, Moscow) in the Russian language are points of attraction for the scholars in Central Asia and Eurasia.

Personal Collections. Acquisition of several personal/rare collections has further enriched the library. These relate to important scholars who have made significant contributions in the field of arts and related studies. These collections include:

- **Suniti Kumar Chatterjee** (1890-1977)**:** The collection consists of a monumental over 30,000 volumes ranging from books, journals, pamphlets, and reports in several languages, including English, Bengali, Greek, Russian, Avesta, Chinese, Japanese, Arabic, Hindi, and Sanskrit. They cover various disciplines such as linguistics, history, art, literature, and archaeology.
- **Acharya Hazari Prasad Dwivedi** (1907-1979)**:** Dwivedi was a well-read man with an encyclopedic mind, a doyen among teachers and a well-known Hindi scholar. His collection contains over 13,000 volumes in the areas of Indian literature, religion, philosophy, history, and science.
- **Heeramaneck:** Son of an art dealer, he himself was a dealer, collector, connoisseur, entrepreneur, and benefactor. The collection consists of about 2,500 books, including books on Asian arts, donated by his wife Alice Heeramaneck.
- **Lance Dane:** Dane's collection of about 5,000 books covers Indian art and architecture. A renowned photographer, his collection includes a number of books on numismatics.
- **Ananda Kentish Coomaraswamy:** The personal collection of the great art historian and man of letters includes 696 folders with correspondence, journals, reprints, and newspaper cuttings; 217 works of art comprising drawings from the Punjab hills, illustrated manuscripts, Indian miniatures from Malwa and Punjab hills, and modern paintings. In addition, there are forty-eight small Indian or South Asian sculptures and decorative objects, forty-eight books, 1,097 journals, 115 letters, 703 music records, 486 glass slides, and 227 photographs.

- **Kapila Vatsyayan:** The founding Member Secretary of the IGNCA, Dr. Vatsyayan is one of the most highly respected scholars in the field of Indian art and culture. Her collection contains about 11,000 books, over 3,000 periodicals, and a large number of seminar reports and papers, photographs, CDs, and other non-book materials such as textiles, ceramics, and slides.

Manuscripts in Microfilm and Microfiche

It is estimated that India possesses over 5 million manuscripts, the largest number in the world, which require preservation for knowledge and reference. Reprographic compilation of valuable manuscripts in Indian and foreign collections from private and public libraries is a unique feature of the IGNCA. The goal is to collate primary sources of the Indian tradition lying scattered, fragmented, inaccessible, or worse, in danger of extinction. So far, over 272,000 manuscripts have been microfilmed in 20,600 microfilm rolls (Figure 3). Scholars and researchers can access this microfilm/microfiche collection and also obtain copies, subject to copyright restrictions. Around 13 million folios of unpublished Sanskrit, Pali, Persian, and Arabic manuscripts are presently available for research and reference. Reprographic material of various primary and secondary texts have been obtained from foreign institutions like the Biblioth que Nationale (Paris), Cambridge University Library (Cambridge, England), Staatsbibliothek (Berlin), INION (Russia), the Wellcome Library for the History of Medicine (London), and the Oriental and India Office Collection, British Library (London). The manuscripts in micrographic records have been acquired from Khuda Baksh Oriental Public Library (Patna), Bhandarkar Oriental Research Institute (Pune), Asiatic Society (Calcutta), Manipur State Kala Academy (Imphal), Atombapu Research Centre (Imphal), Vaidika Samsodhana Mandala (Pune), Saraswati Bhawan Library (Varanasi), Government Oriental MSS Library (Chennai), Oriental Research Institute and MSS Library (Trivandrum), Shri Ram Verma Government Sanskrit College (Tripunithura), Thanjavur Maharaja Serfoji's Saraswati Mahal Library (Thanjavur), and Shri Ranbir Sanskrit Research Institute (Jammu).

Figure 3: Samples of the manuscript collections microfilmed by IGNCA.

Slide Collection and Other Visual Resources

Kala Nidhi holds the largest collection of slides on Indian arts, paintings, sculpture, architecture, illustrated manuscripts, and the performing arts. It is the only center in India equipped with the proper infrastructure for archival storage, computerization of data, duplication, and digitization. Over the years it has acquired and generated about 100,000 carefully selected slides from seventeen centers in India and fifteen abroad. On average about 3,000 slides are added to the collection every year. The notable acquisitions are from the Victoria and Albert Museum (London) and the Chester Betty Collection through the courtesy of INTACH (Charles Wallace bequest). In addition to this, the American Association of South Asian Art has donated a complete set of 8,000 slides to the library. Kala Nidhi also has a collection of more than 1,700 photo-negatives. It has photo-documented the artifacts on display in the Himalayan Heritage Museum (Jammu), Sheeshmahal Museum (Patiala), Qila Androon (Patiala), and the festivals in Kulu. The photo-documented material comprises 699 photographs and 653 slides. Facilities like slide-viewing cabinets, slide viewers, and slide projectors are available under expert guidance. The use and dissemination of most of these are subject to copyright.

The Cultural Archives

The Cultural Archives collects, preserves, and classifies the personal collections of scholars/artists in the originals, in reproductions, and/or in reprographic forms for the purposes of research and dissemination. It is further enriched by personal and ethnographic collections, documentation, cultural exchange, and research in area studies. Some rare collections of ethnography and audio-visual documentation of old masters and rare art forms have been acquired by the archive (Figure 4). The archive also maintains films on research projects taken up by the various divisions of the IGNCA. Some of these films such as *Yelhou Tagoi* by Shri Aribm Shyam Sharma and *Wangala of the Garos* by Shri Bappa Ray have won National Film Awards. The acquisitions have been classified in six categories: *Sahitya* (literature), *Vastu Silpa* (architecture and sculpture), *Chaya Pata* (photographs), *Sangita* (music), *Nrtya* (dance), and *Natya* (theater).

Figure 4: Exhibition space in the Kala Nidhi. Photographs by Ramesh C. Gaur.IGNCA

The literature section consists of forty-one audiocassettes of Dr. R.C. Rangra's interviews of eminent writers in Hindi and other regional languages; twenty audio spools of the voice of Gurudev Rabindra Nath Tagore reciting poetry/songs; and twenty-one audio spools of the Akhilesh Mittal collection of recordings of the renowned Urdu poet Firaq Gorakhpuri. The architecture and sculpture section comprises the Lance Dane collection of over 1,000 sculptures and figures in various media; photographs of the unique terracotta temple architecture of Bengal by Shambhunatha Mitra; and photo-documentation of the Ajanta Caves in 675 color slides by Benoy K. Behl. The photographs section has prestigious collections consisting of 2,700 glass-plate negatives, 2,700 contact prints, and 200 original prints of Raja Deen Dayal, the grand master of Indian photography; Henri Cartier-

Bresson's collection of 107 black-and-white photographs of India comprising candid shots of the Indian freedom struggle; the D. R. D. Wadia collection composed of photographs of political leaders, scientists, diplomats, dancers, and landscapes; and the David Ulrich collection on nature and rock art. The music section comprises S. Krishnaswami's rare compilation of reprographs on the musical instruments of India including research notes made over a period of forty years; musicologist Ranganayaki Ayyangar's collection of V. A .K. Rangarao's and S. Natarajan's music records (78 rpm) of Carnatic music concerts; and Western classical music such as Beethoven and Mozart.

Miscellaneous Collections

This section includes 955 oil, water, and charcoal paintings (1,751 color slides) of the renowned mother-daughter Hungarian painters Elizabeth Sass Brunner and Elizabeth Brunner. Most of these works were created on their journey from Hungary to India via Italy and their sojourn in India, particularly at Santiniketan, Kumaon, Gwalior, and Rajasthan. The collection also includes UNESCO posters containing forty-four color laminated photographs of the important sites of the different parts of the world; two sets of Coorg wedding costumes and forty-three pieces of jewelry; delicate and exquisite *Surahis* of unbaked clay made by the master craftsman Abdul Majid Ansari; and reprints of old (fifteenth century onwards) geographical maps of Europe, Africa, and Asia donated by Professor R.P. Mishra. There are notable additions from other divisions of the IGNCA such as the Harikatha collection composed of 212 books on religion and philosophy (Sanskrit, Tamil, and English) and nine audio spools; the Sadagopan collection containing silent films and audio tapes, photographs, slides, negatives, albums, and notes on the sermons and other religious functions of Jagad Guru Kamakoti Sankaracharya; and puppets from seventeen countries as well as documentation of puppet shows from China, Tibet, and India, among others.

Conservation Facilities

The Kala Nidhi houses a well-equipped conservation laboratory with trained staff (Figure 5) providing conservation services in-house as well as for other institutions in the country and abroad. The unit carries out a regular survey of art objects and provides treatment on a priority basis. Simultaneously, it takes preventive measures against the aggressors that may cause damage to the artifacts. Materials of a delicate nature such as palm leaves, parchment, miniature paintings, textiles, and oil paintings as well as harder substances such as metal and its alloys, stones, terracotta, and ceramics are treated

according to the best practices of conservation. Major conservation projects undertaken so far include the Henry Cartier-Bresson photograph collection, the Yashodhan Mathpal rock art reproductions, the Sharda Lipi manuscripts, the Buddhist Canons, rare books in the reference library, folk paintings, scroll paintings of Santokhba, and wood and terracotta objects.

Figure 5: Conservation staff at work in the Kala Nidhi conservation laboratory. Photographs by Ramesh C. Gaur, IGNCA.

The ABIA Project[35]

ABIA stands for the *Annual Bibliography of Indian Archaeology* that was published by the Kern Institute in Leiden during 1926-73. The *ABIA* is a unique and ambitious scheme of annotated bibliographic compilations on Indian arts and archeology, which used to be released annually by a group of Dutch scholars. The *ABIA* was the brain child of Professor Jean Philippe Vogel, one of the most notable Indologists of the twentieth century.

Though suspended for about a decade in the late 1980s and early 1990s, this pioneering effort was revived in 1995 mainly through the efforts of scholars from the International Institute for Asian Studies in the Netherlands. The Dutch again found willing and enthusiastic partners in Sri Lanka, Thailand, and Indonesia in 1995. In 1996, the International Institute for Asian Studies in Leiden put forward a proposal to resume the bibliography. The new bibliography is called *ABIA South and Southeast Asian Art and Archaeology Index*, and briefly the *ABIA Index*.

The *ABIA Index* supplies annotated and indexed entries on scholarly publications in Asian and European languages relating to prehistory, (proto)historical archaeology, art history (including modern art), material culture, epigraphy, paleography, numismatics, and sigillography. The project

[35] Ramesh C. Gaur, "ABIA South and Southeast Asian Art and Archaeology Index Project at IGNCA, New Delhi," in Challenges for South Asian Resources and Information services: Essays in Honour of Dr. Ravindra N. Sharma, ed. Rajwant Singh Chilana (New Delhi: Concept Pub. Co., 2008

receives scientific support from UNESCO. The database *ABIA South and Southeast Asian Art and Archaeology Index* is fully searchable online and is freely accessible at http://www.abia.net. Extracts from the database are also available in the form of printed bibliographies. Three volumes have been published thus far between 1996 and 2011.

Publication of bibliographies and the development of bibliographic databases are some of the basic objectives of the IGNCA. The IGNCA participates in the *ABIA* project to fulfill the above objectives as well as to strengthen international cooperation among arts and archeology institutions in South and Southeast Asia.

Indian participation in the *ABIA* project started with Professor S. Settar in approximately 1998. In December 2001, the Indira Gandhi National Centre for the Arts was able to associate itself with the *ABIA* project and assisted with getting the required entries on Indian publications. A memorandum of understanding was signed between the IGNCA and the Postgraduate Institute of Archaeology, University of Kelaniya, Sri Lanka, in October 2002 to continue the work. In January 2007, the IGNCA became the coordinating office for *ABIA* project for next five years. The plan is to establish a regional network of individual scholars and institutions in India to gather information on various publications in the areas of interest to the project. Apart from the contributions of records to *ABIA*, this project will also help to strengthen the IGNCA's bibliographic activities. The IGNCA will be planning the publication of the fourth volume of *ABIA* within the next two years of this project.

Database Development and Computerization

The IGNCA has been designated by the government of India as the nodal agency for most matters relating to arts, humanities, and cultural heritage, and it has been charged with providing computerized storage, retrieval, and dissemination of information on all aspects of arts and cultural heritage. Over the years, the IGNCA has developed several multimedia databases and information systems to preserve the vast cultural heritage in various forms and make it accessible for research and dissemination. Special efforts have been made to network these programs with various institutions inside and outside the country. These systems include LMIS (Library Management and Information System) which provides cataloging information on all books and periodicals; CATCAT (Catalogue of Catalogues) which provides information on more than 1,000 catalogs of published and unpublished manuscripts; MANUS which contains descriptive information on about 3,000 manuscripts; PICTO which

includes information on two-dimensional and three-dimensional art objects; SOUND (sound recordings) comprising information on Vedic chanting of the Ranayaniya and Jaiminiya Sakhas of Samaveda and the Paippalada Sakha of Atharvaveda; KK Terms (Kala Kosa Terms) comprising terms for the Kalatattvakosa project which helps scholars in their preparation of comprehensive text references for each term, verification of bibliographic references, and quotations and terms in different texts; BIBL (Bibliography) which provides information on more than 6,000 references (monographs, books, journals, and articles); and THES (Thesaurus) which is made up of key words in some tribal languages and dialects to identify cognate terms relating to the Five Elements[36]. This database has been developed to support the programs of the Janapada Sampada Division of the the IGNCA which aims to understand the arts in their eco-cultural and socio-economic contexts.

Cataloging

The entire collection has been cataloged using the MARC21 format. Some of the unique collections such as the slides, manuscripts, and digital images have been cataloged and are accessible to users over the IGNCA Intranet. The web catalog will be launched shortly. The software used for automation also has the capability of linking the electronic resources to the catalog. The display formats available in the online catalog includes both AACR and MARC. All of the printed publications in the Kala Nidhi Reference Library have been barcoded for the automated circulation system and inventory control.

Abstracting and indexing of Indian journals in the field of art and culture received by the reference library is under process; so far about 8,000 articles have been completed.

Digitization of Cultural Resources at IGNCA

The IGNCA possesses facilities for the digitization of materials, post-digitization editing, high capacity storage and backup systems, design and development of effective retrieval systems, and more. The technology is based on open standards using Unicode, a multilingual standard for fonts, accepted worldwide. Search interfaces are available in both English and Hindi (Devanagari). Users have the option to select material of interest either from a specific collection like books or manuscripts, or from the entire collection.

[36] According to Vastu Shastra, the traditional Hindu system of design, the Five Elements, or pancha maha bhoota, are the building blocks of the universe. In order, the elements are Earth, Fire, Sky, Water, and Air

Figure 6 shows the types of digital resources available at IGNCA.

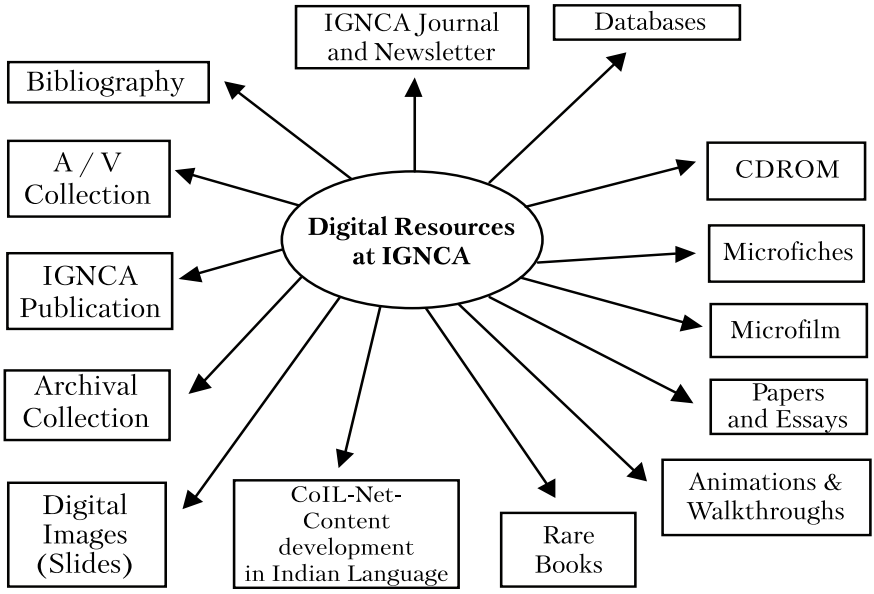


Figure 6: Digital collections at IGNCA.

Digitization standards at the IGNCA follow the UNESCO Guidelines published in 2002[37].

- Slides: Photo CD format (five resolutions normally and six resolutions in specific cases)
- Microforms (microfiche and microfilm): 300 dpi TIFF
- Photographs: 300 dpi TIFF (600/1200 dpi in special cases)
- Books and printed materials: 300 dpi TIFF
- Audio: 44Khz WAV file format
- Video: MPEG-1/2
- TIFF: Uncompressed lower quality derivatives are used for online access of the materials.

[37] Information Society Division, United Nations Educational, Scientific and Cultural Organization, Guidelines for the Preservation of Digital Heritage, http://unesdoc.unesco.org/images/0013/001300/130071e.pdf

The table in Figure 7 provides the details about the status of cataloging and digitization of various materials available at IGNCA. The IGNCA is not only digitizing its own collection but also helping other cultural intuitions in the digitizing of their collections, provided they agree to give a copy of the digitized materials to IGNCA for academic and research purposes. Under this scheme the IGNCA had helped institutions like the National Museum of India, the Archeological Survey of India, Vishva-Bharati, and Shantineketan in their digitization projects. Digitized collections at IGNCA belong to more than sixty institutions and over fifty eminent Indian scholars in the fields of art and culture.

Type of Material	Quantity	Digitization and Cataloging Status	Digitization Formats	Present Preservation and Access Methods	Future Plans
Rare Books	3,500	2,896 Digitized All Cataloged	TIFF, JPEG, and PDF	Online Catalog and DSpace	Development of Digital Repository
Books from ASI Library		All Digitized All Cataloged	TIFF, JPEG, and PDF	Online Access at www.ignca.nic.in	
Manuscripts in Microfilm / Microfiche	252,000	All Digitized All Cataloged	TIFF, JPEG	Online Catalog in MARC21 Format. Access through CD/ DVDs	Conversion to PDF-A, and Development of Digital Repository
Visuals	100,000	All digitized All Cataloged	PCD, JPEG	Online Catalog along with Access to Low Resolution Image	Development of Digital Repository
Photographs, Paintings, and Glass Plate Negatives	10,000	All Digitized Not Cataloged	JPEG	Access through CD/ DVDs	Development of Digital Repository

Books and Journals	170,000	All Cataloged		Online Catalog in MARC 21 Format.	Linking with Digital Repository
Documentation of Monuments of India Both Protected and Unprotected by ASI				Online access at www. ignca.nic.in	

Figure 7: Cataloging and Digitization Status of Materials in the IGNCA Collections.

Various Initiatives Undertaken for the Development of the Digital Repository at IGNCA[38]

The IGNCA's Kala Nidhi has put in place various digitization projects and has been actively curating various digital collections. The collection has grown to over ten million objects. As in the rest of the world, digital technologies have been enthusiastically adopted by India; the lifestyles and culture of Indians are being influenced by digital technologies. This has resulted in digital culture and digital cultural artifacts. IGNCA recognizes that digital platforms are becoming the media for human expression and communication of thought, and it is preparing to curate both digitized objects and born-digital objects. Some of the tasks in hand for setting up a digital repository at IGNCA include:

- Develop a framework and process to support the certification of digital repositories.
- Create tools to identify the attributes of digital materials that must be preserved.
- Design and develop systems for the unique, persistent identification of digital objects that expressly support long-term preservation.
- Capture information about the complex relationship between digital preservation and intellectual property rights.
- Determine which technical strategies best provide for continuing access to digital resources.

[38] Raju Buddharaju, Ramesh C. Gaur, and Vishal Slagotra, "IGNCA's Institutional Digital Repository" (position paper, January 25, 2011, unpublished

- Define the minimal-level metadata required to manage digital information for the long term.
- Develop tools to automatically generate and/or extract as much of the required metadata as possible.

At present the digitized collections are available on CDs and DVDs and also on the server. Users are given access by retrieving details from the online catalog. Some prototypes have been tested but as of now no proper archiving system is in place.

Kalasampada Digital Library: Resources of Indian Cultural Heritage (DL-RICH)

Recognizing the need to encompass and preserve the distributed fragments of Indian art and culture, and to serve as a major resource center for the arts, the IGNCA in collaboration with the Ministry of Communication and Information Technology, initiated a project named Kalasampda (Digital Library: Resources of Indian Cultural Heritage) for the development of a databank of cultural heritage. Kalasampada facilitates scholars' access to the materials including 272,000 of manuscripts, over 100,000 of slides, thousands of rare books, rare photographs, audio and video along with research publications produced by the IGNCA, from a single window. Multimedia computer technology has been used for the development of a software package that integrates numerous sources of cultural information into one place. This will provide a new dimension in the study of Indian art and culture while giving due importance to each medium. The system is a digital repository with a user-friendly interface, and this knowledge base will help scholars explore and visualize the information stored in multiple layers.

More specifically, the contents of the digital repository include over 5,000,000 folios of manuscripts, over 100,000 slides, 4,000 photographs, IGNCA published books, issues of Kalakalp (IGNCA's journal) and Vihangama (IGNCA's newsletter), over 400 hours of audio and video, and approximately fifty walkthroughs. The processes of digitization, post-digitization editing, and integration are continued to encompass all such materials available in the IGNCA. The user interface of the application is very simple, and many cultural institutions have approached IGNCA to copy it. The system is currently available only on the IGNCA Intranet because these materials are prized possessions and covered under intellectual

property rights and copyright law. However, partial information may be accessed from the IGNCA website with special permission.

DSpace[39], the open source software developed by MIT for digital archiving, has been installed at the Kala Nidhi Reference Library. It is proposed that the entire digital collection of Kala Nidhi may be uploaded into DSpace in the future.

CoIL-Net: Content Development in Indian Language Network

Content development in Indian Language Network is a project sponsored by Ministry of Communication and Information Technology (MCIT). The main objectives of the CoIL-Net project are:

- To enhance the access to cultural resources using digital technology
- To develop a reusable "Model Design" and "Development Process" for implementing a user-friendly web-enabled heritage library for the Hindi-speaking population in India and Abroad
- To implement a web-enabled Hindi-based multimedia heritage library which also offers contextual and vetted links to important websites to contribute towards the socio-economic development of the Hindi-speaking region

Conclusion

Digital archives benefit all researchers, scholarly institutions, and the entire research community by avoiding the duplication of effort, broadening the communication process, reducing the time to announce findings, expanding the audience, and above all preserving information assets for the use of future generations. The online catalog of about 500,000 cultural resources in MARC21 format, the installation of a CD mirror server, networking of various the divisions of the IGNCA, and the development of a digital repository consisting of photographs, manuscripts, slides and audio-video materials are just some of the plans being implemented at the IGNCA.

At present about 200 terabytes of digitized material need to be uploaded to a suitable digital archiving system. Several in-house systems as well as some open source tools have been tested without much success. However, under a recently-approved proposal C-DAC Pune (The Centre for Development of Advanced Computing in Pune) with the support of Department

[39] DSpace open source software, http://www.dspace.org

of Information Technology, Government of India will develop digital preservation tools for the IGNCA to create a digital repository of Indian cultural heritage. IGNCA expects its first software tool from C-DAC Pune for the conversion of TIFF files to the PDF format for the manuscripts collection. Presently all of the digital images have been linked with the catalog, and some work is in process to upload the digital rare books into in-house software.

Once all of the collections at the IGNCA are digitally archived, one of the world's largest digital libraries in the field of art and culture will be at the disposal of researchers working in India and abroad. The digital preservation of both digitized as well as born- digital collections is one of the biggest challenges of the twenty-first century.

Sinikka SIPILÄ
Hameenlinna, Finland
President-Elect
IFLA
Secretary General
Finnish Library Association

Esa-Pekka KESKITALO
Helsinki, Finland
Chief Analyst
National Library of Finland

Preserving Digital Information in Finland

This is a short overview on the preservation of digital information in Finland. Various Finnish organisations have met the challenges of digital preservation. On one hand these are the problems related to digitizing, on the other hand - those of securing preservation of the born digital materials.

Cooperation of libraries, archives and museums in document preservation

In 2007, the Ministry of Education and Culture in Finland established a taskforce to analyse the issues related to digital information preservation and use. This taskforce suggested that the library, archive and museum sectors should closely cooperate and develop a joint action plan.

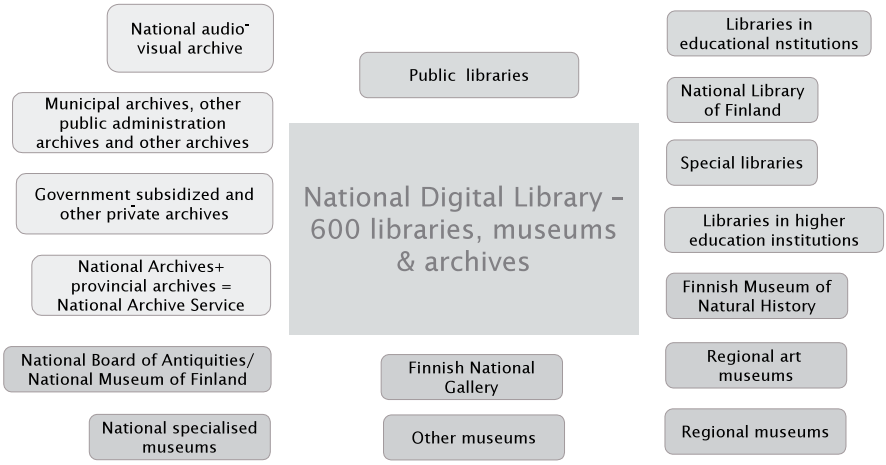
Our objective is to create a centralized or unified long-term preservation system. Under this project, the use and management of the materials will be shared among the organizations; however, only one organization will be responsible for secure preservation and related procedures. These activities should be carried out within a common long-term preservation information system.

National Digital Library Project

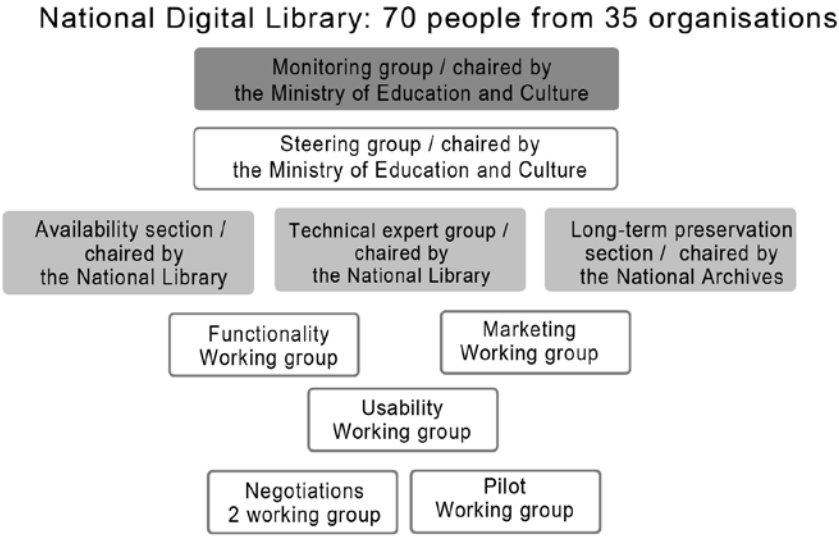
Long-term preservation cooperation has become part of the National Digital Library project.

Information can found at the following address: http://www.kdk.fi/en.

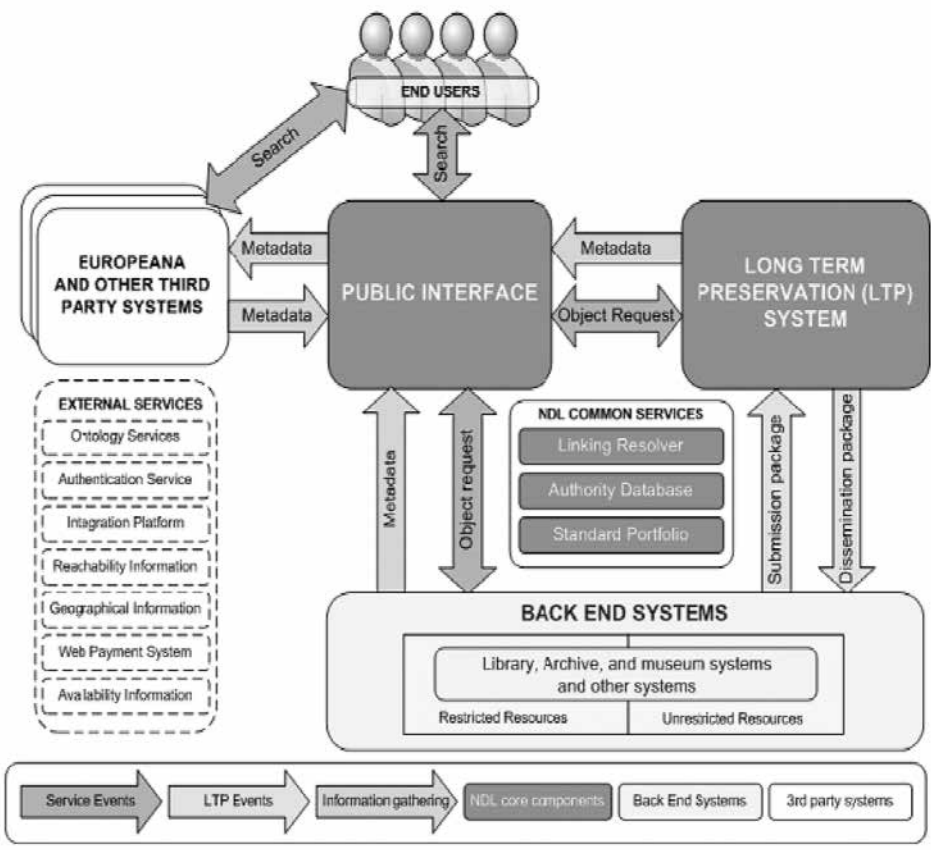
The following diagram gives the idea of how the National Digital Library Project works:



National Digital Library is managed in the following way:



The final report of the first implementation stage of the Long-Term Preservation System project is available in English at http://www.kdk.fi/en/long-term-preservation.



Detailed description of the National Digital Library's enterprise architecture is available at http://www.kdk.fi/en/enterprisearchitecture. The enterprise architecture of the National Digital Library Project:

Digital collections in Finland

Since 2008, the National Library has been receiving mandatory digitized copies of the most remarkable collections, including historic collections of newspapers and periodical, as well as radio and TV archives of the National Audiovisual Archive (since 2008). The amount of archived electronic documents has been increasing as well. Many archives and museums have digitized their historic collections.

Integration of materials

Bit-level data preservation still remains the main priority, which means that, so far, we lack adequate capacities to address procedures for ensuring integrity, such as conversion of large amounts of data to other formats.

Secure long-term preservation of large amounts of data is not an easy task. We believe it technically possible but quite challenging in terms of economic sustainability: it is necessary to understand how to analyse risks, implement the necessary preservation measures and create relevant information systems while staying within reasonable financial expenditure limits.

Long-Term Planning Continued

The long-term preservation system has already been carefully drafted; however, no funding has yet been obtained. Nevertheless, we believe this project to be very promising. Further details are available in the article written by Esa-Pekka Keskitalo, Chief Analyst of the National Library, at http://urn.fi/URN:NBN:fi-fe201107051838. The planning is to be continued.

National recommendations

Decision has been reached to take national recommendations for the National Digital Library as the basis for developing digital information and metadata formats, as well as assisting the activities of minor libraries, archives and museums involved in long-term preservation.

Hard additional work is also required to manage the research data, i.e. first hand information, for instance, measurement results.

For more information, please, contact Esa-Pekka Keskitalo at: esa-pekka.keskitalo@helsinki.fi

Vyacheslau NOSEVICH
Minsk, Belarus
Director
Belarusian Research Center of Electronic Records

The Collection of Archival Records' Digital Copies: Belarusian Experience

Preservation of archival documents and provision of sustained access to them are the main tasks of archivists. Production of digital copies is one of the ways to accomplish these tasks. Digitization is actually the only possibility to build a serviceable archival collection while preserving original copies. In some cases (loss of the original, carrier deterioration, etc.) a digital collection can also act as a backup collection.

It is important not only to create a digital copy but to ensure its long-term preservation. There is always a risk of losing information, and this risk exacerbates due to the fact that archivists, as a rule, have no expertise in the field of information technologies. Therefore, organization of long-term digital preservation calls for serious scientific approach.

In the Republic of Belarus this problem is being solved step-by-step. The main role is assigned to a specialized archival research institution – Belarusian Research Center of Electronic Records (further referred to as "the Center"), established in 1998. In addition to carrying out applied research, the Center serves as a repository of digital copies, and its structure comprises an archive of electronic records.

In 2001, the Center developed the Concept of Document Digital Copying for the National Archival Collections that embraced the main aspects related to digital copying of paper documents in state archives, as well as their preservation and use. International experience was examined, along with a rather poor (and not always positive) experience in digitizing documents accumulated by that time by Belarusian archives. All archives got acquainted with the Concept conclusions, some of them even held discussions on this matter. Thus archivists familiarized themselves with the specific terminology and understood the range of problems related to the task they faced, which enabled them to prepare themselves psychologically for their solution.

At the next stage, since the mid 2000-s, digital copies started to be mass-produced to a certain extent, which brought about the need for organizational measures within the framework of the archival industry. Again, the Center played the pivotal role in the process by accomplishing in 2007 the study «Development and experimental validation of principles of acquisition, preservation, control, and provision of access to the Collection of digital archival copies». In the course of the study, general approaches and principles were defined; the experience of archival institutions in acquiring, preserving and using digital copies of archival documents was analyzed. As a result, the requirements for nine business-processes were formulated:

1. Selection of archival documents to be digitized;
2. Preparation of the selected documents to be digitized;
3. Transfer of archival documents to the Center's division or external agency to accomplish digital copying;
4. Digitization;
5. Identification of digital copies obtained;
6. Their post-processing;
7. Control;
8. Preservation;
9. Provision of access.

The most important requirements have to be characterized.

1. Several approaches **to the selection of archival documents** to be digitized are applicable:
   - Full-scale archival collection digitization;
   - Digitization of the intensively used segment of archives;
   - Building topical collections;
   - Digitization within the framework of cultural and/or educational projects.

Full-scale digitization of archival collections provides expanded access to them for an unlimited number of users and ensures their preservation due to the fact that digital copies of the documents are used while the original is spared. However, this approach should consider the feasibility of digitizing

archival collections that do not contain highly valuable or intensively used documents, and the related financial issues. It is necessary to carry out the monitoring of collections in terms of quality (considering documents value disparity) and intensity of their use.

Digitization of the intensively used segment is ideal for archives because a well-defined criterion is used to select documents that need to be digitized. The objective in this case is to make most relevant documents available for wider community.

Within this approach, much attention has to be paid to the following issues:

- Development of the criteria of collection/part of the collection's intensity of use taking into account experts' recommendations on increasing the efficiency of information use;
- Assessment of demand fluctuations for different branches of knowledge and different types of documents.

Building topical collections is the most applicable approach for the majority of archives. The selection criterion is unequivocally determined by the topic. The subject scope is selected either by the archives themselves or their control agencies depending on the high relevance of a certain topic; there is also a possibility of building digital collections on special order. In this case, external resources – other archives, museums, and libraries – are used more intensively, which enables to build an electronic collection that is more comprehensive than the one based on a single archive collection.

Within the framework of cultural and/or educational projects, selection criteria are determined by the project goals. As a rule, an integral multimedia database on a specified topic is built containing not only archival materials, but also images of physical monuments, printed materials, sound records, etc. This is illustrated by the CD information system "Nesvizh: World Heritage Site and National Monument" in the Belarusian and English languages (designed by the Center in 2005 upon request of the National Commission of the Republic of Belarus for UNESCO) that comprises many digital copies of documents of the Radziwill family collection.

Whatever the approach, the following materials are to be digitized as first priority:

- Highly valuable documents;
- Documents on worn or deteriorating carriers;

- Documents with low-contrast or fading texts;
- Intensively used documents.

Among the equally valuable collections, the collections comprising documents in poor physical condition are to be digitized first.

In any case, after the selection of documents is finished, it is necessary to draw a list of collections, files, individual documents to be approved by an expert commission of higher-level archival administration, or republican archives.

2. **Preparation of documents for digitization** is accomplished by the archives owning the documents in accordance with the document lending procedure. The physical condition of the documents is then checked, and documents that need to be restored, conserved and treated, in particular, to be unbound before digitization and bound again after it is completed, are identified. These documents are subject to supplementary expertise accomplished by digitization specialists who are to conclude on the necessity of restoration, conservation and treatment, or the possibility to digitize the document as it is.

Alongside with archival documents, the inventories are to be digitized as well, and after recognition, with certain skills and knowledge of office applications, their digital copies may be used as electronic reference and search tools.

3. **Transfer of archival documents for digitization** is executed through placing an order for digital copies' production comprising the list of documents to be digitized. The order, if applicable, stipulates patterns and procedure of digitization, image files formats, image files merge mode, image color array. Archives are also to adopt location scheme for digital copies on media (one carrier locates the digital copies of a single collection, or several collections). Upon the completion of work, archives are to verify its completeness and quality.

4. **Digitization** is accomplished by specialized archival institutions and technical units of the archives, if properly equipped. All six central archives in the Republic of Belarus have their own digitization equipment (book and flatbed scanners, digital photo cameras). The goal is set to establish digitization centers in each of the six Belarusian regions. Meanwhile, regional and zonal archives mainly command the services of the Republican Technical Back-up Document Microfilming Laboratory in Gomel, where over 70% of

all Belarus backup copies are produced. Recently, the Central Document Microfilming and Restoration Laboratory of the National Archives of the Republic of Belarus has joined the process of inventories' digitization.

Partly, archival documents are digitized with the equipment of the Center (over 5% copies produced). Meanwhile, the Center's main task is to assimilate new hi-tech scanning hardware, define efficient copying modes and develop user operating instructions. The equipment is further transferred to archives and laboratories. In this manner the National Historical Archives' microfilm scanner and book scanner, as well as Gomel Laboratory's pass-through scanner were put into operation.

After scanning text documents images of 300 dpi (up to 600 dpi in some cases) are mainly produced, while the basic formats are PDF, TIFF, and JPEG. Depending on the original copy, both full-color and monochrome images (grayscale and binary) are created. Belarusian State Archive of Films, Photographs and Sound Records produces digitized copies of films and sound tracks of 1411KbS quality in DVD and WAV formats. Depending on the size, the resolution of scanned photographic negatives makes 300 - 2,400 dpi.

5. **To ensure unambiguous identification**, every digitized file is assigned a unique multilevel ID-number. Identification scheme is part of the "Methodological recommendations to unify archival documents digitization and digital copies identification" (approved by the Department of Archives and Records Management of the Republic of Belarus on December 27, 2007, with amendments and additions adopted on February 6, 2009).

6. When appropriate, graphic files that have been produced are **post-processed.** Post-processing comprises image cropping, angling, stitching, contrast and brightness adjustments, color correction. In some cases, such post-processing significantly increases the quality of the copies and makes visible the elements that are hardly discernible in the original. Meanwhile, the experience proves that, at this stage, unification is hardly achievable. The images produced by different institutions and even different operators often differ significantly in their appearance and quality.

7. To **control digital copies** in archives, the following document formats are used:

- carrier records;
- digitization order log;

- copying order log;
- archive cards;
- folder inventory of digital copies.

Thus, archives control digital copies in accordance with the abovementioned formats by collections or folders.

Archive cards are produced for every collection whose folders are digitized, after the digitization order is executed. Later on, information about the following orders related to that collection is simply added to the card.

Document forms specifying control process have been introduced. These forms are given in the "Methodological recommendations on preparation and transfer of archival documents to be digitized, controlled and preserved" (approved by the Department of Archives and Records Management of the Republic of Belarus on November 25, 2008).

8. Information preservation is achieved through acquiring backup copies and storing them in different places. This principle is declared by regulatory "Provision on the Collection of Digital Copies" approved by the Department on June 22, 2007 (with amendments adopted on October 22, 2008).

The idea is that copies of digitized documents produced by archival institutions or specialized laboratories and recorded on removable storage devices are transferred to a single digital information storage center where two more copies are produced (working copy and backup copy) identical to the original one. They form the **Collection of Digital Copies** (CDC). The Center's Electronic Archive is in charge of the long-term preservation of working and backup copies. Then, the original copies are returned to the owner archives where they flow into **digital user collections**. With the copies being stored by the Center, owners have the possibility to restore digital documents in case of loss or damage.

This approach ensures centralized preservation of two digital copies. In this case archival institutions do not need to spend money on long-term preservation of large amounts of information, which requires expert knowledge and continuous monitoring of physical preservation, multiple copying, migration into new formats and new carriers.

The copies in archival digital collections can be stored on removable media, separate computers, servers, etc. Media conversion also has to be monitored

and documented in digital copy inventory where the route to the digital copy is indicated, up to the subfolder level.

Removable media with digital copies are stored in the archives in designated places: safe boxes, filing cabinets, shelves, etc. The place and access mode are defined in accordance with archive's administrative regulatory order.

The CDC copies are stored in the Center's electronic archive using special equipment (RAID arrays). The transfer to the Center is accomplished according to the schedule annually approved by the Archives and Records Management Department Director. Archives transfer digital copies together with their description files.

By early 2011, the number of digital copies transferred to the CDC exceeded 7,000 items (archive files), or over 1,4 million scanned pages. Besides, 44,000 moving and sound records, and photographic negatives were digitized. The total volume amounts to 755 Gbyte, and 1.5 Tbyte with provision for the second backup copies.

Up to one third of this volume (2,736 archive files, 465,000 pages) was digitized in 2010, with approximately the same figure planned for 2011. Given that the total number of files in Belarusian archives exceeds 13 million, this is a big scale project.

The Center monitors the quality of digital copies being transferred and sends them back for revision, if necessary. Due to the stringent control, 28,500 poor-quality copies were returned to archives in 2010 only. These copies were mainly created by central archives that accomplish scanning on their own equipment with their staff being not always duly qualified.

9. State archives – copy owners are able to provide **access to digital copies** and carry out use control with the help of the copies that make their user digital collections. Experience proves that archives provide access to digital copies mainly in reading rooms. The procedure for printing them out and issuing certified paper copies is the same as for photocopying. It is planned to establish in 2012-2015 within the framework of the national program "Belarusian Archives" an archival open access system to enable remote users to access archives' reference and search tools and to order copies via Internet. Also, the electronic archive of the Belarusian Research Center of Electronic Records is to be upgraded to ensure, among other tasks, preservation of the expanding Collection of Digital Copies. Time will show how far these plans will go.

Giovanni BERGAMIN
Florence, Italy
Head of the Information Systems
National Central Library of Florence

Maurizio MESSINA
Venice, Italy
Head of the IT services
Marciana Library in Venice

The National Digital Preservation Infrastructure for Legal Deposit of Italian Publications: Digital Stacks

This presentation is the result of the common work of Maurizio Messina, head of the IT services at the Marciana Library in Venice, and myself[40]. It outlines the *Digital Stacks (Magazzini Digitali)* project, whose aim is to set up the national long term digital preservation infrastructure for legal deposit of Italian publications – according to the provisions of the new Italian legal deposit law (L. 106/2004, DPR 252/2006).

The project was originally established in 2006 by the *Fondazione Rinascimento Digitale,* a private institution that supports the use of digital technologies in the cultural domain, and by the *Biblioteca Nazionale Centrale di Firenze*.

In the first part of the presentation the technical architecture and the metadata management problems will be outlined; the second part will take into account both organizational model and service model.

Digital Preservation (referred as DP in this presentation) cannot be considered just a technical process. Strategies to avoid bit loss or to prevent hardware and software dependencies are only a part of the issue. DP, and *Digital Stacks*, have also other problems to deal with: legal aspects (mainly copyright), economic implications (sustainability), cooperation between different institutions[41] (legal deposit institutions in the *Digital Stacks* case), because is widely accepted that DP is out of reach of a single institution.

[40] This presentation is mainly based on: G.Bergamin, M. Messina. Digital stacks: turning a current prototype into an operational service in Digital libraries, 6. Italian research conference, IRCDL, 2010, Padua, Italy, January 28-29, 2010. Revised selected papers / Maristella Agosti, Floriana Esposito, Costantino Thanos (eds.), pp.39-46.

[41] Brian Lavoie, Lorcan Dempsey, Thirteen ways of looking at ... digital preservation, <<D-lib magazine>> 10(2004), 7/8 http://www.dlib.org/dlib/july04/lavoie/07lavoie.html

There are many DP definitions, but for the purposes of the project, DP could be defined as a public service to be provided by trusted digital repositories in order to ensure - for deposited digital resources – viability (a bit sequence could be read over time without modifications), renderability (a bit sequence could be successfully interpreted by a computer according to a specific format – for instance PDF), authenticity (identity and integrity) and availability for designated communities[42].

As you see, the name of the project recalls the stacks of the legal deposit libraries. In most aspects digital stacks are comparable to conventional ones: digital resources must be preserved for the long term; digital stacks grow as new resources are added; modification and deletion is not an option; it is impossible to predict the usage frequency of stored digital resources; and it is likely that some resources will be seldom or never be used[43].

The aim of the project was to set up an infrastructure based on a *long term framework*. Component failures in an information system are the norm rather than the exception[44], so the infrastructure is based on:

- data replication (different machines located in different sites);
- simple and widespread hardware components, non vendor-dependent, that can easily be replaced (just simple personal computers: nowadays an ordinary personal computer could easily store up to 12 TB (4 hard disks of 3 TB capacity) using widespread and inexpensive SATA[45] technology).

[42] This definition is based on:
a)*Trustworthy Repositories Audit & Certification (TRAC)* http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf (for the concept of "trusted digital repositories");
b) Luciana Duranti, *Un quadro teorico per le politiche, le strategie e gli standards di conservazione digitale: la prospettiva concettuale di InterPARES,* <<Bibliotime>>, 9(2006), 1 http://didattica.spbo.unibo.it/bibliotime/num-ix-1/duranti.htm (to assess the *authenticity* of a digital resource, the *public service* must be able to establish its *identity* and demonstrate its *integrity*)
c) PREMIS 2.0, 2008, PREsevation Metadata: Implementation Strategies, http://www.loc.gov/standards/premis/ (for the concepts of "viability, renderability, understandability, authenticity, identity");
d) *OAIS. Reference model for an Open Archival Information System,* ISO 14721:2003 (for the concept of *archive* and *designated community:* "an organization that intends to preserve information for access and use by a designated community)

[43] *The large-scale archival storage of digital objects / Jim Linden, Sean Martin, Richard Masters, and Roderic Parker,* 2005, http://www.dpconline.org/docs/dpctw04-03.pdf

[44] *The Google file system* / Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung,, 2003, http://labs.google.com/papers/gfs-sosp2003.pdf

[45] http://it.wikipedia.org/wiki/Serial_ATA

The infrastructure does not rely on custom or proprietary software but is based on open source operating system and utilities (widespread acceptance means less dependency).

Data replication relies on open source disk synchronization utility (rsync[46] for UNIX); to avoid hardware dependencies (ex. g. disk controllers) RAID[47] is not used.

The Digital Stacks system is based on two main deposit sites (managed by the Biblioteca Nazionale Centrale di Firenze and by the Biblioteca Nazionale Centrale di Roma) and a dark archive (managed by the Biblioteca Nazionale Marciana, Venezia).

Each main site is composed by a set of autonomous and independent nodes. In turn, each node on a given site has a mirror node on the other site: *Digital Stacks* service does not rely on a "master site / mirror site" architecture and each site will contain – in a symmetrical way - both master nodes and mirror nodes (*Figure 1*).

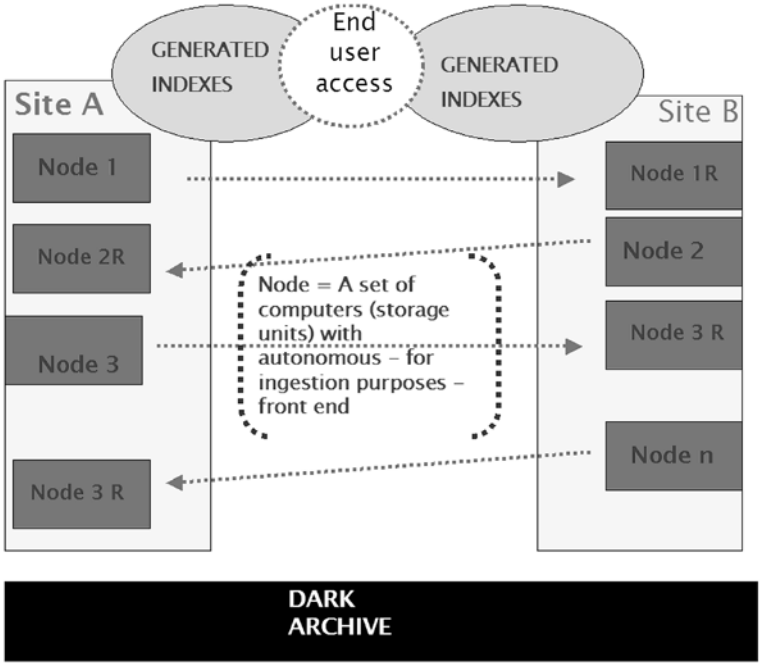


Figure 1: Magazzini digitali technical architecture overview

[46] http://it.wikipedia.org/wiki/Rsync
[47] http://it.wikipedia.org/wiki/RAID

Each physical file is replicated 2 times on different computers within the same node. Dark archive contains also two copies of the same file on two different computers. As a result within *Digital Stacks* each physical file is replicated 6 times.

As said above, the *dark archive* is for preservation and disaster recovery purposes only. The original plan for this site was to use an offline storage system (ex. g. LTO[48] tapes). However for the operational service we decided to use the same technology used in the two "light archives" (i. e. online storage using just simple personal computers). The use of the term online here does not change the purpose of the dark archive that is "to function as a repository for information that can be used as a fail-safe during disaster recovery[49]".

As regards repositories, setting up one main site in Florence, close to the Arno river, with flood risks, and the *dark archive* in Venice with the well known "acqua alta" (or high tide) problem, could not be considered a bright idea: these locations could result in a relevant threat for the security of the overall service. For this reason one important decision was to locate all the hardware on external data center (or collocation center[50]).

Certification to ISO 27001[51] international security standard is the basic prerequisite for the selection of a data center. Each institution (Florence, Rome and Venice) has selected 3 different data centers owned and managed by 3 different companies (to reduce the commercial risk of "domino" effects). Moreover we decided that the 3 collocation centers have to be at least 200 km far away from each other (to reduce the risk of natural threats)[52].

This architecture, based on compliance certification to ISO 27001 international security standard, is the basis for a domain specific certification

[48] http://en.wikipedia.org/wiki/Linear_Tape-Open
[49] http://www.webopedia.com/TERM/D/dark_archive.html
[50] http://en.wikipedia.org/wiki/Colocation_centre
[51] ISO/IEC 27001:2005 "specifies the requirements for establishing, implementing, operating, monitoring, reviewing, maintaining and improving a documented Information Security Management System within the context of the organization's overall business risks"
[52] The collocation center:

- for Rome is located in Bologna (CINECA);
- for Florence is located in Turin (COLT TELECOM);
- for Venice is located in Rome (ALAMAVIVA).

of *Digital Satcks* as trusted digital repository (during the test phase we tried to apply DRAMBORA[53], but also TRAC[54] was taken into account).

Digital Stacks is trying to take into account also the Archivematica[55] model that "uses a micro-services design pattern to provide an integrated suite of free and open-source tools that allows users to process digital objects from ingest to access in compliance with the ISO-OAIS functional model".

The *Digital Stacks* core is quite simple. *Digital Stacks* could ingest to date two kinds of files:

- data wrapped in WARC containers: WARC (ISO 28500) container aggregates digital objects for ease of storage in a conventional file system[56].
- metadata wrapped in MPEG21-DIDL containers[57]: MPEG21-DIDL (ISO 21000) is a simple and agnostic container suitable for the representation of digital resources (sets of metadata compliant to different S*chemas*)

It is important to note that Digital Stacks could take into account in the near future other kind of aggregation formats, i.e. conventions for wrapping primary (data) and secondary content (metadata).

Digital Staks is also facing the *metadata management problem (*also known as *lake or river model[58]).* A *long term archive*, and particularly a *legal deposit long term archive*, can not rely on a *lake model* (stores of metadata based on few *schemas* and fed by a few principal sources). A *long term archive* has to face stores of metadata based on *schemas*[59] that can change over time and which are fed by many streams. It could be based only on a *river model.*

[53] http://www.repositoryaudit.eu/

[54] *Trustworthy Repositories Audit & Certification (TRAC)* http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

[55] http://archivematica.org/wiki/index.php?title=Main_Page

[56] IISO 28500:2009 "specifies the WARC file format: to store both the payload content and control information from mainstream Internet application layer protocols, such as the Hypertext Transfer Protocol (HTTP), Domain Name System (DNS), and File Transfer Protocol (FTP); to store arbitrary metadata linked to other stored data"

[57] ISO/IEC 21000-2:2005: "The Digital Item Declaration Model describes a set of abstract terms and concepts to form a useful model for defining Digital Items [...], is based upon the terms and concepts defined in the above model. It contains the normative description of the syntax and semantics of each of the DIDL elements, as represented in XML".

[58] http://orweblog.oclc.org/archives/001754.html

[59] Schema is used here as http://www.w3.org/XML/Schema: "XML Schemas express shared vocabularies and allow machines to carry out rules made by people"

In a *long term archive* we have to assume there will be different metadata *schemas* originating from – using the PREMIS language - different *agents* (ex. g. OAI-PMH metadata harvesters, metadata extractors like JHOVE, Librarians, etc). Each *schema* is subject to change over time. Semantic overlap elements belonging to different Schemas (ex. g. PREMIS, MIX) will be probably the norm rather than the exception.

Since *metadata* are the only mean to control *data*, it is essential to have the *control of metadata* to avoid the risk of a *Babel Tower model*.

We are working on that taking into account the fact that there are no tools available. There are some interesting directions, but perhaps the most promising research activities are about "linked data"[60]

As promised the second part of this presentation will take into account briefly both organizational model and service model of Digital Stacks.

Digital Stacks realizes what the most recent Italian law on legal deposit (L. 106/2004, DPR 252/2006) provides for: a trial period for legal deposit on a voluntary basis of electronic documents that are defined by the law as "documents disseminated via digital communication network". This legislation can be regarded to as a strong commitment for national libraries to set up national long term digital preservation infrastructure for legal deposit of Italian publications. As we know, the "commitment" is one of the requirements of a trusted digital repository[61].

Fundings come from the Italian Ministry of Culture, with the support, in terms both of human and financial resources, of Fondazione Rinascimento Digitale (FRD).

The following three main goals have to be pointed out:

- To implement an *organizational model* suitable for creating the national and regional archives of electronic publishing production, as provided for by the law, and for being extended on a larger scale
- To implement a *service model* suitable for balancing the right-holders interests in contents protection with the final users ones in contents access

[60] http://www.w3.org/DesignIssues/LinkedData.html , http://en.wikipedia.org/wiki/Linked_data

[61] *Trustworthy Repositories Audit & Certification (TRAC)* http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

- To implement a sustainable *system* suitable for ensuring long term preservation and access to digital contents, as well as their authenticity (identity and integrity), as we have seen above.

In order to achieve these goals a legal and agreements framework is needed, also for balancing different interests of all the involved stakeholders. Two agreement ha been signed.

The first one (signed in November 2010) is an agreement between the three National libraries and the FRD in order to set specific roles and responsibilities of each institution from different points of view: scientific, technical, operational and financial, and to set up a steering committee for all management, monitoring and results assessment activities. It will be also of utmost importance to define an organizational and financial sustainability plan, after the 36 months trial period. The second one is a framework agreement between National Libraries and the Italian associations of electronic publishers (signed in July this year). This agreement sets out the following points:

- Digital Stacks will collect the agreed electronic documents (harvesting operated by Digital Stacks is the recommended way, but also other ways of deposit could be agreed);
- Digital Stacks will ensure long term preservation and access of the deposited documents, and will track any changes in the same documents;
- Digital Stacks will be allowed to perform any necessary actions (refreshing, duplication, migration etc) in order to achieve long term preservation and access of the deposited documents;
- only registered users and only from national libraries premises will be allowed to access deposited documents;
- access to deposited documents will also be allowed to regional deposit libraries, in the same way, but only to deposited documents of publishers whose registered office is in the same region of the deposit library;
- files printing and/or downloading will be subject to a specific license: in this license the publisher can allow also the delivery of documents to other libraries or even the access to the documents for registered users outside national libraries.

A final remark about sustainability. As you know, access to born digital e-journals is normally subject to a license. A typical provision of these licenses concerns the perpetual access to the licensed contents. It's a provision of the utmost importance for libraries and their users, and the only way for libraries to maintain over the time the availability of the contents they have paid for. But at the same time it is a provision that can be fulfilled only through a dedicated organizational and technical infrastructure, i.e. a trusted digital repository. It's unlikely the publishers will manage such an infrastructure. So this kind of service could be provided from the legal deposit libraries network, and its value could be part of the negotiation with publishers[62].

[62] A comparative study of e-journals archiving solutions. A JISC funded investigation. Final report, May 2008 / Terry Morrow, Neil Beagrie, Maggie Jones, Julia Chruszcz. http://www.slainte.org.uk/news/archive/0805/jiscejournalreport.pdf

María ALARCÓN RENDÓN
Mexico City, Mexico
Information Technology Director
National Archive of Mexico

Strengths and Weaknesses of the Situation of Management and Preservation of Digital Documents in Mexico

Introduction

The technical challenges of managing and preserving digital documents in Mexico do not differ too much from those being faced by other countries around the world. However, Mexico faces a specific situation regarding the way in which digital documents are and should be treated as means of preserving the country's documentary memory.

The purpose of this document is to present a general diagnose of the situation of digital records and archive management in Mexico taking into consideration both technical and human aspects concerning this matter.

SWOT Analysis

The method used to present the current situation of digital documents in Mexico is using a SWOT (Strengths, Weaknesses, Opportunities and Threats) Analysis. This type of analysis is centered in dividing the different aspects that influence the specific topic (management and preservation of digital documents in Mexico) in order to evaluate how the different issues can be combined and made into strategies to achieve a specific objective. The reach of this document is to describe the elements of the analysis that will describe a diagnostic of the current situation to help records management and archive institutions in their decision making processes.

In order to carry out this analysis, the objective proposed is that Mexico can attain an optimal level of efficiency in the management and preservation of digital documents.

Strengths

In order to attain the objective mentioned, the strengths of Mexico in the subject of management and preservation of digital documents describe the internal factors that currently affect positively the way these activities are being done.

Human Resources

In Mexico there is a strong group of Information Technology (IT) professionals that have very good knowledge of technology and Business Process Management (BPM) as well as Enterprise Content Management (ECM). Universities around the country, as well as companies and government institutions are interested in aligning their business processes to IT assets. In that same way, business and public administration professionals and Industrial Engineers are going deeper into process standardization knowledge. Finally, there are national groups of records management professionals and archivists with a good knowledge of process management.

Education

As far as education goes, there are a few initiatives that are starting to gain importance such as the "Diploma Course on Digital Information" done by the National University of Mexico (UNAM).

Regulation

In Mexico, there is a constitutional statement indicating the obligation of public entities to maintain "updated records". Thus, the importance of records management has a strength given by the Constitution itself.

A part from this mandate, the existence of some regulation concerning digital documents and electronic certification of documents has permitted a widespread usage of this type of technology. An example of this is the tax payment which is being done with electronic signatures and certificates.

Within the public administration, a step forward towards the proper management of digital records has been the "General Aplication Administrative Manual of Information Transparency and Records". In this document there is a section that refers to the management of electronic records. Even though it is just a start, the importance of having a section on this matter represents an advantage.

Records Management

It is a fact that the majority of the federal government institutions in Mexico have already a records classification scheme that serves as the basis for all the records management in each institution. This represents a strength even though the majority of these schemes are thought for paper but at least there is a basis with which to start when going on to electronic documents.

Also, there is an interest of public and government institutions to use digitization as a means of preserving and spreading current and historical information that records management and archives manage.

Weaknesses

On the other hand, there are the internal aspects that are affecting negatively the achievement of the goal. Here are some aspects to take into consideration:

Human Resources

Even though there are groups with a certain level of expertise in each of their areas of specialty, there is a lack of adequately trained professionals with the necessary knowledge of records management, IT and Process Management. There is a sense of urgency of trying to get the IT professionals to understand the logics of the records managers and archivists because if this is not achieved, the risk of losing an important amount of digital documents still exists.

Education

As part of the strengths, it was mentioned that there are a few initiatives that are currently starting to set interest in digital preservation matters. However, there is a lack of widespread training on digital records management and archiving. Thus, preservation in the long term is seldom regarded as an issue to take into consideration when studying records management or IT science. The efforts are just starting in this matter.

Records management professional curricula in universities (bachelors and masters degrees) are seldom considering digital file management. Much less are they considering digital preservation in the long term. Archival techniques and methods are basically oriented towards traditional formats (paper). The IT training given in these programs are more oriented towards technical usage of computers.

Regulation

Even though there is some ruling about digital documents, in Mexico there is no national law for the regulation of records management and archiving. The good part of this "weakness" is that the National Congress is currently about to approve the National Records Law in which electronic documents are considered.

Another area that can be improved is that of e-Government services. As mentioned before, there is an interest of the country in making processes more efficient by using electronic means to make procedures easier for the citizens. However, legislation should be moving towards including standards for long term preservation of all the electronic information being generated by these systems.

Records management

In institutions and companies around the country, there is a lack of a widespread use of international standards such as ISO 15489 in records management and digital archiving. As mentioned before, there is an interest in preserving documents by using digital images of documents. However, few projects can assure they are using proper metadata, procedures and standards when implementing them.

Some institutions are currently working on very good examples of the implementation of standards in their digital document management systems. However, these efforts are isolated and very specific to each institution. There is a need to work on measures towards an effective digital records management system country wise.

Opportunities

There are several opportunities that records management institutions and archives in Mexico can take into consideration to attain an optimal level of efficiency in the management and preservation of digital documents.

An important element that could give the country a strong support would be the creation of an intergovernmental discussion group that would work towards the standardization of digital records and archive management. The idea of this group would be to create a group of "rules" or proposed procedures for the "organization and preservation of records and archives in the institutions that form part of the public administration."[63] Efforts are starting to be done in this area.

[63] Barnard, Alicia. "Proposal of rules for electronic record documents and their preservation in the long term". Mexico, 2011.

Another external opportunity is the acquirement of knowledge of those organizations at an international level that already have had an experience in digital records and archive management. The participation of Mexican records keeping institutions and archives in international groups such as InterPARES or UNESCO IFAP is very important because they can give a great value to the efforts being made in these topics.

Threats

Finally, the existence of unavoidable external factors that may represent an obstacle in attaining the objective mentioned, basically come from the end users of the information.

One of the main threats that affect digital document management and preservation is that there is a growing tendency of digital information to grow. According to a study made by IDC and EMC[2] in 2008, "the digital universe is bigger and growing more rapidly than original estimates as a result of accelerated growth in worldwide shipments of digital cameras, digital surveillance cameras, and digital televisions as well as a better understanding of information replication trends."[64] The same tendency is occurring and affecting Mexico. This threat comes, not from the weakness mentioned before (where digital images are not correctly generated by institutions), but an inherent property of digital systems and from the people's increasing use of electronic means to carry out their everyday activities.

This is a situation to be taken into consideration especially because it is information that is being generated in a disorderly manner, which might be causing "abrupt losses of important documents"[65]. On the other side, there is the increasing amount of "e-Waste"[66] being generated by non-properly managed digitization projects and irrational use of "permanent" storage space where there is no space at all. Even though, disc space and other means of storage may not be too expensive, the risk that comes with waste should be taken into consideration as a threat.

[64] "*New Study Forecasts Explosive Growth of the Digital Universe; Spotlights Worldwide Phenomenon of "Digital Shadow*" http://www.marketwire.com/press-release/new-study-forecasts-explosive-growth-digital-universe-spotlights-worldwide-phenomenon-nyse-emc-831713.htm (accessed September, 2011).

[65] "*Digital Records in Mexico: Achievements? Perspectives...*". Article written by Alicia Barnard Amozorrutia and Aurora Gómez Galvarriato Freer. December, 2010.

[66] "*New Study Forecasts Explosive Growth of the Digital Universe; Spotlights Worldwide Phenomenon of "Digital Shadow*" http://www.marketwire.com/press-release/new-study-forecasts-explosive-growth-digital-universe-spotlights-worldwide-phenomenon-nyse-emc-831713.htm (accessed September, 2011).

Summary

The following figure shows a diagram of the SWOT analysis of the situation of management and preservation of digital documents in Mexico:

	INTERNAL	EXTERNAL
POSITIVE	**STRENGTHS** □ Human Resources: IT, BPM, Record keeping and archival professionals. □ Education: Few initiatives in digital document management. □ Regulation: Constitutional statement, regulation concerning the use of electronic procedures. □ Records management: use of digitization as a means of preserving and making archives public.	**OPPORTUNITIES** □ Creation of an intergovernmental discussion and cooperation group. □ International cooperation.
NEGATIVE	**WEAKNESSES** □ Human Resources: professionals must work together or new multidisciplinary professionals must be trained. □ Education: Lack of widespread training on digital records management. □ Regulation: efforts must be made to achieve a National Records Law. □ Records management: lack of a widespread use of standards.	**THREATS** □ Exponential growth of the "Digital universe". □ Creation of "e-Waste".

The next step in order to achieve the efficient management and ways of preserving digital documents in Mexico would be to take each of the elements presented here and evaluate what strategies can be drawn in order to achieve the objective mentioned at the beginning of this analysis. Some of conclusions that might develop into further strategies are the following:

- It is clear that cooperation and education must be implemented to take advantage of the knowledge current professionals have.
- The strength given by the National Constitution and the emerging efforts in electronic document regulation must be the basis to support the National Records Law and help it come to reality.

- Team work must be implemented between government and private organizations taking into consideration the experience of those that have made isolated efforts in this matter.

References

1. "*Digital Records in Mexico: Achievements? Perspectives...*". Article written by Alicia Barnard Amozorrutia and Aurora Gómez Galvarriato Freer. December, 2010.
2. "*New Study Forecasts Explosive Growth of the Digital Universe; Spotlights Worldwide Phenomenon of "Digital Shadow*" URL: http://www.marketwire.com/press-release/new-study-forecasts-explosive-growth-digital-universe-spotlights-worldwide-phenomenon-nyse-emc-831713.htm (accessed September, 2011).
3. Political Constitution of the United States of Mexico. Secretaría de Gobernación. Mexico, November, 2009.
4. "General Aplication Administrative Manual of Information Transparency and Records". URL: http://www.normateca.gob.mx/Archivos/50_D_2755_11-08-2011.pdf (accessed September 2011).
5. Barnard, Alicia. "Proposal of rules for electronic record documents and their preservation in the long term". Mexico, 2011.

Section 1. Digital Preservation of Text Documents

Elena KOZLOVA
Director
INFORMREGISTR Research and Technology Centre
Moscow, Russian Federation

The Preservation of Digital Content within National Information Library Collection

The broad dissemination of digital resources led to the appearance of an information environment containing numerous objects of varying levels, as far as the content and technology are concerned. Contemporary digital information resources possess varying formats, are meant for use in diverse program environments, and notably differ from each other in the processing level and quality of content. That is why there is no unified solution for the problem of digital resource preservation.

The progress of information technologies necessitates permanent evolution of data formats, software and hardware, which in its turn demands permanent development of preservation technologies, closely connected with the semantic aspects of building storage arrays. There are two priority issues to be solved for successful development in this field:

1. The selection of digital objects for storage
2. The organizational and legal aspects of preservation

The formal change of presenting information with the development of IT is reflected in an acting law of the Russian Federation – Federal Law No 77, of December 29, 1994, On Legal Deposits. The extensive public need for electronic documents demands their preservation and bibliographic control. With this goal in view, electronic forms of documents (publications) are included in the list of materials to be delivered as legal deposits – though only those on removable media. Selection criteria have been established to build national library information collections.

The following criteria are applicable to all objects. The law pertains territorially only to documents produced in Russia, and concerns a restricted number of electronic distribution technologies (the law

"does not apply to ... electronic documents distributed exclusively via information and telecommunication networks"). Electronic documents are also selected from the point of:

function (official documents, standards and patents – without specifying the media; software and databases);

presentation format (sound and visual products, and electronic documents);

method of publication (unpublished documents without explicit medium, and electronic publications).

These criteria are applied to the distribution of digital documents among authorized organizations and aim to establish the rules of their bibliographic control.

The classification stipulated by the federal law On Legal Deposits has led to the determination of selection principles. A library information sector is being formed purposefully and regularly to preserve digital documents pertaining to the cultural heritage. Naturally, such objects demand lasting preservation and access.

There are two basic aspects of digital document preservation in national library information collections:

- Preservation during the collection formation – the establishment of a metadata system and their preservation;
- Preservation in storage – support of digital storage and provision of access.

Bibliographic control stipulated by the law guarantees the formation of a metadata system in which every document's metadata contain unambiguous index and bibliographic information, which allows the control, retrieval and monitoring of the national collections. All subsequent preservation measures are planned and implemented proceeding from metadata.

Though the federal law On Legal Deposits makes explicit requirements to the preservation of digital documents, the bylaws on preservation in storage do not stipulate practical action, so every authorized agency for legal deposit acquisition and storage preserves digital documents independently.

The Informregistr Research and Technology Center, one of the sectors of the national library information collections of digital documents, presently possesses more than 24,000 titles of federal legal deposits the metadata on which are found in the Russian Electronic Publications catalog on the Informregistr website.

Analysis of digital legal deposit collections shows that the present procedure for metadata creation and resource selection for the national library information collections deserves to be spread to a more extensive number of digital objects.

The new formats of digital documents require extending the action of the federal law On Legal Deposits to other kinds of documents – web-documents and immutable files on dedicated hardware.

There are no legal and organizational provisions for web-publications control as yet, so there are no national union registers either. The same holds true for immutable files on dedicated hardware. Users can't get complete and reliable information on the available choice of resources, while individual publishing institutions fail to provide resource preservation in the full range of procedures.

The quality of documents in information web resources differs, and not all can be regarded as cultural heritage. It is impossible to collect all digital documents. That is why we need standards for the selection of documents deserving preservation.

Analysis of digital documents necessarily includes its origin. Just as digital documents on removable media, there are two kinds of web-documents – born-digital documents and digital copies of documents originally created on other media. When digitized from hard media, a document also gets a hard copy, therefore such information has larger capacity for survival while documents with no hard copies are more endangered and so should hold priority.

This classification is of special importance in preparing metadata and building accounting arrays as digital copies on other media contain original identification data. It is difficult in certain instances to distinguish between data on the original and digital copies – which means that the metadata system should contain information about the origins of the digital document.

Imprints of digital document on removable media comprise data of producer, which helps to identify resource authorship. Information resources in information and telecommunication networks do not always comprise data on publication time and place, and so impede reliability assessment. To identify web-documents, each of them should comprise a set of imprint data comprehensively characterizing digital documents: title of digital document; name of individual or legal entity on whose behalf it is prepared; data on subject scope, readership and purpose, web-site address; access right limitations.

Life length is an essential aspect of document preservation. Every object is intended either for short-term or long-term storage. The former envisages storage without an account for technological changes, and preservation is achieved by archiving and backups. The latter envisages preservation for an indefinite period and future access. These goals are not achieved by simple backing up and demand an account for the prospects of changes in software and hardware, and an evolution of media and formats.

Digital document preservation has three vectors. Depending on the goals of a particular depository for particular arrays, we may preserve:

- content;
- function; or
- both content and function.

Digital documents on removable media can be preserved currently at the physical level due to the lack of right to copy and re-record. Besides, producers protect their products from unauthorized circulation through one-time installation to rule out the use of backup technology. This contradicts to principal requirements and technological trends for digital preservation. It particularly impacts the demand for content preservation together with functionality. Designing virtual devices to reproduce software environment of the obtained information resource appears the most efficient option.

The preparation of digital arrays for preservation as the national library information collection is formed demands the development of legal, organizational and technological conditions of depositing alongside setting selection standards.

The national preservation programs abroad offer the following web resource selection criteria:

- Collection of all e-documents, e.g., Internet Archive;
- Collection of e-documents of the national domain;
- Collection of e-documents on a particular theme/thematic range;
- Website selection and monitoring their development;
- Event-driven acquisition of information.

The European Union countries' national legal deposit projects select web resources according to:

- Domain names.
- IP addresses.
- National affiliation.

These criteria guarantee complete acquisition of information resources within a thematic area – a relevant criterion for national digital document archives. However, they do not provide a selection of resources pertaining to the national cultural heritage in the library community.

The preservation of digital documents is the duty of various organizations – libraries, archives, government agencies and educational institutions, and every group of organizations requires its own set of standards for selecting digital documents to qualify as national heritage.

The formation of a national library information collection of web documents for long-term storage requires additional standards of mandatory selection according to the availability of metadata (completeness of document description) and the relevance of documents for library and information collections.

The building of a common information environment comprising information about the content of national depositories/archives is a goal shared by them all, of whatever type. To meet this goal, they need national standards of paperwork and presenting identification data and information resources for long-term storage.

It takes legislative amendments and the drawing of regulatory acts on digital information preservation to streamline the formation of a national library information collection. It is necessary, above all, to expand digital acquisition to web editions and E-Reader files.

The expansion of the digital collection range requires:

- The development of selection standards for digital legal deposits belonging to cultural heritage;
- The development of techniques for collection/delivery of web resources by their manufacturers.
- The preservation of digital resources requires:
- The development of the legal basis of preservation;
- The development of standards of information resource presentation for long-term storage;
- The development of storage standards: life, format and access;
- The development of standards for long-term storage, and transfer to repositories; and
- The establishment of government information resource repositories.

Digital heritage preservation demands the development of relevant strategies at the national and international levels and the drawing of a final document setting basic goals for the years to come:

1. The elaboration of methods and strategies of national digital heritage preservation;
2. The establishment of a common information environment for access to metadata on the preserved digital heritage;
3. The promotion and coordination of research on digital heritage preservation, including search for technologies guaranteeing its long life.

Alisher ISHMATOV
Tashkent, Uzbekistan
Director
Information Library Centre of the Republic of Uzbekistan

Uzbekistan's State Policy on Information and Library Collections Preservation

Over the recent decade Uzbekistan's library community has witnessed and participated in nation-wide reforms in librarian science. These reforms have covered mainly the normative and legislative basis of the library activities; introduction of information and communication technologies in libraries; review of the current and development of new educational standards and programs in the field of information and library training, etc.

Impulse to the reforms was given by the Decree of the President of Uzbekistan on *Provision of the Republic's Population with Information and Library Services* adopted on June 20, 2006. According to the Decree, information and library centers (ILC) and information resource centers (IRC) were established on the basis of the libraries under the Ministry of Culture and Sports. These centers were placed under the charge of Uzbekistan's Agency for Communications and Informatization (UzACI) and the Ministry of Public Education and the Secondary Special and Vocational Training Center, respectively.

The main goal set in the Decree was to intensify the introduction of ICTs in library processes and promote the leading role of Uzbekistan's Agency for Communications and Informatization in providing advanced training of library ICT personnel.

Another major task to be resolved by the libraries of Uzbekistan is to preserve books and manuscripts and to make them widely accessible to the users. This concerns primarily the Alisher Navoi National Library of Uzbekistan and the information and library centers of the Kara-Kalpak Republic, Uzbekistan's regions of and the city of Tashkent which are freely accessible to all groups of users. The National Library, information and library centers of Bukhara, Samarkand, Khwarezm and Tashkent, as well as information and resource centers of the National University and the Samarkand University boast rich collections of local historical and ethnological artifacts and rare books. The republic counts over 2 million items in total, of which 400 thousand are rare, highly valuable and unique

publications. All these precious materials are in great demand and of high scientific and historical value, since many of them exist in a single copy only. It is extremely important to make these rich collections accessible to users in Uzbekistan and other countries while securing their preservation.

However, documents that are in long-term storage in libraries and other institutions are gradually deteriorating. In the poorest condition are documents, which were published on acid paper in the second half of the 19th - early 20th centuries, and manuscripts with damaged covers and separated pages due to long storage record. Additionally, such poor condition of documents is the result of lack of adequate equipment and materials. The study of the current situation has shown that the main attention should be given to manuscript preventive treatment, adequate and timely preservation (conservation) and, finally, creation of digital copies of manuscripts. This leads us to the conclusion that the better the conservation procedures, the less serious deterioration of a manuscript with time and the less the need for restoration. The originals should be exposed to contact as little as possible in order not to lose their historical value, and their digital copies could be provided for use instead.

In order to preserve the national documentary heritage as part of the global intellectual assets stored in the ILC and The National Library of Uzbekistan, it is essential to introduce and use modern information technologies and tools and build a solid material and technological base to allow the implementation of all the necessary document conservation and preservation measures.

Of utmost importance for the library community of Uzbekistan was the adoption of the Information and Library Activity Law, which was developed with a focus on ICT, and Decree of the President of Uzbekistan On Measures for Further Quality Development of ICT-Based Information-Library and Information-Resource Services for 2011—2015. Adopted on February 23, 2011, the Decree included provisions on building up a collection of digital resources, inter alia by creating digital copies of printed publications and manuscripts, primarily of unique and highly valuable publications and publications that have moral, cultural, historical, scientific or educational significance for the creative development of intellectual and harmonious individuals. This strategic plan served as a basis for a draft executive order on the launch of information and library resource restoration and conservation departments under the Uzbekistan National Library and the Institute of Oriental Studies at the Academy of Sciences of Uzbekistan.

Apart from that, the plan is to the organize archaeographic expeditions and the establishment of an expert group for collecting and processing privately-owned manuscripts, lithographic publications and historical documents in order to convert them to the digital format and create backup and working copies. Thus, the double aim of preserving these materials and providing the users in Uzbekistan and elsewhere with the access to them will be achieved.

However, special measures are required to preserve unique and highly valuable archives. It is necessary to:

- Establish specialized departments within information institutions and libraries;
- Equip laboratories with adequate technologies and materials for conservation and restoration, as well as creation of digital copies;
- Create digital (backup) copies.
- Train library staff on the creation and storage of digital copies;
- Create conditions necessary for preserving antique documents;
- Provide for safe use of such materials and their restoration in case of damage or loss.
- Ensure full user access to digital copies.
- Ensure ideal storage environment for antique and highly valuable documents.

The first steps in this direction have already been taken: training of the personnel of information institutions and libraries has been organized in collaboration with the Goethe Institute in Tashkent and the U.S. Embassy in Uzbekistan. Workshops, videoconferences and master classes on advanced technologies in the field of document conservation and restoration have been held.

Initiated by the President of Uzbekistan, the construction of the National Library is nearing completion. The new building will have premises for the restoration and conservation department and specialized well-equipped laboratories. A similar center is to be opened at the Institute of Oriental Studies at the Academy of Sciences of Uzbekistan.

In 2012-2013, a new restoration and conservation specialization will be launched at the Tashkent State Institute of Culture.

Lyubov Kazachenkova
Moscow, Russian Federation
Editor-in-Chief
Sovremennaya Biblioteka (Modern Library) Magazine

Five-Item Not-to-Do List of Digital Information Preservation

Knowledge becomes real knowledge only when it is spread, when the new generation develops, expands, deepens and transfers it to the succeeding one. Knowledge ceases to be knowledge when it falls into oblivion. This obvious and simple truth is evident; the modern day, however, makes us remember it all the time. Data, information, knowledge... I'm not going deep into terminological details - let the philosophers, sociologists, or information specialists do this for me. Being user and owner of printed and digital information, I am interested in other matters: to what degree the legacy created over the latest two decades of tremendous ICT boom, remarkable technological progress, modernization, globalization of economies and virtualization of everyday life would become knowledge for the future generations. Or whether it will become an information dump where everything will be jumbled and the only option left will be to bury and forget everything.

Today, the amount of information is growing at a cracking pace. The researchers from the USC Annenberg School for Communication & Journalism have evaluated the total global technological capacity (http://www.sciencedaily.com/releases/2011/02/110210141219.htm). The study aimed at finding the answer to a simple question: how much information humankind is able to store, communicate and analyze. "We live in a world where economies, political freedom and cultural growth increasingly depend on our technological capabilities," said lead author Martin Hilbert. – "This is the first study to quantify humankind's ability to handle information." Researchers calculated that at both digital memory and analog devices humankind is able to store at least 295 exabytes of information (a number with 20 zeros in it). That is 315 times the number of grains of sand in the world, but it is still less than one percent of the information that is stored in all the DNA molecules of a human being.

The year 2002 could be considered the beginning of the digital age, the first year worldwide digital storage capacity overtook total analog capacity. As of 2007, almost 94 percent of our memory was in digital form. In 2007,

humankind successfully sent 1.9 zettabytes of information through broadcast technology such as televisions and GPS. That's equivalent to every person in the world reading 174 newspapers every day. On two-way communications technology, such as cell phones, humankind shared 65 exabytes of information through telecommunications in 2007. Another interesting fact stressed by the researchers is that in 2007, all the general-purpose computers in the world computed $6{,}4x10^{18}$ instructions per second, in the same general order of magnitude as the number of nerve impulses executed by a single human brain. Doing these instructions by hand would take 2,200 times the period since the Big Bang. "These numbers are impressive, but still miniscule compared to the order of magnitude at which nature handles information", Hilbert said. - "Compared to nature, we are but humble apprentices".

Yes, we are just learning to analyze our capabilities, develop strategies, and maintain real-life contact with each other despite intensive communication in social media that seem to be gradually replacing everything else. And time and again, we face the same questions. What do we retain in our memory? How do we store the knowledge: in our mind, in book cases and on bookshelves, in strongboxes or depositories? Do we really need to preserve all these riches? Who might be in need for them? Who will do this?

The problems of digital information preservation are just emerging and humankind has not yet fully comprehended them. Many aspects remain unclear and call for joint efforts of technologists, sociologists, politicians, cultural researchers, librarians, etc. Of course, I am not going to instruct anyone on what they should do. Held in Moscow on October 3-5, 2011, the International Conference "Preservation of Digital Information in the Information Society: Problems and Prospects" gathered experts from all over the world to became the first interdisciplinary thematic forum. It considered quite a number of problems: from political ones that define preservation strategies to specific technological aspects, such as scanning standards, etc. The Conference adopted the Moscow Declaration on Digital Information Preservation, which highlights the importance and relevance of the issues raised and proposes a set of measures and steps in politics, awareness-raising and education, scientific research, economy, ICT industry, partnership and coordination. The document was prepared by a group of experts from Argentina, Austria, Denmark, India, Italy, Russia, South Africa and other countries.

And still. Times goes on, but travelling a lot across this country, I can sadly admit that little has changed yet. Of course, such problems cannot be solved overnight, but we could at least outline the strategy for their solution. This

reminds me of something I experienced in the past. Library journalist for 25 years, I could witness many transformations in our library world. This is why when speaking about digital preservation I will tell you about my five-item not-to-do list including the things that I in no way want to happen when pursuing this goal.

Another déjà vu

My first item on the list is never to go through the experience of the 1990-ies again. Let's refresh our memory on it. The country was automating its libraries. What was the procedure? At first, believing themselves capable of it, almost everyone started to develop their own integrated library systems and came up with a great many of homemade designs. Each and every programmer who ever came to work for a library (and at that time they were recruited under an employment contract) committed to demonstrate their skills and undertook to design a unique masterpiece. Sometimes, however, the situation was different: the newcomer was unable to cope with the ILS designed by his or her predecessor, and therefore started to develop the program anew. Unfortunately, this trend is still alive in web-design. Many library web-sites, though attractive and functional, can hardly be upgraded or updated just because the open codes where ignored during their creation, which means that a new person can change virtually nothing without the help of the original designer of the web-site.

Almost 20 years ago, one could only wonder at the names given to the ILSs. The legacy from the past days can still be found in the Russian libraries. I believe everyone is aware of the consequences that followed. I hope library management students know it too.

Let me focus on some of the problems: system incompatibility, impossibility of data exchange and union catalog acquisition or their use, formatting, isolation in spite of the seeming progress. Swords crossed when it came to a common machine-readable communication format of bibliographic description. Much like today no professional meeting can leave legal issues aside, the formats were heart and soul of the epoch.

Finally, the publication of the book by Yakov Shrayberg and Felix Voroysky "Automated Library-information Systems of Russia: current state, selection, implementation, development" in 1996 straightened everything out. The book described the most reliable ILSs of the time. The libraries undergoing automation process gradually recognized the advantages of the existing systems and began to purchase complete programs that had already been tested.

And then followed the adoption of the Russian communication format to represent bibliographic records in a machine-readable form, the emergence of the LIBNET program, the conception and the evolution of the Russian Union Catalog, and the development of authority files and other civilized automation features.

It seemed things started to look up...

Everyone can notice that technologies develop much faster than it takes us, the users, to adapt to them, and definitely way faster than the technologies finally reach libraries. Technological advancement and improvements, however, were accompanied by an "unexpected" trouble - digitization. Time and again we would learn from our own mistakes rather than those of others. Once scanners purchased (putting quality aside), digitization spread all over the country. There was nothing bad in libraries converting books, newspapers, and leaflets to digital formats. That was a pure supply-demand model. Until there were no legal obstacles, everything went well. It was OK for a single library. Nationwide, though, the efforts resulted in absolute chaos. I wonder how much money was spent on digitization of one and the same work on a national, or at least, regional level? Then the Civil Code was adopted and the digitization process slowed down. It looked like the right moment to review digitized collections and build up a register of digital documents in library collections, as well as to develop a strategy for digital conversion of library collections. As far as I remember that was a goal pursued by the Strategy for Developing Information Society in the Russian Federation. The question "What's next?" remained unanswered.

Today, we face the problem of digital preservation. I do not want to repeat the story of automation and digitization campaigns. There is definitely the need for a head agency that would take decisions on the strategy and principles, organize expert meetings and forums, etc. Since the Russian Committee of the UNESCO Information for All Programme has offered to organize such meetings, and since that it is the second time that the Russian Federation, represented by Evgeny Kuzmin, has chaired the Intergovernmental Council for the UNESCO Information for All Programme, it would probably be reasonable for the Russian Committee to head these activities, and even more so since it has already taken the first steps.

A sow is no match for a goose

Another not-to-do item is linked to industrial egoism. To a less extent this applies to libraries, or at least public libraries: they happen to be the most

open for cooperation. However, besides libraries, other structures such as higher educational institutions, museums, archives, various government-funded and commercial companies are also engaged in the digitization process. I am convinced that the national program for digital information preservation should take this into account. Moreover, it should be library-oriented, or should make them the basis for the process. Let me give you an example. A family creates its own digital photo archive and would like it to be accessible in 10, 20, or 100 years. Whom should this family give it to? I am not even talking about estimating its worth. The item that has little value now could well become a real treasure in 50 years.

Let's turn back to the past, though. Not such distant as 1990ies, the more recent one - back to the year 2010, i.e. the 65th anniversary of the victory in the Great Patriotic War. People from all over Russia were bringing their photos to libraries, making, mostly at home, their digital albums. Whom can they pass these valuable materials to?

Another requisite of our days is social media. In fact, they have become a kind of educational platform where people of various professions post materials, hold discussions and undertake case studies. Is this huge collection of documents subject to analysis and storage?

Or else, think about photo, video and presentation web-hosting services. Picasa alone is able to replace art department of a city library. On the amount of loaded files, Youtube exceeds musical collections of a lot of libraries. The amounts of data stored by SladeShare would surpass the collections of scientific and educational materials, and its list of topics would beat many academic libraries. Pinterest, aggressively entering into the social media market, will soon be able to replace library arts departments, as well as galleries. Where should these valuable resources be accumulated, evaluated, selected and stored? The answer is obvious – in libraries! Libraries have been recognized as the national document depositories meant to last for centuries. One can argue that archives perform the same functions, however, public access to archives is restricted while that of libraries it not. This makes the latter stand out in comparison to other institutions. It is their special, exclusive feature, and for this very reason they should become digital data storage areas.

One in the hand or two in the bush

If one cannot do much, one tries to achieve at least something. This, however, does not work for digital information preservation. Just a small

remark. Web-archiving was one of the topics discussed at the conference mentioned above. Grethe Jacobsen, head of one of the departments at the Danish Royal Library (Copenhagen), shared the ways of how it is done in her library. She noted that in accordance with the version of the Act on Legal Deposit of Published Material (part 3) as of 2004, all Danish material published in electronic communication networks is subject to legal deposit if it is published from Internet domains etc. which are specifically assigned to Denmark, or if it is published from other Internet domains etc. and is directed at a public in Denmark. The focus is on publications covering a wide range of topic: from social and political issues to sports through culture, economy and law. The copies are deposited in three libraries: the Royal Library, the State Library and the Aarhus University Library. However, Netarkivet.dk has been chosen as a third-party virtual agent for web-archiving. At present, the data is accessible for researchers only, which has already caused concern among the experts. Grethe Jacobsen highlighted the role of international cooperation in selecting, preserving digital resources, as well as providing access to them. Web archiving was also examined in the statement by Barbara J. Ford, director of the Mortenson Center for International Library Programs and Distinguished Professor at the Library of the University of Illinois at Urbana-Champaign, USA. She spoke about educational initiatives of the U.S. Library of Congress. Following one of them, personal archiving recommendations have been developed for the population to help them preserve their personal information (photo albums, audio, etc.). Another web archiving project engages students in selecting and archiving web resources according to certain criteria.

After the conference I have visited quite a number of Russian regions. What have I discovered there? When digital information preservation is brought up, the only subject discussed is web archiving. But how various organizations and institutions are to handle piles of their documentation? What, for instance, am I, editor-in-chief and publisher of the Sovremennaya Bibkioteka Magazine and the Igrovaya Biblioteka Magazine, to do with the materials sent to me by the authors? Where should I store them? A strict procedure was once established for storing printed materials: they were stored in the editorial office for a few years and then packed and sent to special depositories. And now, does anyone need this at all? Where should I deposit the digital issues of the magazines? No laws, norms, or standard procedures exist for that anymore! There is, of course, the National Library resource where the authors can bring their works to. But it is not a public organization, it is not in charge of long-term preservation and it is not a

library. And what about publishers' electronic archives? Following the existing law, it is easier to corrupt, erase, or destroy them rather than transfer them somewhere for storage.

This is my third not-to-do item: not to follow the principle "an integrated approach does not work? Let us do at least something". It is possible, of course, but then we are sure to return to the two previous items.

We will see in ten years

In 10 years, we definitely do not want to see anyone sorting out the piles of materials that we have digitized but for various reasons failed to preserve. It is common truth that it is harder to restore than to create. Indeed, there are a lot of problems today, many technological issues remain unsolved: formats, preservation media; scanning and storage standards are not defined, absence of principles and criteria for selecting digital information to be preserved, lack of appropriate infrastructure (just think of how long it took to develop infrastructure for preserving printed library collections). In fact, hardly anything has been done there except for few attempts of national and foreign libraries to do something. But we have to start at some point, anyway.

It will not straighten out by itself

I believe many people will support my next not-to-do item. I do not anticipate that someone will come and solve the problem for us, or that it will straighten out by itself. Noting like this will ever happen. But we should not jump from the frying pan into the fire, either. The society and those who accumulate, create, provide access to and preserve the created resources have to undertake conscious efforts. The development of a coherent national policy is urgently needed, which should involve various ministries and agencies. Let me stress once again that it is the libraries that should serve as platform for targeted measures. Moreover, the function they would perform should be listed among the obligatory public services that the libraries are to provide. These services should be government funded only! Of course provided that the State cares at all about leaving anything to the future generations.

I hope that the efforts of the Russian Committee of the UNESCO Information for All Program, Intergovernmental Council for IFAP, and widely known experts will give the expected results and a clear-cut action plan will soon be elaborated.

Igor KISELYOV
Moscow, Russian Federation
Head of Information Systems
Russian State Archive of Social and Political History

Digitization of Archive Documents: Retrieving the Value

A few years ago, when digitization was a novelty, archivists saw its goal and meaning as transparent and simple: to create a collection for use and, perhaps, a backup collection of archive documents. The procedure seemed quite easy: to insert a document into a scanner, then press a couple of buttons. However, as time has revealed, the digital preservation of archive documents is a complicated, multifaceted and multilayer task which includes many processes and reasons. This task has not been sufficiently studied in archiving so far, so let us try to grasp its general notions and determine factors that are significant for archive cultural heritage preservation.

First and foremost, we need to answer the following question: what heritage are we going to preserve, digital or traditional? The answer of the archiving community today is clear: traditional. There is no digital archive heritage in Russia yet. Digital content, which is produced in abundance by organizations and individuals in their daily routine, has not only been denied archival storage, but, at the legal or methodological levels, has not acquired the status of a valid document. This content consists today of files and file collections. These files should be preserved, too, because, though not considered valid documents, they contain information which, contrary to formal definitions, acquire with time the functions and meaning of a historical source.

Let us look more closely into the nature of the preserved traditional archive heritage. Mass cultural heritage properties are considered those preserved in libraries, museums, and archives – all the same, comma separated. When people speak about the preservation of cultural heritage, they mean scanning of library, museum and archive collections. The scanning itself is carried out the same way - one and all, comma separated or no comma at all.

In one aspect, which is very significant for preservation, these properties are differing. Each museum item is unique, since its value lies in its visualized component. Libraries store copies of mass publications whose value consists in the information which may not be linked directly to a specific

visual form (or format). Archive documents are in between: each archive document is unique, but its main load is information. This means that an archive document alongside with information value and have material value (market or auction value) as a unique artifact.

Here of course we mean archives, libraries and museums in their «pure» form and do not take into account manuscript collections of museums and libraries, collections of books and journals and museum items stored in archives. It should be also noted that the so-called duplicate archive materials (for example, one and the same directive sent to several subordinate organizations) do not mean documents of mass production, since each recipient organization's copy can bear resolutions and represent therefore a unique document.

As for archivists, they used to regard, until recently, preservation of cultural heritage as routine measures aimed at preserving archive documents. However, digitization of archive documents has been acquiring more and more relevance. There is no doubt that something is preserved following digitization, but we have to determine definitely and clearly which items representing traditional cultural heritage are to be preserved.

The answer would seem obvious – archive documents are to be preserved. Let us consider the situation, or, rather, the archive document itself, more closely. The definition of the archive document has been fixed in legislation as 'a material medium which carries information recorded on it, has requisites allowing to identify it, and is subject to preservation because of the value of the medium and the information on it for the public, society, and state' (The Federal Law 'On Archiving in the Russian Federation,' Art. 3). In the dual nature of the document – the medium and the information, the primacy of the medium which is mentioned twice in the above definition, stands out: it is the key element of the definition, and the identifying requisites also refer to the medium. Strictly speaking, the properties of the medium may help identify the document (for example, the study of the dryer fabric nature when making paper, X-raying of palimpsest), but these cases are scarce.

The existing standard of 1998 entitled 'Record Management and Archiving. Terms and Definitions' offers a somewhat different interpretation of the document, defining it as: «Document; documented information: information which is recorded on a material medium and contains requisites allowing for identifying it» (State Standard RF GOST P 51141-98, Art. 2.1). In this definition, the main emphasis lies on the information, both by the place given to it in the definition and by the affiliation of the requisites.

A short background study reveals that this disagreement in definitions is not occasional. Our first domestic standard of 1970 (GOST 16487-70. Record Management and Archiving. Terms and Definitions) specifies that the document is "a means of fixing information on facts, events and phenomena of the objective reality and intellectual activity of humans on a special material medium". Since this definition was obviously inappropriate, only 13 years had passed before a new, «materialistic», wording appeared (GOST 16487-83. Record Management and Archiving. Terms and Definitions): «the document is a material object with information which is recorded on it using a man-made process for the purpose of transferring it through time and space».

In the end, the «information» approach, which recognizes the primacy of information, took over 15 years later in GOST 1998.

Even the Law on Archiving did not avoid these fluctuations. The draft Law on Archiving of the Russian Federation published in November 2003, which was submitted to the State Duma by Government Edict No. 1743-p of November 29, 2003 (already updated in accordance with recommendations made by the Institute of Legislation and Comparative Law of the Government of the Russian Federation and the Private Law Research Center State Institute under the President of the Russian Federation) reads that «the archive document is information accompanied by requisites, which allow for its identification; this information is recorded on a material medium and is subject to preservation with due account for its value to the public, society and state». In the end, the State Duma decided that the document is a medium with information and that it is the medium that should be identified.

Putting aside the subtle matters of correlation between the value of the item and its place in the unnumbered list, we cannot but observe that, since the wording offered by the Law covers the notion of the electronic document (Art. 5), it implies the identification of the electronic document by the requisites of its medium. *Dura lex*...

Let us also observe that the Fundamentals of the Archive Fund and Archives of the Russian Federation (1993), which were replaced by the Federal Archiving Law in 2004, avoided the issue of the document proper keeping themselves to the following definition in the glossary: "The archive document is a document preserved or subject for preservation due to its value to the society and the owner."

The above variations and fluctuations of the wording confirm the dual nature of archive documents. And it is their digitization for the purpose of preservation of cultural heritage that unambiguously reveals this duality.

There is yet another reason for the above study of definitions. The created digital heritage must have an officially recognized and legitimate status that confirms the authenticity of both the digital copy of informational (visualized) component of the original and the reliability of descriptive metadata. The above formulations obviously cannot provide a solid ground for digital cultural heritage legitimization.

So let us return to the question: what is preserved as a result of digitizing traditional archive documents? Both components of the document are preserved, i.e. the informational one and the material one, but they are preserved differently, in different places and conditions.

The informational component of the documents is preserved directly in the digital image of this document. The original archive document itself is withdrawn from everyday use and, thus, preserved. As a result, the original archive document becomes a museum item, and its informational cousin begins living in the archive's local network and on the Internet.

This explains the archive development trend: archives will gradually develop into museum storages, while digital copies of archive documents, brought together in an integrated automated system, will be in active use on-line.

In the recent years, national archives, both federal and regional, have been digitizing their documents building on two main principles: digitization within separate projects or as a result of creating document copies upon one-off user requests (more and more often users prefer files to photocopies). Though the latter trend has become quite popular, this flow is insignificant, and fragmentary copying cannot be regarded as a serious factor of preserving archive documents.

It is clear that considerable, even huge resources, are needed for mass and target-oriented digitization projects. For example, 75 million Euro are required to digitize 1% of documents of the State Archives of Germany (Bundesarchiv) including input of metadata necessary for orientation in digital materials. Digitization costs of Russian archives are, at least, comparable to those of European archives, and the mass of archive collections is so immense that even the national budget would not hypothetically cover the digitization of the entire lot.

In these conditions, the sound management of funds obtained by archives for digitization of their archive documents seems highly important and pragmatic.

In other words, a detailed elaboration of digitization projects of archive documents is meant.

Any project starts with the selection of documents for digitization. Since this stage is subjective, it is useful to consider whose and what kind of interests may influence the choice of documents for scanning. At a first approximation, we can single out three groups of actors or, in other words, three groups of motivated participants in the decision-making process.

The first group is archivists. They are interested, primarily and directly, in ensuring the preservation of documents, which implies the scanning of unique and especially valuable documents (as a means to prevent theft of originals and a guarantee against damage); materials in a poor physical condition; documents in great user demand, i.e. documents whose condition deteriorates as a result of frequent use. Moreover, in the conditions when archives cannot recruit enough staff for their low-paid positions of keepers – employees responsible for carrying documents from storage rooms to reading rooms and back, the development of digitized documents helps reduce demand in keepers. The latter though influences only general interest in scanning, not to the selection of documents process.

The second group comprises users. It is miscellaneous and has rather wide and ambiguous needs. We should admit that users in reality have no ways of influencing the selection of documents for digitization because they do not represent an organized team capable of articulating and lobbying its demands.

Finally, the third group includes the administration of archives, i.e. officers who can (and must) engage in fundraising for digitization purposes. Their motivation in the selection of documents is mainly subject-oriented. In fact, widely popularized projects, projects with high sociopolitical profile, have more chances to secure funding; this means that results of the digitization must be accessible on the Internet. In such projects, the openness of archive documents and a broader access to them are of key importance, the preservation proper representing a secondary, though indispensable element.

Interests, motives and weight categories of these groups need to be examined separately.

The selection of documents must be followed by the project planning stage; this takes account of technological processes, metadata composition, specially trained staff, time costs, equipment used and software required. Unfortunately, a somewhat solid methodological basis which could be used for such planning is still lacking.

We can start with the issue that seems simple: the cost of digitization of one page from an archive file within mass digitization. No such data on Russian archives have been published.

An approximate and rough assessment may be found in the pricelists of services offered by archives. The resulting picture is rather versatile. The comparison data on digitization costs (in Rubles) of one A4 page without graphics processing are listed below.

In federal archives:

The State Archive of the Russian Federation, Russian State Archive of the Naval Forces – 100;

Russian State Archive of the Contemporary History – 86;

Russian State Archive of the Sociopolitical History – 60;

Russian State Archive of Literature and Art – 50-500;

Russian State Military Archive – 50-100;

Russian State Archive of Economics – 20-150;

Russian State Archive of Scientific and Technical Documents (Moscow) and its Samara branch – 550.

In the state and municipal archives:

The State Archive of Perm Oblast - 88;

The State Archive of Kostroma Oblast - 39;

The National Archive of the Tatar Republic - 75;

The State Archive of Murmansk Oblast - 73;

The State Archive of the Modern History of Perm Krai - 28;

The State Archive of Perm Krai – 88;

Komi-Permyak Regional State Artchive – 20;

The State Archive of Vladimir Oblast. – 27;

The State Archive of Tver Oblast – 34;

The State Archive of Moscow Oblast – 25;

The State Archive of Primorsky Krai – 21;

The City Archive of Perm – 258;

The State Archive of Voronezh Oblast – 34;

The State Archive of Chelyabinsk Oblast – 176.

In conclusion of this list we will bring here the cost of digitization of one page in the Samara Laboratory for Archive Document Preservation: 9.2 rubles for organizations and physical persons alike. However, this figure should not be confusing: it does not include the description of electronic images (input of metadata), integration of images into the automated system, transfer of documents from the storage room, as well as binding and unbinding of files (if applicable). The cost of metadata input alone can exceed many times that of scanning of archive documents.

The above figures allow for numerous "guess-how-much" interpretations, but if we put aside the ungrounded suspicions of greediness or altruism, we have to admit that this versatility of prices might be caused by difference in digitization and metadata input procedures, which can hardly be explained by local peculiarities.

The real costs of mass digitization of archive documents within various big projects should be, presumably, more homogeneous, but we have no grounds for objective judgment. The cost of scanning and metadata input stipulated in an agreement between an archive and a financing institution, is a commercial secret, let alone the archive's factual expenses on scanning and input of descriptions...

In fact, archivists need a basis of calculations, not certain figures, to design their projects; i.e. they need a detailed step-by-step description of the optimal planning and implementation of mass digitization projects.

Unfortunately, at present, there are no methodological guides to be offered to Russian archivists. Scarce handbooks are very specific (for example, Recommendations for scanning archive documents of the Russian State Archive of Scientific and Technical Documents, which have been developed for the departments of this archive only).

To enable an effective digitization, it would be very useful to highlight the following key points:

- Resolution and image chromaticity, with account of the purpose of the material digitized and the type of the document scanned.

- Correlation of images with software which contains both images and metadata. According to scarce descriptions of digitization projects of archive documents, there may be cases when descriptions are input into the database, then the documents are scanned, and only afterwards integrated into an integrated resource. However, software allows to send images of documents directly from the scanner to the target cell in the database and immediately links the image with its description.
- Selection of scanning equipment. The best option would be regular and comprehensive monitoring of equipment coupled with the assessment of some parameters, namely, cost-effectiveness, which may be achieved with the help of expensive book scanners with a V-type cradle that allow to exclude unbinding and binding of files thanks to the application of, or an in-built graphic software capable of correcting the image on a concave page. (It is difficult to assess the physical damage inflicted on an archive document by photostress when it is rolled through a flatbed scanner, unless by calculating the costs of physical and computer restoration).
- Metadata composition and formats of scanned images which ensure a satisfactory description of archive documents, accessibility for the Internet search engines, interaction with similar systems.

Another challenge should be highlighted, whose solution is crucial for the development of cultural heritage digitization projects. Work is underway on elaborating online systems into which archive and book resources will be integrated. Finding solutions to the issues of metadata composition and formats within these systems is hampered by many obstacles and painful moments, because libraries and archives have their own traditional methods of description enshrined in databases. It is quite obvious that the next step in creating such systems will be inclusion of museum resources into them; in this case the issue of metadata will become an even more relevant problem.

These brief comments highlight some of the key issues which determine, to a large extent, the successful preservation of archive cultural heritage. We hope that the UNESCO Information for All Programme will contribute to an active and professional search of solutions to these issues and the beneficial exchange of best practices.

Hayk Ojen KHANJYAN
Yerevan, Armenia
President
"Bi Line" Ltd
Head of the Culture and Language Division
Information Technologies Development Support Council
of the Republic of Armenia

Proposals to Develop a Multilevel Model of Digital Content Archiving

"...disappearance of any element of the cultural heritage in whatever form constitutes an impoverishment of the heritage of all nations of the world"

UNESCO's General Conference Charter on the Preservation
of the Digital Heritage

The exceptional importance of the UNESCO Information for All Programme and the fundamental principles stated in the UNESCO Charter on the Preservation of the Digital Heritage (Paris, 2003) for the development of knowledge-based information society, for the preservation of the rich multinational cultural, scientific, and the whole diversified human heritage and passing it over to descendants is unprecedented.

This program is particularly valuable for the Armenian people who make one of the most ancient nations with the history dating back to the 4th millennium B.C., with the original alphabet developed in 405 A.D. which is still in use, with its ancient handwritten heritage (over 30,000 manuscripts, 15,000 of them outside Armenia) beginning with the 5th century till the 18th century, early printed books (over 1,200 volumes, 1.5 million copies) starting from 1512 up to 1800. To celebrate the 500th anniversary of book publishing, UNESCO declared Yerevan Book Capital 2012.

All the above mentioned is probably to be referred to as universal human heritage which has to be preserved for many centuries and made accessible to any interested user in digital format.

Unfortunately, due to certain external and internal reasons, digitization of the Armenian heritage which is acknowledged to be of value is not a top national priority, and no national programs embrace digitization as an instrument

of long-term preservation and support of universal accessibility, though we are aware of the experience of the United Kingdom, Russia, New Zealand, Australia, Poland, etc., where these issues are of national importance.

Still, there are certain examples of cultural and literary heritage digitization in Armenia, but it is spontaneous and accomplished as separate undertakings, or with external support. In particular:

- Digitization of the ancient manuscripts of Matenadaran Museum has been initiated (approx. 1,000 manuscripts digitized, unfortunately, below high contemporary standards), with the Museum's collection comprising over 18,000 manuscripts dating back to the 5th century A.D., 15,000 of them of Armenian origin and 3,000 – of Chinese, Persian, Arabic, Jewish, Indian, etc., origin (www.matenadaran.am),
- A group of Armenian experts supported by the American University of Armenia carries out a rather successful project of Armenian literary heritage digitization – www.digilib.am,
- Another project is Eastern Armenian National Corpus being accomplished by Moscow Russian Language Institute with the support of several institutes in Yerevan www.eanc.net; it is aimed at preserving the Western-Armenian language which is included on the list of 100 world disappearing languages even though it is spoken by 5 mln out of 7 mln Armenians living in many countries outside Armenia,
- 1,000 of 15,000 ancient Armenian manuscripts in the British National Library, Bibliothéque Nationale de France, libraries in Poland, Germany, Vienna, etc., have been digitized automatically.

On the one hand, the abovementioned solutions are individual examples of preservation of the Armenian heritage, but on the other hand, they reveal certain general problems to be considered while developing strategies, policies, procedures, standards, and selection criteria for long-term preservation of heritage materials (further referred to as materials) that need to be digitized, let us call them "tangible materials" (handwritten, printed, graphic, video, architectural monuments, etc.). On the contrary, we shall call born digital materials "virtual materials".

Acknowledging and accepting the priority and significance of the problems and ways of their step-by-step discussion and solution as they are stated in the Strategic Plan for 2008-2013 developed by the Bureau of the Intergovernmental Council for the Information for All Programme, let me emphasize and draw your special attention to the following facts:

- "Tangible" materials, regardless of their origin and value, are often stored outside the countries' territory or ethnic and national communities.
- Many "tangible" materials with their definite historical, scientific, religious and other value for one state and/or nation, are created by representatives of another state or transnational company, and probably published in a third state.
- Digitization of "tangible" materials often remains unknown to the country of their origin and their ethnic background. In particular, HathiTrust Digital Library has digitized over 9.7 mln items of almost every country and nation, for instance, 224,836 Russian and 8,641 Armenian items, though free access is provided only to 250 records.
- "Virtual" materials may have certain value for a country and/or nation; however, these materials are often published and preserved without reference to their geographic and national origin.
- The protection of publishers' rights and copyright is another even more difficult problem, in particular for virtual materials. It is unfeasible to expect that creators or cessionaries (countries, publishers, others) will track and secure their rights to publication and digitization in the nearest future. It takes quite a lot of time not only to develop new international rules and provisions for the protection of copyright and right of publication on the Internet along with the software tools to identify the cases of breach of these rights, but also to instill discipline and responsibility to many millions of creators and publishers.

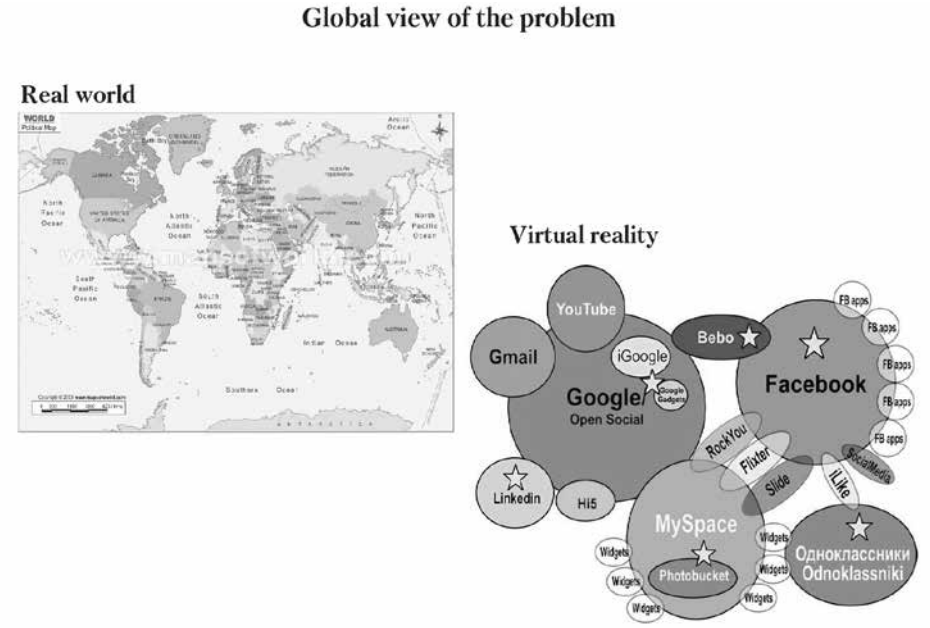
The copyright and identity of both "tangible" and "virtual" materials are among the main problems in terms of responsibility for digital heritage preservation and future access to it but also in terms of the possibility of emergence of disputable situations, including at the interstate level. The table below presents our assessment of the current situation.

From particulars to generals – Definitions

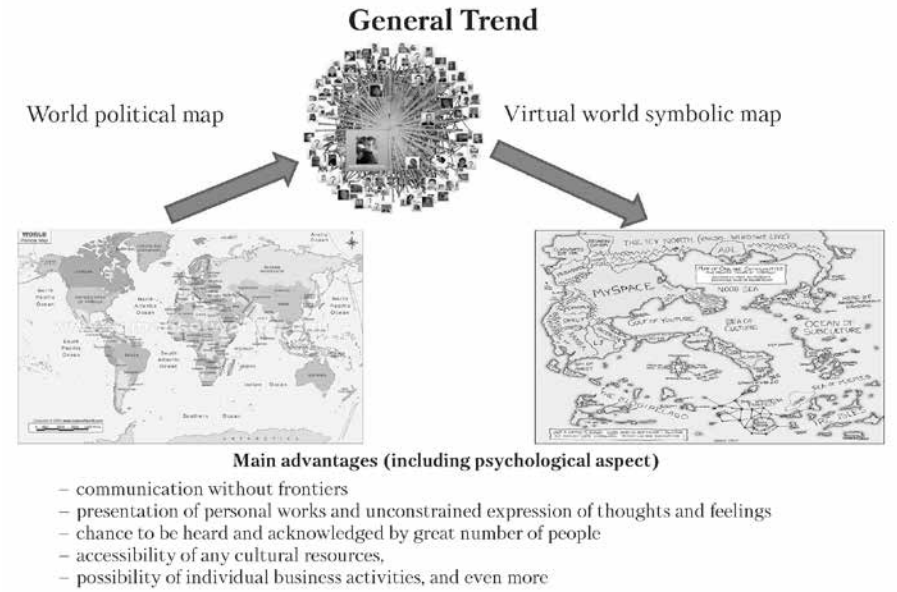
- "**tangible**" materials – manuscripts, printed, painted, graphic, video, architectural monuments and other materials to be digitized
- "**virtual**" materials – born digital materials
- *categories related both to "tangible" and "virtual" materials*

Tangible and virtual materials	Estimation of tangible materials status	Estimation of virtual materials status
Nationality and/or ethnic origin	Mainly defined	Mainly undefined
Creator – Authorship (unavailable in some cases)	Mainly defined	In some cases not specified
Publisher	Mainly defined	Overlooked
Location – storage	Mainly defined	Overlooked
Value (human, regional, ethnic, etc.)	Partly acknowledged	Overlooked
Type (manuscripts, printed m., paintings, photos, music records, architectural monuments, etc.)	Known	Mainly known
Digitization status	Few	
Who digitized	Mainly known	Mainly known
Digital copy owners	Disputable	Disputable
Other		

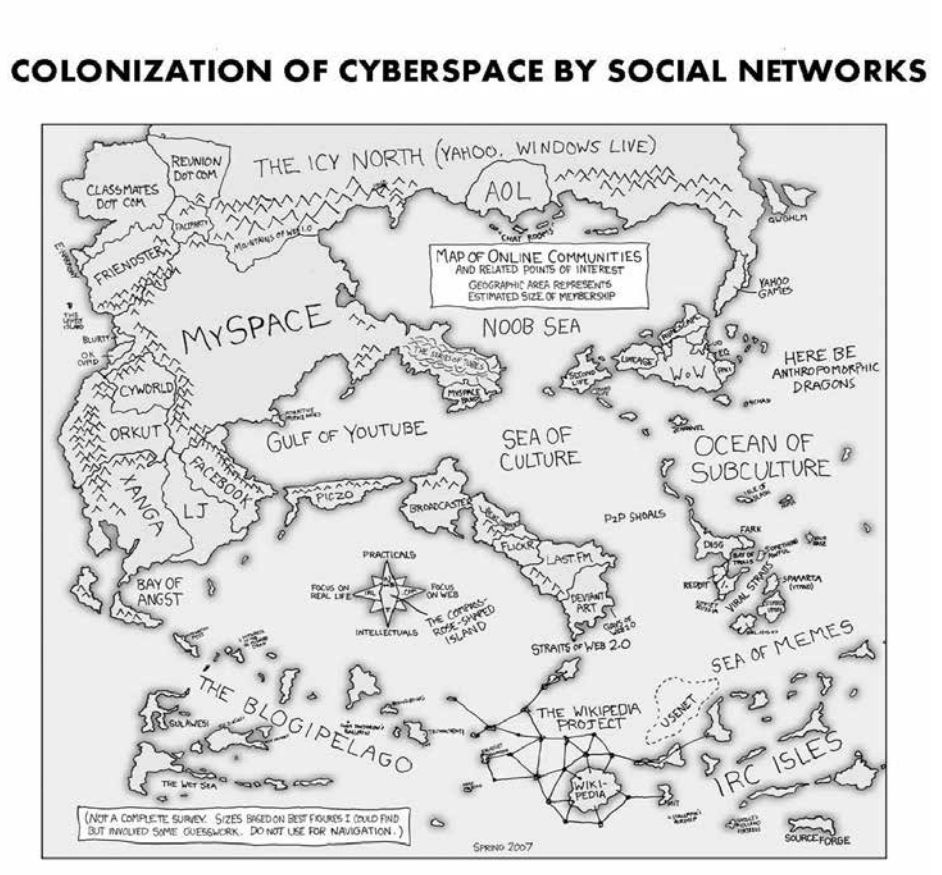
Globally, in order to solve most of the above-mentioned problems it is necessary to understand our new current reality – along with the real world, a more advanced and "aggressive" virtual reality has taken shape in the recent decades. This reality has to be taken into consideration and the infinite possibilities that it offers have to be used.



Due to the virtual world advantages, shown in the slide below, the general trend in the real world results increasingly in people's "migration" towards Virtual Reality in the Cyberspace.



Meanwhile, well-known social networks have been increasing their influence on the processes of reality virtualization and cyberspace "colonization".



Several important conclusions have to be made:

- Nothing like a "digital heritage preservation network" to balance commercial social networks has been proposed or initiated;
- Tangible and virtual materials are scattered across different states, private organizations and on the World Wide Web, which significantly impedes their search, preservation, digitization, etc. by governmental agencies, libraries, and institutions that are in charge of heritage preservation;

- A considerable number of UNESCO member states, as well as non-members, will not be able (due to external and internal reasons, i.e. economic, technological, technical, etc.) to assume full responsibility for the preservation of "their" national heritage which is, to a certain extent, the universal human heritage;
- Development of strategies, principles, policies, procedures, standards and selection criteria to provide long-term preservation of materials possessing universal, national, ethnic, etc., value requires considerable effort. We do not have much time for their approval and adoption by the UN, intergovernmental, governmental and nongovernmental organizations.

Therefore, when developing strategies, principles, policies, procedures, standards and selection criteria, it is advisable to consider the following proposals:

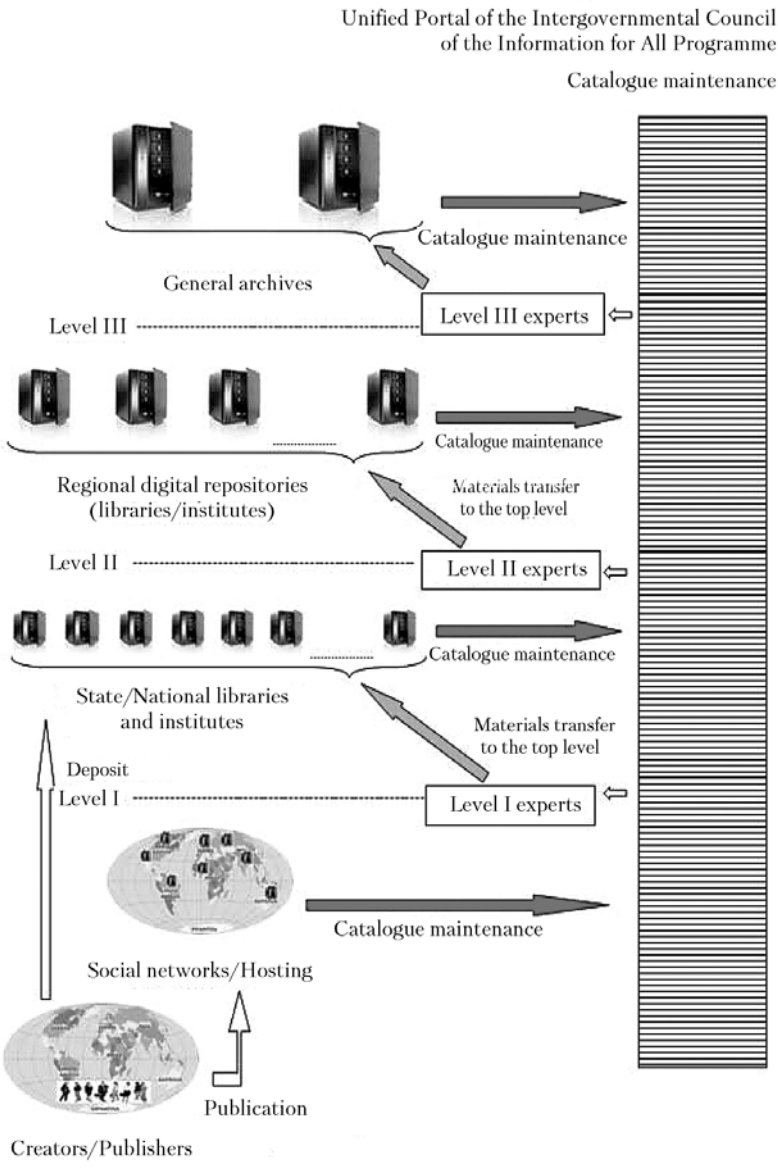
1. To develop a system of multifaceted classification for digital heritage materials to ensure a uniform material indexing based on:
    - value (universal, regional, ethnic),
    - affiliation (national and/or ethnic, private, etc.),
    - purpose (culture, arts, religion, history, policy, etc.),
    - creation date,
    - type of primary source (manuscript, publication, document, graphic image, video, digital, etc.)
    - organization that accomplished digitization,
    - digitization quality,
    - storage location, etc.
1. To establish **multilevel** system of expert groups – interstate, regional, national, ethnic – to consider and assign one or more value statuses to born-digital and digitized materials;
2. To develop a distributed multilevel archiving network on the basis of state and/or national libraries and archives, a number of specialized research institutes, privately-owned and global social networks and commercial hostings;

3. To introduce reliability appraisal and minimum standards of preservation length (short-term, medium-term, and long-term) for every archiving level;
4. At the intergovernmental level, to design a unified portal, for example, within the framework of the Intergovernmental Council for the Information for All Programme, where every organization belonging to the multilevel archiving network shall catalogue its digital materials in accordance with the accepted classification system;
5. To propose to establish a fund affiliated to UNESCO and/or Information for All Programme Intergovernmental Council to provide financial, consultative, technical and technological support to the states that are unable to accomplish digital preservation programme on their own;
6. To attempt to establish a social network to engage interested individuals and legal entities in building archives of cultural and other heritage.

Recognizing that the complicated task of digital content preservation has to be dealt with today (it had to be dealt with a long time ago, in fact), we propose a multilevel model for digital heritage archiving; according to the model, the initial (short-term) preservation is to be imposed on hosting companies, i.e. the participants of information exchange who actually acquire, store and, in many cases, publish content materials irrespective of their purpose, value, or affiliation.

The main idea of the model shown in Picture 1 is to build a multilevel model for archiving global digital heritage regardless of geography, language and origin, and a model for a step-wise selection of materials depending on their value – universal, national, ethnic, etc., with further storage in more compact volumes and for definite longer periods of time in relevant institutional repositories of higher level.

MULTILEVEL DIGITAL HERITAGE ARCHIVING MODEL



Pictured by Hayk Khanjyan, Merujan Karapetyan, Oganes Kizogyan, Ruben Tarumya

Section 2. Digital Preservation of Audiovisual Documents

Dietrich SCHÜLLER
Vienna, Austria
Vice Chair
Intergovernmental Council of the UNESCO Information for All Programme
Consultant
Vienna Phonogrammarchive of the Austrian Academy of Sciences

The Role of Digitization in the Preservation of Audiovisual Documents

This paper is an updated and expanded version of "Socio-technical and Socio-cultural Challenges of Audio and Video Preservation", a paper originally presented at the 3rd Memory of the World Conference in Canberra 2008, published in International Preservation News 46, 2008. Still concentrating on audio and video documents, it also includes the latest development in film archiving.

Audio and video documents are the most significant primary sources of linguistic and cultural diversity. With all respect to the role of language and written texts in human communication, the limits of these traditional tools to communicate and describe cultural phenomena are obvious and undisputed. It must be noted that scientific interest was the driving force for the invention of audiovisual recording technology: the study of language and the human voice paved the way for the invention of sound recording while the interest to analyse fast movements, which could not be explored by the naked eye, triggered the invention of cinematography. Several disciplines like linguistics, ethnomusicology and parts of anthropology did not really flourish until the advent of audiovisual documents which - more or less perfectly and more or less objectively - permit the creation of adequate primary sources of or about the phenomena of interests themselves: language, music and dance, rituals, artefacts etc. Consequently, it was the academic world that installed the first sound archives, 1899 in Vienna, 1900 in Paris and Berlin, 1908 in St. Petersburg.

Commercial exploitation, though not at the cradle of the new recording technologies, started even before 1900: the products of the phonographic and film industries soon quantitatively surpassed the academic activities. It is noteworthy, however, that systematic collection and archives for the products of the entertainment industry emerged only by and by in the 1920s and 1930s, as libraries and archives started to include audiovisual materials in their collections. In those years, independent units in form of sound archives were created (e.g. the Discoteca di Stato in Italy or the French Phonothèque National), while film archives were founded in the Netherlands, the UK, the Soviet Union, France, and Germany. As Radio Broadcasters also developed from 1922 onwards, radio sound archives came into existence.

The consolidation of audiovisual archiving only happened since the later 1940s, heavily supported by the flourishing phonographic industry, specifically after the introduction of the microgroove ("vinyl") disc. The other push came from the international spread of magnetic audio tape recording, in Germany already in use since the mid-1930s, which after World War II was immediately widely adopted for the production of unique documentary and cultural recordings. From 1956 onwards, magnetic video recording became available and gradually replaced film recording in television stations. Outside broadcasters in the academic world, magnetic audio tape recording, specifically the availability of battery-operated portable equipment, enormously furthered the production of research materials, as it became possible to record language, music and rituals everywhere in the world in good quality. This also created the corpora that constitute the primary source materials of our present-day knowledge of the linguistic and cultural diversity of mankind. While the creation of film documents for research was not very widespread because of the considerable costs involved in the production and development of film, moving image documentation for scholarly and cultural purposes mushroomed, since in the 1980s true portable video recorders became available which permitted the creation of video documents in a fashion similar to what had already existed over the past decades for audio.

These three creative sectors in audiovisual production - the record and film industry, the radio and televisions broadcasters, and the academic and cultural bodies - have accumulated a remarkable legacy of primary source materials, which form the most significant sources of cultural and linguistic diversity of mankind. They are partly artistic creations in their own right,

like films and music productions, and partly documents of political, historical and cultural events and phenomena. Most justifiably, audiovisual documents have been called the media of the modernity: no adequate understanding of the past 100 years would ever be possible without them.

The long-term availability of this important group of documents is, generally, more problematic than that of traditional text documents:

Historical cylinders become brittle and mouldy, unique instantaneous disks deteriorate beyond retrievability, life expectancy of magnetic tape can be assumed to be only in the order of decades[67], and recordable optical disks must be considered to be at great risks, unless produced under tight quality control, which practically can hardly be met. The long-term preservation of original photographic materials is not so much of a problem, provided stringent storage and handling conditions are in place.

Carrier instability, however, is only part of the problem. As machine-readable documents, all audio and video recordings depend on the availability of format-specific replay equipment, some of considerable sophistication. Because of the technical development over the past 20 years, we have experienced ever shorter commercial life cycles of dedicated audio and video tape formats. Whenever a format had been superseded by the next, industry swiftly ceased production of equipment, spare parts, and professional service support for the older format.

Around 1990, this foreseeable development led to a shift of paradigm amongst sound archivists: it was realised that the classic aim to preserve the document placed in the archives' care would ultimately be in vain, because even if carefully kept carriers survived over longer periods, the unavailability of replay equipment would soon make these stocks irretrievable, and thus useless. Audio preservation has to concentrate on the safeguarding of the content, not of the original carriers, by copying contents losslessly from one digital preservation platform to the next. Analogue contents have to be digitised first.

[67] A specific problem in the Russian Federation and in former Eastern Bloc countries is the instability of acetate cellulose tape, which has been used in great quantities from the 1950s to the early 1970s. Such tapes become brittle with age, which makes their reproduction difficult, often impossible. However, most recently a method has been developed to re-plastify these tapes (Wallaszkovits 2011).

This new paradigm was met with some scepticism from traditionally-minded archivists. However, German radio broadcasters took the lead to develop digital mass storage systems, which soon became state of the art in audio archiving. The incentive for their installations was not so much preservation, but automated access to huge archival holdings, which was considered to become a strong weapon in the fight of these previously monopolistic institutions against upcoming competition from private broadcasters. Video archiving is following that path, with some time delay however, as storage quantities for video are significantly higher than for audio. Outside the broadcaster's world, national archives and libraries, but also research archives are also following.

This development has also been furthered because, faster than originally anticipated, dedicated formats have been gradually given up in favour of true computer file formats, audio fist, followed meanwhile by video. Thus, recording, postproduction and archiving became part of the IT world. Today, there is unanimous agreement that the time window left to keep machines in operating condition for all pre-IT, single carrier based audiovisual, specifically the tape based formats, is not more that 15 years, if at all.

Most recently, this change of paradigm has also happened for film preservation. Here it is not so much the problem of fading reproduction equipment, but simply the foreseeable termination of chemo-optical film production as a consequence of digital film projection, which is rapidly replacing classical film projection in cinemas worldwide. Also, film production is increasingly changing from traditional to electronic recording. Because digital film preservation is in demand of even higher storage quantities than video, it is obvious that film is the last of the audiovisual media to enter the digital domain.

Concentrating now on audio and video recordings, the worldwide holdings are estimated to be 100 million hours for each of the two categories.

Feeding analogue and single digital carriers into digital repositories is a demanding and time-consuming process. Principles for audio documents have been standardised by the International Association of Sound and Audiovisual Archives (IASA), which had also issued practical guidelines for the production and preservation of digital audio objects. The respective guidelines for video are presently under preparation. The transfer of originals is in need of modern replay equipment, of test equipment and expertise for their proper maintenance. The time one operator needs for

transfer must be estimated to be at least triple the duration of audio, and even significantly more in the case of video documents. Bigger radio and national archives are solving that problem by simultaneously transferring three or four audio tapes at one time, making use of special quality control software to replace the aural control of the operator. This works well with fairly homogeneous source material as typically available in radio archives. Holdings of research materials, because of their diverse technical nature, hardly lend themselves to this kind of "factory transfer".

Yet there is more to it than solving the transfer of originals. Digital preservation is equally demanding, as it requires an ongoing investment to keeping digital data actively alive. Appropriate professional storage technology and management software is expensive and needs subsequent renewal at least in the pace of migration intervals, which are generally in the order of five years. It must be clearly stated that the use of recordable optical disks as sole digital target media constitutes a great risk, although it is unfortunately still widespread, specifically amongst small and less wealthy institutions. Professional digital preservation costs have steeply come down over the recent years and have reached approximately 1USD/GB/year[68].

It can be assumed that the challenges as outlined above will be met by the radio and television archives as well as the national collections of fairly wealthy countries within the next 15 years. Because of the impending unavailability of replay equipment, this is the time window generally considered to be available for safeguarding what we have accumulated so far. Several developing countries, however, will face considerable problems in safeguarding their holdings, even in a selective manner. The most significant problem is lack of funds. While it is fairly popular to finance digitisation projects in the course of international development cooperation, the lack of commitment to finance long-term preservation of the digital files makes many such projects a dead end road.

The great majority of small and hidden collections in all parts of the world, which preserve a considerable part of the world resources of cultural and linguistic diversity, have a different, generally much greater problem. The

[68] In 2007, when the actual costs still were in the order of 5-8USD/GB/year, the IFAP Working Group on Information Preservation targeted this 1USD level, which today applies for great digital storage units in the order of 100 Terabytes (TB) and beyond. This cost development is a result of the tumbling prices of hardware and storage media. Because of the stable, if not increasing costs for building, energy, and personnel, however, this slope of the prices will to some extent flatten out. On the general problems of digital preservation, see specifically the contributions throughout these proceedings

first is awareness. While generally "digitisation" is recognised to be an action to be carried out, there is little knowledge about prerequisites needed and standards to adhere to. Most typically, inadequate replay equipment is regarded to be sufficient, and there is no realistic perspective about standards and costs to preserve the digitised documents. The other notorious problem is lack of money, which mainly - apart, of course, from unfavourable general economic situations - means lack of awareness on the part of parent organisations, governing financing bodies, and/or of the public at large.

From the technical prerequisites, the required expertise, and the necessary financial resources it becomes clear that autonomous audio and video preservation requires critical masses amounting to several thousands of carriers within each format. As many important collections are held by relatively small institutions, many even still at the private homes of the researchers that had recorded them, the only viable solutions for these holdings are cooperative projects, which can be arranged in different forms: the transfer of original contents to digital files should thereby be seen separately from digital preservation

First and foremost, such cooperation should be established on a national basis, whereby audiovisual archives should avoid competition by specialisation in complementary sectors. Globalisation, however, increasingly creates international cooperation amongst institutions, including systematic training of audiovisual archive staff under the auspices of the international audiovisual archives associations. Cooperation with commercial service providers is also an ever increasing factor[69].

At present, systematic film digitisation is still in its infancy. As film archiving is generally in the hands of well organised film archives, film digitisation will happen centrally in a professional environment.

In summarising, it can be stated that the greater part of audio and video collections, held by the broadcast and national archives of wealthy countries, will be safeguarded and made available in the long term. Whether also in developing countries these kinds of institutions will be able to solve their problems within the time window of the next 15 years, remains open. Much will depend on the political will of these countries to safeguard their audiovisual cultural heritage and to allocate the necessary funds. There is some reason to hope that the development in this field in the West over the

[69] For detailed discussion see IASA-TC 04, chapter 9. Presently, European Commission funded project PrestoCentre is establishing a network of European competence centres for audiovisual preservation http://www.prestocentre.org/.

past 20 years can be optimistically extrapolated to other parts of the world. This, however, concerns only part of the accumulated heritage.

In terms of importance, a major part of the entire audiovisual heritage is still held in small and scattered, often hidden and notoriously underfunded research and cultural collections all over the world[70]. Without them, our view of the cultural and linguistic diversity of mankind would be incomplete. Their loss would mean a substantial deprivation of cultural, linguistic, and ethnic minorities in terms of their heritage, their history and their identity. The veritable challenge of a worldwide strategy of audiovisual preservation is to spot these collections and to organise their physical survival.

References

1. Boston, George (Ed.): Safeguarding the Documentary Heritage: A Guide to Standards, Recommended Practices and Reference Literature Related to the Preservation of Documents of All Kinds. UNESCO, Paris 1998 http://www.unesco.org/webworld/mdm/administ/en/guide/guidetoc.htm. Extended CD-ROM version UNESCO, Paris 2000 (available from a.abid@uneso.org).
2. Bradley, Kevin: Risks Associated with the Use of Recordable CDs and DVDs as Reliable Storage Media in Archival Collections - Strategies and Alternatives. UNESCO, Paris 2006. http://www.unesco.org/webworld/risk.
3. Bradley, Kevin: Towards an Open Source Repository and Preservation System. Recommendations on the Implementation of an Open Source Digital Archival and Preservation System and on Related Software Development. UNESCO, Paris 2007. http://www.unesco.org/webworld/en/mow-open-source.
4. Edmondson, Ray: Audiovisual Archiving: Philosophy and Principles. UNESCO, Paris 2004. (CI/2004/WS/2). http://unesdoc.unesco.org/images/0013/001364/136477e.pdf.
5. Henriksson, Juha, und Nadja Wallaszkovits: Audio Tape Digitisation Workflow. Digitisation workflow for analogue open reel tapes (2008). http://www.jazzpoparkisto.net/audio/.

[70] It is estimated that 80% of the total holdings worldwide are still outside professional archival care.

6. IASA Technical Committee: The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy, edited by Dietrich Schüller. (= IASA Technical Committee - Standards, Recommended Practices and Strategies, IASA TC-03), Version 3, 2005. http://www.iasa-web.org/de/IASA_TC03/TC03_English.pdf. Also available in German, French, Swedish; Italian, Spanish, Russian, and Chinese.

7. IASA Technical Committee: Guidelines on the Production and Preservation of Digital Audio Objects, edited by Kevin Bradley. (= IASA Technical Committee - Standards, Recommended Practices and Strategies, IASA TC-04), second edition 2009. http://www.iasa-web.org/de/tc04/audio-preservation.

8. Schüller, Dietrich: Preservation of Audio and Video Materials in Tropical Countries. International Preservation News 54, 2011, 31-34. http://www.ifla.org/files/pac/ipn/IPN 54def.pdf.

9. Schüller. Dietrich : Audiovisual Research Archives. Report WP 6 of EU-funded Project TAPE. ECPA 2008. http://www.tape-online.net/docs/audiovisual_research_collections.pdf.

Nadezhda BRAKKER
Moscow, Russian Federation
Senior Expert
Centre of Informatization in the Sphere of Culture (Centre PIC)

Leonid KUIBYSHEV
Moscow, Russian Federation
Director General
Centre of Informatization in the Sphere of Culture (Centre PIC)

Preservation of Digital Information in Cultural Institutions: Tasks and Barriers

Introduction

The electronic collections of the Russian depositary libraries and archives have been built in compliance with the legal deposit copy law. Federal Law No. 77 On the Legal Deposit Copy serves as a basis for the national library and information collection of the Russian Federation.

In addition to the legal deposit copies, digital collections of depositary and other libraries are built up in two ways: (1) externally, through standard acquisition channels (purchase, donation, exchange); and (2) through digitization of traditional collections.

In this paper, the authors will dwell upon the compilation of non-depositary collections and long-term preservation of digital information resources in Russian libraries, archives and museums (hereinafter, cultural institutions).

What Should We Preserve?

Cultural institutions should preserve documents held in their collections, including digital documents. This responsibility of cultural institutions has been stipulated in the national law and ministerial decrees and instructions. The collections of cultural institutions include audiovisual documents, electronic publications on external media and born-digital

documents, all of which collected according to the acquisition profile of libraries, museums or archives and accepted for preservation according to the applicable regulations.

Additionally, cultural institutions should ensure preservation of digital documents and information resources that are not part of formal collections. The government is interested in building digital information resources of cultural institutions and should be interested in their long-term preservation. Thus, according to the Strategy for the Development of Information Society in the Russian Federation, 100% of library catalogues should be digitized until 2015. The information resources of cultural institutions are accumulated mainly through federal and local funding and are of a lasting value.

This category includes:

Digitization Results

For the last 10-15 years, Russian cultural institutions have been actively engaged in digitization. However, information on the state of preservation, the amount of digitized documents in relation to the total volume of collections, and the total number of digitized items has not been either collected or studied. Procedures or terms of digital preservation in cultural institutions have not been developed, either. Neither have been determined the status or the accountability methods for digitized documents in cultural institutions.

It is still unclear whether they form part of the collections; whether they shall be accounted for separately from their traditional original or together with it; how account should be taken of documents which are the result of digitization of the same original but are presented in different formats.

Information Systems

In the process of informatization, cultural institutions create significant reference and bibliographic databases, e-libraries and e-catalogues, which tend to grow steadily.

Table 1. Electronic bibliographic databases of cultural institutions

	2008		2009		2010	
	Electronic bibliographic databases of Cultural Institutions, thou.entries	**Inclusive of the electronic catalogue**	Electronic bibliographic databases of Cultural Institutions, thou.entries	**Inclusive of the electronic catalogue**	Electronic bibliographic databases of Cultural Institutions, thou.entries	**Inclusive of the electronic catalogue**
Federal libraries	14072	12747	15754	13348	19524	16808
Generally accessible local libraries of the Russian Ministry of Culture	111916	57416	129673	66978	141164	73444
Total	125988	70163	145427	80326	160688	90252

Collections of cultural institutions also include complex multimedia electronic publications. However, preservation of the information systems of cultural institutions is not regulated in any way, persons responsible for their preservation are often not designated, and sanctions for their damage or loss are not established.

Internet resources of cultural institutions

Internet resources of cultural institutions include mainly an increasing number of their own web-sites and portals, whose quality is been constantly improved. For example, according to the Catalogue of Library Websites posted at *library.ru*, there are 1139 web-sites of libraries of various types and affiliations, and this list is far from being complete.

Of particular value for the society are those web-sites of cultural institutions which provide remote access to electronic catalogues, reference and bibliographic databases, full-text electronic libraries, virtual exhibitions, static and moving images, as well as web-sites providing well-developed services (for example, enable the users to order literature, use a virtual reference service, etc.).

Financial and management documentation

Preservation of electronic financial and management documentation is another distinct challenge. E-documentation in Russia has become increasingly popular and has been introduced from above through the e-government programs. Russian domestic business already actively uses e-documentation. Courts, which are gradually turning to electronic modes, are getting accustomed to them and regularly take decisions based on the evidence submitted electronically. Introduction of electronic invoices is expected in Russia in the near future.

Working papers of the staff

The working papers of the staff are the major information resource of any cultural institution. These are texts of reports, lectures, presentations, photo and video records of events, draft documents of electronic publications, electronic correspondence, etc. They should also be preserved along with other library information resources. The loss of working papers leads to ungrounded duplication of effort and inefficient use of financial resources.

Preservation of Electronic Documents and Information Resources at Cultural Institutions: Goals and Tasks

The goal of preserving electronic information resources at non-depositary cultural institutions is to ensure efficient preservation of both collections in digital formats and cultural institutions' own electronic resources not included

in the collection; ensure access to them and their use until they migrate to modern technological platforms; and ensure long-term preservation of cultural institutions' own electronic resources not included in the collection.

Certain archiving and storage requirements should be met in the process of creation and use of electronic information resources at the cultural institutions to guarantee their expedited preservation. Expedited preservation means application of a set of measures which ensure preservation of information resources until their migration to new technologies.

To preserve digital pieces we have to identify and understand the **key challenges and goals** of this process. To this end, it is indispensible to:

1. Plan the activities aimed at preserving electronic information; **assign those responsible for** preservation**; develop and introduce regulations** on preservation and emergency response and **exercise management control** over their observance.

2. Apply a design approach. Cultural institutions should develop project documentation in compliance with the federal standards, as it is the practice in engineering. **E-data preservation activities should be included in any project aimed at creating electronic information resources** (digitization; e-library, web-site, e-catalogue creation, etc.). It is necessary to (a) calculate and substantiate the appropriate memory capacities and the set of hardware; (b) select and acquire system-level and special-purpose software (including antivirus, anti-hacker programs, etc.); (c) determine the need in electric power and labor costs to create resources in compliance with the established standards; (d) determine and substantiate the competences of the staff; (e) determine the sources and level of funding needed to maintain and preserve information resources following the termination of the project. The quality of project documentation and project execution should be monitored by the funding agency which should impose strict requirements on cultural institutions for the quality of design activities.

3. Use **hardware and software** adequate for the technological cycle of protection, preservation, access provision, preservation control, rerecording, etc.

4. Hire **highly qualified employees.** The competence and reliability of employees is a very important issue for long-term preservation of electronic information. Practice shows that employees should be highly qualified, trustworthy and responsible because electronic information preserved in libraries, museums and archives is open to abuse.

Additionally, employees should assume responsibility for the observance of the copying, updating, rerecording and other technological cycles. It is essential to provide regular retraining to the employees due to the ever-emerging novel technologies and information preservation facilities, as well as threats and risks.

5. Create special storage facilities or rooms to ensure:

- Protection, security, up-to-date fire alarm and fire extinguishing systems.

Today, virtually all cultural institutions (the major ones, at least) are equipped with security and fire systems, metal detectors at the entrance, etc. However, it is not just the presence of these devices that matters, but a regular check of their operability, special instructions and procedures developed for standard and emergency situations, and skills of the staff responsible to properly use them. Moreover, there is need in modern fire-extinguishing devices, those that use inert gases and powders, since the major and irrecoverable losses are often caused by incorrectly selected fire-extinguishers (water, foam and others), rather than the fire itself.

- Solid, uninterruptible power supply ensured by backup power systems (in case of blackout), UPS, filters, power stabilizers and the like.

Uninterruptible power supply in the storages used to preserve electronic information resources is essential because power failures may lead to irrecoverable loss of huge arrays of information preserved.

- Limited access of employees in compliance with the duties performed.

Such limited access in compliance with the duties performed can and should be established for various premises and departments of cultural institutions. This is particularly important for the premises used for long-term preservation of electronic information since any carelessness (intended or unintended) when handling storage devices may result in irrecoverable loss of information.

- Adequate physical and environmental storage conditions.
- Unfavorable environmental or physical storage conditions of optical disks may result in quite a fast loss of information recorded on them. This refers not only to optical or magnetic media. Other

special-purpose information storage media can be subject to adverse effects as well. Since modern devices allow for storing huge arrays of electronic information (up to hundreds and thousands TB), creation of special-purpose containments with the required physical and environmental conditions is important and vital, especially in large storage facilities.

To ensure long-term preservation of electronic documents and information resources it is essential to:

- Implement preparatory measures which are necessary for subsequent transition to new formats, standards, and soft- and hardware platforms (including development and introduction of standards, preservation metadata, methodological guidance);
- Carry out, when necessary, transition to new formats, standards, and soft- and hardware platforms in a way which allows for retaining all meaningful and functional features of original digital information resources, as well as for search, presentation and interpretation options for subsequent access and use.

Preservation of digital heritage cannot be a final result of any initiative or project, but it should become a regular, on-going and mandatory process for any institution engaged in the creation of its own digital information resources.

Practice and Challenges of Electronic Information Preservation in Cultural Institutions

The authors have conducted a sample survey among the major libraries, museums and archives, as well as the leading experts in informatization of culture and studied the documents of the major professional conferences and Internet resources.

The results of the survey have demonstrated that:

- All respondents consider the challenge of long-term preservation of electronic information highly topical;
- The major federal cultural institutions set up on their premises modern facilities to store electronic information, have the required technical and technological capacities, rooms, highly skilled employees and regulatory documents;

- These storage facilities are designed for preserving electronic information and the institution's own Internet resources in the process of their creation and use, and they do not help to handle the problems of long-term preservation of electronic information,
- Migration to new formats and technological platforms is or will be carried out when the need arises, and by cultural institutions themselves (provided the respective funding is available) in cooperation with the developers of software installed at cultural institutions (provided such developers are available),
- Cultural institutions do not deal with external Internet resources and do not ensure their long-term preservation;
- Cultural institutions have no procedures for long-term preservation of electronic collections on optical disks and do not handle the problems of either physical ageing of media or technological ageing of such publications,
- Experts draw attention to the burning necessity of introducing serious changes to Russia's legislation in order to establish a legal ground for long-term preservation of electronic information;
- Not all cultural institutions, even those of the federal level, can set up storage facilities for electronic information preservation, primarily, due to lack of funding, technical and technological capacities for safe long-term preservation of electronic information, appropriate administrative structure, required premises, high qualified employees and, in some cases, inability to understand the problem at all;
- Cultural institutions lack budget resources to establish and maintain such facilities; they need financial support to purchase and upgrade the necessary equipment, purchase and update on a regular basis the licensed software, provide adequate remuneration to highly qualified IT specialists;
- The efforts of libraries, museums and other cultural institutions to create and store digital copies are not coordinated;
- It is required to provide ongoing monitoring of dying technologies and formats and to develop methodological guides - neither of these tasks is within the scope of cultural institutions;

- In order to make grounded administrative decisions on the federal level, it is required to monitor electronic information resources of cultural institutions and to ensure preservation of electronic information.

Conclusions

In spite of rich experience accumulated by economically developed western countries and the growing acuteness of the problem of electronic information preservation, this issue remains rather obscure for Russia's scientific and cultural community; it is rarely discussed at conferences, workshops, professional fora; it is insufficiently highlighted in professional mass media and **has never been analyzed at either political or administrative levels (including high level). One can even state that the importance and scale of this problem, or at least its mere existence are hardly recognized in Russia.**

The legislative basis for information resource creation, registration, expedited and long-term preservation, access provision, creation of a legal deposit copy, protection of copyright and neighboring rights has not been developed or adapted to the changing environment.

The organizational structure of cultural institutions does not correspond to the challenges they face and the functions they should perform in order to preserve electronic information. There are no relevant regulations, instructions, procedures, as well as duly qualified specialists.

The responsibility for the preservation of electronic information at various levels and at various institutions has not been determined, as has not been determined the responsibility of officials for the loss of electronic information and the responsibility of producers and keepers of the digital content.

The funding of activities aimed at preserving electronic information (both newly generated and already existing) within programs, initiatives or events has not been envisaged.

Based on the experience in designing and implementing the National Program for the Preservation of Library Collections, **the authors propose to develop, discuss and adopt an elaborate and financially supported Russian national program for long-term preservation of cultural electronic information, which should be focused on and implemented by libraries, archives, museums and other cultural institutions.**

The Russian national program for long-term preservation of cultural electronic information should involve the implementation of the national strategy of cultural electronic information preservation, including:

- Development and adoption of general principles of creating and digitizing cultural electronic information resources;
- Compilation of inventory and creation of registers of electronic information resources;
- Development and adoption of general principles of selection of information resources for long-term preservation;
- Regular monitoring of the preservation of electronic information resources in various fields;
- Elaboration of measures to preserve the most valuable materials facing extinction, primarily in the audiovisual field, and born-digital materials;
- Assigning of persons and structures responsible for the preservation of electronic information, as well as personal responsibility for its loss;
- Development of regulatory documents on expedited preservation of electronic information and physical preservation of media (security and alarm systems, premises, climate, control over the physical condition of media, antivirus protection, protection of data transmission channels, storage of copies in different rooms, etc.);
- Development of special instructions and procedures to respond to standard and emergency situations;
- Scientific studies and technological developments on certain subjects;
- Coordination of efforts to preserve electronic information;
- Determination of the status of different types of electronic information;
- Technical, software, technological and methodological support of digitization, as well as of expedited and long-term preservation of electronic information;

- Enhancement of the existing automated document management systems (elements responsible for electronic information management) in cultural institutions;
- Development and introduction of normative and methodological materials, standards and metadata for the preservation of electronic information resources;
- Creation of mobile structures for digitization of documents to ensure their preservation within the collections of various cultural institutions;
- Improvement of normative and legal support of the processes of registration and management of electronic information;
- Proposals for improving the electronic information preservation laws and submission of these proposals to the legislative authorities in the form of a legislative initiative;
- Development of the training and retraining curricula for specialists in electronic information preservation and provision of access to it; unification and standardization of requirements for the level of knowledge of electronic information preservation and access matters; development of teaching and methodological materials;
- Information support to fulfill the task; conferences, workshops and round tables; publication of articles, books, and materials in professional and the public mass media;
- Provision of annual funding to hold events and undertake projects within the Russian national program for the preservation of electronic information in the cultural sphere;
- Wide-range discussions of the issues related to electronic information preservation;
- International cooperation in digitization, electronic information preservation and access provision.

Mechthilde FUHRER
Strasbourg
Council of Europe
Programme Adviser
Directorate of Culture and Cultural and Natural Heritage
Council of Europe

The Council of Europe contribution to the Preservation of Digital Information with a Focus on the European Convention for the Protection of the Audiovisual Heritage

The Council of Europe

The Council of Europe, founded in 1949, is the oldest European intergovernmental organisation, and is located in Strasbourg. Our mission is to strengthen democracy, human rights and the rule of law. The European Cultural Convention (1954) is one of our founding texts, along with the European Convention on Human Rights (1950).

The Council of Europe encourages strong dialogue for active democratic citizenship and human rights, thus promoting social interaction. The Council of Europe strives for a Europe where each person is granted with equal participation and creativity, access to culture and cultural rights. The intercultural approach, based on Europe's founding values, favours dialogue, based on equality, mutual recognition and non-discrimination. Two important texts have been issued thereupon, the *White Paper on intercultural dialogue of the Council of Europe* published in 2007 and the report *Living together – Combining diversity and freedom in 21st-century Europe* in 2011.[71]

The Council of Europe also gives high priority to co-operation with key partners. In the cultural field, we carry out a wide range of activities with UNESCO, the European Union, ALECSO, ALF, and, of course, INGOs. We are pleased to make a contribution to this International Conference.

[71] This report, drafted by a 'Group of Eminent Persons' was presented on the occasion of the 121st session of the Committee of Ministers in Istanbul on 11 May 2011. It will generate a higher political commitment for a range of current and future Council of Europe activities, against the backdrop of the Organization's reform process.

Standards

Now I would like to talk about preservation of digital information in the area of culture and cultural policies in the information society. Many departments within the Council of Europe deal with issues directly related to the theme of this Conference. The Council of Europe has created a long list of recently adopted recommendations, resolutions and declarations in the media field, dealing with human rights and the rule of law in the information society: there are documents on independence of public service broadcasting in member states, children rights in the new information and communication environment, digital agenda, increasing participation of member states in Internet governance, measures to develop public service broadcasting, etc.[72]

The Council of Europe member states have agreed to adopt a convention to protect their heritage. These conventions provide a common framework of action for policy makers responsible for safeguarding and enhancing heritage. The Council of Europe sets standards and monitors their application through the Convention for the Protection of the Architectural Heritage of Europe, the Convention on the Protection of the Archeological Heritage, the Convention on the Value of Cultural Heritage for Society.[73] Article 1 of the European Cultural Convention stipulates that "each Contracting Party shall take appropriate measures to safeguard and to encourage the development of its national contribution to the common cultural heritage of Europe."[74] The protection of cultural heritage is one of the priority areas and one of the pillars of cultural co-operation in the Council of Europe member states. They then translate the respective cultural policies into national laws which provide protection, renewal, restoration, exploitation and presentation of cultural heritage.

Rich audiovisual space

Defending Europe's cultural resources through its diversity of film and audiovisual projects is a vital means of reflecting democratic Europe.

[72] http://www.coe.int/t/dghl/standardsetting/media/Doc/CM_en.asp

[73] The Council of Europe prepared a report criticizing the Council of Europe Conventions at the end of September 2011. This report was urged to identify key conventions that might constitute a legal platform for all member states, as well as those conventions which have become obsolete, require reviewing or updating. This report illustrated that the Conventions in the cultural field are quite relevant.

[74] European Cultural Convention, CETS no. 018, Paris, 19.XII.1954.

Supporting creativity also helps us to better understand our heritage and pass on rich European cultural heritage to future generations. The Council of Europe assists the audiovisual sector with standards and guidelines, offers financial support to film production, and provides key information on audiovisual industry.

The **European Audiovisual Observatory** fosters greater openness of audiovisual sector in Europe and meets the information needs of audiovisual professionals. **Eurimages** is the European Fund to support Co-Production and distribution of Creative Cinematographic and Audiovisual Works. Two international treaties guide member states in promoting co-productions and in preserving Europe's extensive heritage in moving images: the **Convention on Cinematographic Co-Production** (1992) and the **European Convention for the Protection of the Audiovisual Heritage (EAHC) with its Protocol on the protection of television productions** (2001)[75].

The European Convention for the Protection of the Audiovisual Heritage (EAHC) and the Protocol thereto are the first binding international instruments to provide systematic archiving of audiovisual works, contributing to

- the systematic preservation of cinematographic work as a first step towards its access and use by future generations;
- building a compulsory legal deposit as a basic principle for the safeguarding of audiovisual heritage;
- paving the way for co-operation between film archives in Europe.

The Convention was drafted in 2001 and came into force in 2008. The European Broadcast Union, which has recently signed a Memorandum of Understanding with the Council of Europe, was also invited to join the co-work in the framework of the EAHC as an observer.

The second meeting of the Standing Committee on the European Convention for the Protection of the Audiovisual Heritage has just taken place.[76] The Secretariat distributed a Questionnaire to the members of the Steering Committee for Culture asking for information on why member

[75] European Convention for the protection of the Audiovisual Heritage, CETS no. 183, Strasbourg, 8 November 2001.

[76] This meeting took place in Budapest on 24 September 2011 upon invitation of the Hungarian authorities.

states have not ratified the EAHC and what hinders them from "translating" the Convention into national law. The Questionnaire revealed that there will be more signatories to the Convention in the near future, meaning also that the Protocol will be able to enter into force. But there is still a need for more international co-operation and sharing of experience on respective practices in the field of preservation. This need for more international co-operation has been stated several times at this Conference.

Indeed, it is crucial to promote the Convention and to transfer the knowledge about audiovisual heritage protection to various countries. Having ratified the Convention every country agreed on the need to ensure preservation of their cinematographic heritage, but, in fact, practices differ. There are differences in a comprehensive legal deposit scheme, an obligatory deposit (for publicly funded cinema work), and voluntary schemes of deposit or only through selection procedures. It is equally important to collect and disseminate information about these different schemes, their strengths and weaknesses, and their impact on collection, preservation and availability of film heritage. The true benefit of the Convention can only be recognized in comparison with other schemes. By showing how the system works in those countries that have ratified the Convention, others may be convinced to join the Convention as well.

The Secretariat has therefore developed a special section on Monitoring the CoE EAHC and Related Policies in Europe in the *Compendium of Cultural Policies and Trends in Europe*, online information constantly updated and monitoring system on cultural policies, one of the Council of Europe's leading projects in the cultural field. It allows conducting comparative studies on diversity, systematically monitoring cultural policies in more than 40 countries, to communicate via the web and network to exchange practical experience.[77]

The Parliamentary Assembly of the Council of Europe has discussed a draft report on *Access to and protection of the audiovisual cultural heritage* at their October session and the input of the Standing Committee on EAHC will be vital. However, the Council of Europe to participate in other projects carried out in the audiovisual field by the European Union, e.g., the EU study *Digital Agenda for European Film Heritage*. There is Europeana

[77] http://www.culturalpolicies.net

and the recommendations of *Comité des sages* feeding into the European Commission's broader strategy, under the Digital Agenda for Europe, to help cultural institutions make the transition towards the digital age and to search for new and effective business models that accelerate digitalization and allowing fair remuneration for copyright holders (where necessary).

Work prospects

The Convention for the Protection of Human Rights and Fundamental Freedoms states in Article 10:

"Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This article shall not prevent States from requiring the licensing of broadcasting, television or cinema enterprises."

This article grants the right to freedom of expression, the right to seek and receive information. The ultimate goal is to grant access to information, culture and heritage. The question is how to get it, how to achieve this in our quickly changing society, where libraries and archives are closing down due to financial restraints. In that context, what is the role and the 'added value' of this Convention? How does this text help to promote and guarantee basic rights and access to information?

Many other issues need to be touched upon in the larger context of preservation and protection. In the area of online media, what is the role of digital archives, digitalization, on-line libraries? What is the role of the state, companies and public service media?

The right to access information stated in Article 10 also implies preservation. The principle is important. What does it mean? The text of the Convention gives the answer. The Standing Committee certainly intends to undertake a detailed analysis of the Convention text in order to tackle future digital preservation and restoration of audiovisual heritage in Europe, facing the challenges of technical standards for long term archiving, such as deterioration of originals, new support materials and intellectual property rights.

In the context of the Convention for the Protection of Human Rights and Fundamental Freedoms, there are issues yet to be tackled, such as

freedom of expression[78], the role of the state in the new context of the online world with social media and access to Internet being in the process of becoming a fundamental right. What does this mean to audiovisual heritage? Article 1 of this Convention[79] provides the right to intellectual property. Indeed today the question of copyright is a major issue, libraries are closed, and books are digitized and are available online. Online archives are on their way forward. But how will this be managed? Is it in the public interest? The Rec 16 (2007) on public service value of the Internet deals with the issues on the use of the public sector information, access and public service media in audiovisual heritage.

All these themes illustrate the very timing of this International Conference, when freedom of expression and copyright issues need to be carefully monitored, as well as is the case for public service media and the public value of the internet and its role to be considered in order to promote common values and to maintain the culture's key role in forming a Europe of solidarity and shared standards. Such interdisciplinary forums as this one are necessary to bring together professionals, government representatives and international organizations in order to animate exchange and co-operation, set standards for future work and confirm the vital role of culture in fostering democratic processes.

[78] Cf, Copyright and Human Rights. Report prepared by the Group of Specialists on Human Rights in the Information Society (MC-S-IS), September 2008, Directorate General of Human Rights and Legal Affairs, Council of Europe, Strasbourg, June 2009.
[79] Article 1: Obligation to respect human rights.

Vladimir KOLYADA
Moscow, Russian Federation
Director
Russian State Archive of Phonodocuments

Experience in Digital Information Archiving

Preservation of digital information is becoming increasingly topical for the institutions which are responsible for archiving such information as the contemporary society heritage. The acuity of this task is increasing together with the impact of scientific and technical progress on documentation, communication and archiving of technically produced content. Certain experience accumulated in this field by the Russian Archive of Phonographic Documents allows me to suggest some milestones.

The Russian State Archive of Phonodocuments focuses on audio documents as a type of modern documents most sensitive to the scientific and technological changes. Intrinsically, a phonographic document is a trinity of technological progress, historical evidence and public phenomenon. Each of these components exists and makes itself evident in inextricable connection and interdependence with the other two, in permanent evolution of the internal and external forms throughout the life of a sound recording. To a large extent, the technological part of audio recording of information has determined not only the main object of the activity of the Archive of Phonodocuments but also the structure of its internal operations. At present, the digital audio information processing technology occupies the leading place at the stages of item registration, communication and archiving. As a result, we have to amend many provisions and methods of the archive practice and introduce visible changes into the Archive's position in the information environment.

In the modern world, in the conditions of globalization of public processes, audio information is becoming a phenomenon which determines the development of civilization. Growing flow of audio information and broadening of its functional field are stimulated by a simultaneous and interrelated impact of such a powerful factor as public demand in efficient and consumer-accessible information combined with rapid progress in the means of registration and communication of sound on the basis of the latest hardware. Audio information is becoming increasingly important because of its influence over public psychology and political and aesthetic

environment; more and more often it has been used not only to record social processes but to manage and manipulate them. This attracts attention of many governmental, public, political and social institutions, as well as business circles engaged in public information, audio publications and distribution of digital content. All aspects of this situation are becoming more significant as digital information is gradually occupying a more visible place in the activity of governmental and public institutions, media structures and cultural sphere. Consequently, we can witness the growing topicality of archiving of this information at all stages of its creation and circulation for the purpose of preserving it as a historical memory.

Introduction of digital technologies into recording processes has led to drastic quantitative and qualitative changes in the environment where audio information is produced. The affordability of digital documenting technologies makes them accessible for wider circles of users of various legal statuses. As a result, we are witnessing a sharply increasing information flow and diversity of its content. However, the volume of digital audio information poses a serious problem for those who are involved in singling it out in the content which could be of interest for permanent archiving. Additionally, the digitized information content exists in a certain virtual form which is not related directly to a physical information medium and remains in a latent state as an element of a certain electronic system. This state is going to be the main one for audio information at all stages of its existence, thus aggravating the problem of its archiving and searching for the most rational ways and methods of its preservation.

The contemporary information environment is characterized not only by speedy changes in quantitative parameters but also by highly expedient changes in the internal content: replenishment and correction take place in real time and depend entirely on this flow manager. Reforming of the information environment mates with certain changes in the so-called external function of an archive and its relationship with other subjects operating in the given environment.

The first and the main task of an archive is formation of a representative array of documents reflecting the contemporary historical process. For a phonographic archive this task looks and will look as a formation of a certain array of audio information, both in a physical format and as a digital content (mind that the latter is due to become the main one with time). This means that, in future, an archive will change qualitatively from a

physical collection of documents as a tangible proof to an electronic bank of virtual audio information. This forecast is based on the examination of the current state of the information field and the introduction prospects of modern technologies.

The Archive of Phonodocuments accumulates in its collection such audio information which has been made generally accessible. The electronic and broadcasting media have transferred practically to a digital format, and audio publishing companies are following them, increasingly distributing their products in the network environment, which are the main sources of replenishment of the Phonographic Archive.

Preservation of historical memory about the sphere of culture and art is the mission of the State Archive of Phonodocuments; more and more facts and events in this field are frequently created, recorded and distributed in a digital format. Advancement of digital technologies gave birth to specific lines of creativity which can be completely realized only in electronic environment. At that, the archiving of such information remains on the "outskirts" of interest, and replenishment and amendment of such information are conducted rather spontaneously. Appearance or disappearance of information about any cultural event or fact is not stipulated by any framework, provisions or the like. Such digital content emerges and exists in the electronic networks exclusively and is not registered, as a rule, in any physical format on any information medium. Its extraction for archiving and preservation as an evidence of a cultural and historical process becomes entirely the responsibility of an archive because the attention of its creator is not focused on such details, as a rule.

The bulk of audio information is being created in organizations which differ by their status, scope and profile of activity, and which become the source of the collection of the Archive of Phonodocuments. The use of digital technologies is increasing and this more and more often drives out traditional audio documents. The progress in electronic communication systems excludes the necessity to register audio information on a physical medium; registration, communication and conservation of audio information take place in the virtual electronic environment and acquire a material form only in case of necessity. As for the mission of the Archive, this situation means a quality change in generating the array the Archive is responsible to preserve. A traditional connection of the Archive with acquisition sources, which implied preliminary selection, expert examination and physical

acceptance of phonographic documents on various media, is giving way to digital content to be included in the Archive collection through widely used communication channels. In view of growing globalization of information environment this signifies a serious broadening of formal and subject-oriented frameworks of the Archive collection, because a free exchange of information is a basic principle of the contemporary civilization. At the same time, this principle is accompanied by many limitations, in the first place, IPR limitations, which result in an ambiguous understanding of norms, their voluntary interpretation and selective application. Growing flow of such specific digital information adds to the complexity of archiving of this information and to the routine functions of the Archive.

Transformations reached even such a traditional function of the Archive as expert examination and selection of documents for permanent preservation - information, in our case. Digital technology allows performing these operations remotely, without a direct contact with the information source. On the one hand, this contradicts with the applicable norms, but on the other, this is the only way to accumulate the material of cultural and historical value. In many cases only the operative initiative of the Archive, i.e. immediate extraction of information from the IT network and registration in the archive's system, allowed us to archive and preserve information. It is obvious that this task may be fulfilled only if the Archive is permanently present in the information environment of a certain profile. The immensity of the present information environment poses a task of specifying the Archive's focus within the conglomerate of growing information flow, which in practice means determination of the priority sources of information of interest.

Thus, digital technologies lead to transformation of the position of the State Archive within the information field. The traditional scheme, according to which the Archive was the final element in the chain and selected and preserved documents (information) after a certain period of time, is displaced by the active role in the process which implies permanent presence in this process and equality with other participants. This transforms the expert examination of the overhanging information flow and forces us to conduct it in a more efficient way minding the above variability of this flow. In these conditions, the successful building of the Archive collection depends directly on the professionalism and skills of the personnel, which is actually another burning problem of ours today.

The tasks to be fulfilled by the State Archive in this field bring to the forefront the legal issues connected with the authority of the archive to conduct specific operations, the more so as the source of information itself may be formally out of the Archive's powers currently in force but the information it offers may be undoubtedly important for its preservation. This moment is especially significant due to the federal status of The Archive of Phonodocuments, which outlines the legal frameworks of its competences, as opposed to an individual user of communication networks. At present, there are no strict legal provisions concerning the handling of digital information for archiving purposes, which is fraught with acute and ambiguous situations which may occur when audio information concentrated in the Archive in this way is made generally accessible. These problems are of exceptional significance for The Federal Archive of Phonographic Documents because it deals with IPR-protected documentation and information.

Digital technologies do not only exert their influence on the environment of audio information creation and distribution, but demand certain amendments to be made in the Archive's internal operations.

The Archive's audio documentation, in any form, needs to be organized and classified for the purpose of sorting, guaranteed preservation and general accessibility. Digital information poses additional issues connected with the determination and registration of its volume, consolidation of search and discount images, creation of conditions for its access and comprehensive use. Here, we consider expedient to use similar electronic resources and produce in a traditional form only that set of documents which is needed to formalize the parameters of the Archive collection. The growing digital information flow demands that we amend the classification system of the array of documents, which has been adopted by state archives, at the level of 'registration unit/stock-keeping unit' in the first place and within the Archive collection measuring system.

Application of digital technologies allows us to improve the functions of the Archive in terms of including its collection into scientific and public flow. It opens up an opportunity to build multi-subject reference formats, create a comprehensive awareness system for a researcher informing him/her about the composition and content of the archive collection of audio information. We think it beneficial to include such developments into a network resource. However, this trend implies certain legal constraints.

The legal maintenance of the activity of the State Archive becomes of paramount importance when dealing with digital information, both generally and in specific ways.

Internal functions of The Archive of Phonographic Documents when using digital technologies acquire a new quality, in the first place, a quality needed for collection preservation. An opportunity to use digital technologies in the maintenance of documents and restoration of the tone quality allows the Archive to enhance the technical standards of audio information preservation. It is well-known that that sound-recording, in the course of its development, has produced documents on various physical media of different parameters, whose stability, longevity and tone qualities are far from being unambiguous, and demand permanent collection maintenance operations. Electronic archiving of audio information allows us to conduct these operations on a qualitatively new level and to improve tangibly the guaranteed preservation frameworks of audio collections.

Lev NOL
Moscow, Russian Federation
Head of IT Department
The Pushkin State Museum of Fine Arts

Vladimir OPREDELENOV
Moscow, Russian Federation
Deputy General Director
"EPOS Group" CJSC

Preserving Digital Arrays of A Major Museum (Experience of the Pushkin State Museum of Fine Arts)

In recent years the problems of preservation and presentation of global cultural heritage to the wide public by means of information and communication technologies (ICT) have been considered among the most important for the world civilization. These problems are given the most serious attention both nationally, and internationally.

Unfortunately, non-coordination of these efforts results in the lack of national and international standards in many areas of this activity, and, consequently, low integration of existing information systems.

The Government of the Russian Federation gives full attention to preservation of the Russian national cultural heritage by means of ICT. It is subject both to immovable historical and cultural monuments, and to museum collections. With many Russian museums using AIS and/or building their own databases, the general status of ICT acquisition and maintenance remains low and calls for systematic development.

Building the National Catalogue of the Museums Funds of the Russian Federation is still on the agenda, and the Ministry of Culture of the Russian Federation has taken this issue under its control.

For many years, the Pushkin State Museum of Fine Arts (further referred as Museum) uses ICT in its almost every practical activity – control, collection building, conservation, exhibitory, etc. Within the Museum network infrastructure, since 2003 the Complex Automated Information System KAMIS-2000 has efficiently operated. As of September 23, 2011, the database numbers 538,890 records with museum objects descriptions,

68,162 of them contain digital images. The data on collections (both text descriptions, and digital images) have been intensively increasing.

Under these circumstances, it is vital to use and develop ICT in the Museum's activities properly. Unfortunately, hard- and software, existing procedure and technologies do not completely meet growing standards and have to be streamlined fundamentally. To a significant extent, it applies to processes of acquisition, preservation and use of digital images. Thus, digital shooting with semi-professional equipment is accomplished in different Museum divisions, and procedure related to the operation quality and digital array preservation is not always observed. Until recently, the Museum lacked single infrastructure and regulations for some activities, in particular, those related to digital data control and preservation. Digital arrays were stored separately; no equipment to support centralized storage and use of digital images was available.

Obviously, Electronic Depository (ED) of the Museum's digital resources (DR) has to be built to ensure the system of centralized digital data acquisition and preservation management, and to make systematic rules for digital graphic and audiovisual resources use, devoted procedural, hardware and software complex. The ED Project has been being developed since 2009.

Electronic Depository

The goal of the electronic depository is to solve the totality of problems related to acquisition, preservation and use of the digital resources necessary for the Museum's activities (primarily, graphic, audiovisual information).

The Museum's digital resources comprises approximately: digital images (80%), video materials (10%), sound records (7%), other documents (texts and graphic documents, publications layout files, etc.) make just 3%.

The ED Project has been considering the special features of existing digital resources and those being created. The main focus of the Project is on bringing the Museum's activities within the single system of proper preservation and use.

The ED DR main functions are:

- Manageable accumulation of the database of digital images and audiovisual materials;

- Assured preservation of digital resources;
- Controllable data images use;
- Manageable accumulation and acquisition;
- Shaping dedicated automated information system;
- Controllable use.

Project stages

Stage 1 (2009) – analysis of the Museum's digital resources, study and analysis of global experience in the area, development of the concept, Part I Preliminary Design development, purchase of hardware. This stage has been completed by now.

Stage 2 (2010) – Main directions:

- Part II Preliminary Design development, including procedure for the Museum's departments using DR ED,
- DR ED software purchase,
- Customization of DR ED software and hardware,
- Direct location of high-precision images into DR ED hard- and software ,
- Creation of ED reference and search tools based on KAMIS and web-technologies,
- System testing,
- Testing procedures for Museum departments working with DR,
- Trial run, drafting of recommendations for adaptation and improvement.

In 2011, main efforts have been focused on refining adopted technological solutions, improving DR ED hardware and software, downloading Museum's digital resources into ED.

For 2012, the production run of existing ED version is planned, along with further improvement of soft- and hardware complex, including provision for long-term data storage.

Functional outline of the current version of the Museum's DR ED hardware and software

Server system

WWW-server 1 (KAMIS) – the www-server is intended to work on ORACLE database platform with preserving JPEG pyramid images inside the database. The server supports KAMIS at authorized staff workstations.

WWW-server 2 (KAMIS backup) – the server duplicates **Server 1.**

WWW-server 3 (ED administrator's production www-server) – the server will enable to work directly with ED file storage and will be provided with software for professional image and audiovisual data processing. The server will be available exclusively from the workstations at Visual Information Department.

WWW-server 4 (ED Reference and Search Tool) – the reference and search server based on KAMIS database will be based on web-technologies and will enable to get information on the collection and ED images from any computer. Interface will be as user friendly and understandable as possible and will work as dedicated intranet site via a standard Internet browser.

The Moscow Declaration on Digital Information Preservation

Final Document

International conference

Preservation of Digital Information in the Information Society:
Problems and Prospects

October 3–5, 2011

The preservation of digital information – a new problem of our era – concerns the whole world, and every man and woman. It arose at a time when electronic gadgetry is used in every sphere of life; when the volume of born-digital or digitized information is growing exponentially; when the variety of formats of information presentation, and the types and classes of digital objects is increasing apace; technology and software are being rapidly updated; the lifetime and reliability of present-day digital media do not meet the challenges of long-term information preservation; and last but not least, the virtual space of digital information impedes law enforcement and obliterates the borders between national jurisdictions. Despite all efforts, the preservation of digital information lags behind technical development and social changes.

For two millennia – ever since the Library of Alexandria was established – humankind has been learning to preserve information on analogue carriers. With that goal in view, all developed countries had established by the mid-20th century a ramified network of comprehensive and specialized libraries, archives, museums, and scientific and technical information centres. They shared basic functions and actively interacted on their own level and between levels nationally and transnationally. They all were parts of a sophisticated infrastructure comprising the press, research institutes, method-setting centres, the educational and personnel retraining system, national and international professional associations, equipment manufacturing industries, standardizing organizations, etc. The methods and criteria of information source selection, and sophisticated and globally standardized methods of, information accounting and storing and metadata creation are regularly updated.

Unlike the situation in analogue data storage, a majority of countries have not yet:

- elaborated the philosophy of long-term preservation of digital information;
- established a regulatory legal basis and efficient policies leading to its establishment; and
- created an infrastructure for digital information preservation (all too often, its relevance is not even realized).

Traditional memory institutions of these countries – libraries, archives and museums – cannot cope with the snowballing amount of digital information, and are making their alarm public.

The international conference "Preservation of Digital Information in the Information Society: Problems and Prospects" was initiated to enhance the understanding of the importance, scope and topicality of preserving digital information in the professional milieu, at the political level and among the public-at-large; to promote the evaluation of priority problems and political and professional strategies in the sphere of digital information preservation and streamlining global, regional and national policies, legislation and practical activities in this field.

Convened in Moscow, the conference took place on October 3–5, 2011 to gather approximately 150 participants representing 37 countries: heads and leading experts of major libraries, archives, museums, research institutes, universities, international organizations, government bodies, media outlets, publishers, research and technical information centres, the ICT industry and other entities interested in the elaboration of the theme of digital information preservation.

The Ministry of Culture of the Russian Federation, the Federal Agency for Press and Mass Communications, the Commission of the Russian Federation for UNESCO, the Russian Committee of the UNESCO Information for All Programme, the Interregional Centre for Library Cooperation, and the State Tretyakov Gallery jointly organized the conference within the framework of Russia's chairmanship in the UNESCO Information for All Programme.

The conference concluded on the necessity of urgent political and practical measures. Otherwise, the world may discover in the near future that:

- a huge amount of Internet and other information, including information pertaining to history, culture and research (the content of portals, websites, blogs, social networking sites, celebrities' electronic correspondence, private and institutional electronic archives, etc.) has never been collected and so is irretrievably lost to future generations;
- a vast amount of digitized information has been lost because either its digitization had no adequate support or it was not catalogued correctly and in time, and so cannot be retrieved, or it was not properly stored or, again, was not transferred into new digital formats in due time.

By guaranteeing reliable passage of information from person to person and from generation to generation, we can prevent the advent of the so-called Dark Information Age, also described as "digital Alzheimer" and "digital amnesia". Efforts to preserve digital information should start the instant an information object emerges.

Proceeding from the above, conference participants declare the necessity of the following measures in the following spheres:

1. **Politics**. Support of the formation and development of the philosophy, strategy and policy of preserving digital information at the national and international levels, which comprise the socio-cultural, ethical, legal, economic, administrative, personnel, technical, technological and other aspects. The preservation of digital information must become an inalienable part of cultural, educational, research and information policy, and the policy of information society building.
2. **Information/education**. Promotion of the awareness of decision-makers and the public-at-large on potential risks and basic principles pertaining to digital information storage, including long-term storage.
3. **Education**. Promotion of the elaboration, development and circulation of educational curricula and training courses of varying levels for the heads and experts of memory institutions (mainly libraries, archives, museums, and scientific and technical information centres) and other institutions that are (or should

be) involved in activities toward the preservation of digital information. Inclusion of basic knowledge and competences connected with digital information preservation in activities/programmes/curricula enhancing the media and information literacy of professionals in the information field and the public-at-large.

4. **Research**. Promotion of research in the philosophical, political, economic, socio-cultural, organizational, legal, personnel, technological, methodological, method-setting, ethical and other aspects of the preservation of digital information. Support of national and transnational cooperation to elaborate decisions and standards, and experience exchanges for the preservation of digital information.

5. **Economy**. Elaboration of basic organizational principles of funding the long-term preservation of digital information by traditional memory institutions. Inclusion of the evaluation of efforts, approaches and decisions from the feasibility point in the number of mandatory components of strategies and activities in the preservation of digital information, and raising relevant public awareness.

6. **Cooperation**. Promotion of interdepartmental cooperation of memory and educational institutions and administrative bodies with private businesses and other stakeholders of digital preservation processes, including public and private initiatives and projects; development of international cooperation.

7. **ICT industry**. Establishment and strengthening of cooperation with the ICT industry to include procedures promoting/guaranteeing long-term preservation of digital information in operating systems and basic supplementary packages. Promotion of the elaboration and implementation of free and open software for the preservation of digital information.

The conference addresses the following proposals:

to UNESCO:

- To use the UNESCO Information for All Programme, which includes information preservation in its five top priorities, as an

international interdisciplinary and inter-institutional platform for the formation of the political framework, for discussions and experience exchanges.

- To update the Charter on the Preservation of Digital Heritage, and upgrade it as a regulatory instrument of a high political level.
- To consider the preparation of a World Report on the Preservation of Digital Information with the following goals in view:
  - ▷ all-round and maximally comprehensive definition of problems and challenges connected with the preservation of digital information;
  - ▷ analysis of the present state of activities to preserve digital information that seeks to meet such challenges and addresses arising problems; identification of the perpetrators of these activities; assessment of their goals, and means and methods of such activities;
  - ▷ awareness of the actual/desirable patterns of coordinating relevant efforts at the national, regional and global levels;
  - ▷ assessment of the level of familiarity with and apprehension of the various aspects of information preservation by decision-makers, heads and experts of memory and other relevant institutions, and the public-at-large;
  - ▷ elaboration of action plans of various levels and on various approaches pertaining to the sphere of digital information preservation, and of a system of indicators of the success of such plans implementation.

to IFLA:

- In cooperation with UNESCO, to contribute to the elaboration and circulation of quality educational programmes and curricula on the preservation of digital information for library managers and personnel.
- To encourage the inclusion of components pertaining to the preservation of digital information in information literacy programmes and curricula.
- To actively participate in research aimed at:

- ▷ elaboration of criteria for the evaluation and selection of electronic content for long-term storage;
- ▷ responsibility delineation of memory institutions;
- ▷ exposure of gaps (unprotected digital heritage);
- ▷ creation of a regulatory framework for activities at various levels to preserve digital information; and
- ▷ exchanges of work experience, methods and technologies.

to national governments:

- To include the preservation of digital information as an inalienable part of cultural, educational, research and information policy, and as part of any national information society policy;
- To support research in the various aspects of the preservation of digital information.

to governmental and non-governmental institutions which are involved in digitization projects:

- To include the long-term preservation of digital content produced by digitization projects as an integral component of project planning and execution.

Authors

María ALARCÓN RENDÓN – Information Technology Director, National Archive of Mexico (Mexico City, Mexico)

Giovanni BERGAMIN – Head of the Information Systems, National Central Library of Florence (Florence, Italy)

Tatyana BOGOMAZOVA – Head of IT Department, Museum of Anthropology and Ethnography (Kunstkamera), Russian Academy of Sciences (St. Petersburg, Russian Federation)

Nadezhda BRAKKER – Senior Expert, Centre of Informatization in the Sphere of Culture (Centre PIC)

Susana FINQUELIEVICH – Senior Researcher, National Council for Scientific and Technical Research; Director of the Research Program on Information Society, Instituto de Investigaciones Gino Germani, University of Buenos Aires (Buenos Aires, Argentina)

Barbara Jean FORD – Member, U.S. National Commission for UNESCO; Director, Mortenson Center for International Library Programs; Distinguished Professor, University of Illinois at Urbana-Champaign (Urbana-Champaign, USA)

Mechthilde FUHRER – Programme Advise, Directorate of Culture and Cultural and Natural Heritage, Council of Europe (Strasbourg, France)

Ramesh GAUR – Librarian & Head, Kala Nidhi Division, Indira Gandhi National Centre for the Arts (New Delhi, India)

Natalia GENDINA – Director, Research Institute of Information Technologies in Social Sphere, Kemerovo State University of Culture and Arts; Member, Russian Committee of the UNESCO Information for All Programme (Kemerovo, Russian Federation)

Alisher ISHMATOV – Director, Information Library Centre of the Republic of Uzbekistan (Tashkent, Uzbekistan)

Grethe JACOBSEN – Head of Department, The Royal Library (Copenhagen, Denmark)

Lyubov KAZACHENKOVA – Editor-in-Chief, Sovremennaya Biblioteka (Modern Library) Magazine (Moscow, Russian Federation)

Esa-Pekka KESKITALO – Chief Analyst, National Library of Finland (Helsinki, Finland)

Hayk Ojen KHANJYAN – President, "Bi Line" Ltd; Head of the Culture and Language Division, Information Technologies Development Support Council of the Republic of Armenia (Yerevan, Armenia)

Igor KISELYOV – Head of Information Systems, Russian State Archive of Social and Political History (Moscow, Russian Federation)

Sergey KLESHCHAR – Head of Department, Research Institute of Reprography (Tula, Russian Federation)

Vladimir KOLYADA – Director, Russian State Archive of Phonodocuments (Moscow, Russian Federation)

Elena KOZLOVA – Director, INFORMREGISTR Research and Technology Centre (Moscow, Russian Federation)

Leonid KUIBYSHEV – Director General, Centre of Informatization in the Sphere of Culture (Centre PIC); Member, Russian Committee of the UNESCO Information for All Programme (Moscow, Russian Federation)

Evgeny KUZMIN – Chair, Intergovernmental Council, UNESCO Information for All Programme; Chair, Russian Committee, UNESCO Information for All Programme; President, Interregional Centre for Library Cooperation (Moscow, Russian Federation)

Mikhail LARIN – Director, All-Russian Research Institute of Archival Science and Records Management (Moscow, Russian Federation)

Peter Johan LOR – Extraordinary Professor, Department of Information Science, University of Pretoria (Sedgefield, South Africa)

Maurizio MESSINA – Head of the IT services, Marciana Library in Venice (Venice, Italy)

Sanjaya MISHRA – Programme specialist, UNESCO Sector for Communication and Information (Paris, France)

Oleg NAUMOV – Deputy Head, Federal Archival Agency (Moscow, Russian Federation)

Lev NOL – Head, Information Technology Department, The Pushkin State Museum of Fine Arts (Moscow, Russian Federation)

Vyacheslau NOSEVICH – Director, Belarusian Research Center of Electronic Records (Minsk, Belarus)

Vladimir OPREDELENOV – Deputy General Director, "EPOS Group" CJSC (Moscow, Russian Federation)

Alfredo RONCHI – Secretary General, MEDICI Framework of Cooperation (Milan, Italy)

Sabine SCHRIMPF – Scientific Officer, Deutsche Nationalbibliothek (Berlin, Germany)

Dietrich SCHÜLLER – Vice Chair, Intergovernmental Council, UNESCO Information for All Programme; Consultant, Vienna Phonogrammarchive of the Austrian Academy of Sciences (Vienna, Austria)

Yakov SHRAIBERG – Director General, Russian National Public Library for Science and Technology; Member, Russian Committee of the UNESCO Information for All Programme (Moscow, Russian Federation)

Sinikka SIPILÄ – President-Elect, IFLA; Secretary General, Finnish Library Association (Hameenlinna, Finland)

Andris VILKS – Director, National Library of Latvia (Riga, Latvia)

UNESCO Information for All Programme (IFAP)

IFAP was established as a leading intergovernmental programme in 2001 by UNESCO. IFAP is aimed to assist member states in planning and realizing a balanced national policy on national information policies framework. The Programme is based on interdisciplinary, inter-sector and integrative approaches. Information preservation is one of five IFAP priority lines of action, along with information accessibility, information literacy, information ethics and information for development.

http://www.unesco.org/new/en/communication-and-information/intergovernmental-programmes/information-for-all-programme-ifap/homepage/

Russian Committee of the UNESCO Information for All Programme

Russian Committee of the UNESCO Information for All Programme was established in 2000 and became the first national committee of the Programme. Committee members represent state bodies, educational, research, cultural, communication, public and business organizations. The Committee provides assistance in implementing the Programme's ideas, tasks, concepts and priority lines of action at both national and international levels. It participates in the improvement of policies and legislation in the fields of culture, education, communication and information for the purpose of building inclusive information society/knowledge societies. The Committee contributes to the preparation of analytical reports and elaboration of practical guidelines, to the establishment of the best practice centres and to the improvement of activities of various institutions within its competence.

www.ifapcom.ru

Interregional Library Cooperation Centre

Interregional Library Cooperation Centre (ILCC, established in 1995) is the working body of the Russian Committee of the UNESCO Information for All Programme. In addition to contributing to the IFAP implementation, ILCC participates in (a) drafting and implementing in

Russia the governmental library policy and national programmes aimed to preserve library collections; (b) developing all-Russian public centres of legal and other socially meaningful information; (c) reading promotion; (d) advancing professional library training and (e) developing multilingualism in cyberspace. On a regular basis, ILCC drafts, publishes and disseminates information and methodological materials on development of librarianship, culture and information policy.

www.mcbs.ru

Preservation of Digital Information in the Information Society

Proceedings of the International Conference
(Moscow, Russian Federation, 3–5 October, 2011)

Compilers: *Evgeny Kuzmin* and *Tatiana Murovana*

Translators: *Elena Malyavskaya, Tatiana Zverevitch, Tatiana Butkova* and *Ekaterina Komarova*

Responsible for issue: *Sergey Bakeikin*

Publisher:

Interregional Library Cooperation Centre

2A, bldg. 1, 1st Basmanny Lane, Moscow, 105066

Tel.:+7(499) 267 33 34, Fax: +7(499) 263 26 61

www.mcbs.ru