COMMISSION OF THE EUROPEAN COMMUNITIES

Brussels, 30.9.2005
COM(2005) 465 final

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS

i2010: DIGITAL LIBRARIES

{SEC(2005) 1194}

{SEC(2005) 1195}

Table of Contents

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS

i2010: DIGITAL LIBRARIES

1.   INTRODUCTION

In a letter of 28 April 2005 to the Presidency of Council and to the Commission, six Heads of State and Government advocated the creation of a virtual European library, aiming to make Europe's cultural and scientific record accessible for all. The Commission has welcomed this plan and will contribute to it through the i2010 flagship initiative on digital libraries.

The present Communication outlines the vision underlying the digital libraries initiative and presents its main building blocks. At the same time it constitutes the first element of the initiative, dealing with the digitisation, online accessibility and digital preservation of our cultural heritage**.** The Communication analyses the challenges for bringing out its full cultural and economic potential and proposes a first set of actions which will contribute to overcoming the present fragmentation of efforts in Europe.

2.   i2010: DIGITAL LIBRARIES

The **digital libraries initiative** aims at making European information resources easier and more interesting to use in an online environment. It builds on Europe's rich heritage combining multicultural and multilingual environments with technological advances and new business models.

Digital libraries are organised collections of digital content made available to the public. They can consist of material that has been digitised, such as digital copies of books and other 'physical' material from libraries and archives. Alternatively, they can be based on information originally produced in digital format. This is increasingly the case in the area of scientific information, where digital publications and enormous quantities of information are stored in digital repositories. Both aspects – digitised and born digital material – are covered by this initiative.

Three main strands will be followed to realise the potential of digital technologies for widespread and easy access to information:

- **Online accessibility**, a precondition for maximising the benefits that citizens, researchers and companies can draw from the information.

- **Digitisation of analogue collections** for their wider use in the information society.

- **Preservation and storage** to ensure that future generations can access the digital material and to prevent precious content being lost.

The digital libraries initiative will address these strands in several building blocks. Whereas the present Communication deals with Europe's **cultural heritage**, a second key area is **scientific information**. This is an area with its own specificities and dynamics in view of the need to handle and store huge quantities of digital data, and the rapid growth of publications only available in digital form. A Communication foreseen for 2006 will set out the specific challenges for digital libraries in this area, including the role of the supporting high capacity infrastructure and the actions to be undertaken at European level.

An **online consultation** accompanying the present document will deal with some of the major challenges influencing digitisation, online accessibility and digital preservation. Replies will be an important input for a possible Recommendation on digitisation and digital preservation. The replies will also feed into other relevant Community initiatives such as the review of the copyright framework.

In summary:

The present Communication 'i2010: digital libraries' with a focus on **cultural heritage** (2005)

An **online consultation** (2005) on questions related to digitisation, online accessibility and digital preservation. Replies will feed into a proposal for a **Recommendation** on digitisation and digital preservation (2006) and other relevant Community initiatives such as the **review of the copyright framework** (2006)

A Communication on the accessibility of **scientific information**, focusing in particular on the issue of born digital information (2006)

---

3.   ONLINE AVAILABILITY OF EUROPE'S CULTURAL HERITAGE: WHAT IS AT STAKE?

Cultural and social aspects

European libraries and archives[1] contain a wealth of material – including books, newspapers, films, photographs and maps – representing the richness of Europe's history, and its cultural and linguistic diversity. The online presence of this material from different cultures and in different languages will make it easier for citizens to appreciate their own culture heritage as well as the heritage of other European countries, and use it for study, work or leisure. It will thus contribute to complement and support the objectives of European Union action in the field of culture.

Economic aspects

Libraries and archives are major sectors of activity in terms of investments and employment. In 2001 European libraries employed 336,673 full time equivalent staff with more than 138 million registered users.[2] Their impact on the economy at large is substantial.

---

[1]   Many of the issues raised are also relevant for museums presenting their collections online.
[2]   *International library statistics: Trends and Commentary based on the Libecon data*, D. Fuegi and M. Jennings, 2004.

Digitisation of their resources could considerably increase this impact. Once digitised, Europe's cultural heritage can be a driver of networked traffic. It will be a rich source of raw material to be re-used for added-value services and products in sectors such as tourism and education. If properly preserved, the material can be used time and time again. Furthermore, digitisation efforts will have considerable spin-offs for firms developing new technologies.

4.   DIGITISATION

The quantity and diversity of material in Europe's libraries and archives is impressive.

> *The total number of books and bound periodicals (volumes) in European libraries (EU 25) was 2,533,893,879 in 2001.[3]*

The main reason for digitising this material is to make it available to users in an online environment. In some cases, however, digitisation is not primarily used to make content more accessible, but to guarantee its survival. This concerns in the first place audiovisual material, where analogue formats deteriorate with time, and precious material is lost.

> *A survey of ten major broadcasting archives found 1 million hours of film, 1.6 million hours of video recordings and 2 million hours of audio recordings. Total European holdings of broadcast material are probably 50 times larger. Most of the material is original and analogue. 70% of the material is at risk, because it is decaying, fragile or on obsolete media. Every year Europe's audiovisual archives lose 10,000s of hours of the oldest part of their collections.[4]*

At present, only a small part of European collections has been digitised. Digitisation activities exist in all the Member States, but efforts are fragmented and progress has been relatively slow. This was underlined by the announcement of the Google initiative to digitise 15 million books from four major libraries in the US and one in Europe. If realised as planned, the Google initiative by far exceeds the efforts at national level in any of the Member States. Also in other parts of the world digitisation efforts are being stepped up. In India and China there are, for example, ambitious digitisation agendas in place covering material in different languages.

A number of challenges impact the pace and efficiency of digitisation in Europe.

Financial challenges: Digitisation is labour-intensive and costly. It takes a considerable upfront investment, which in most cases goes beyond the means of the institutions holding the information. Digitising all relevant material would represent an impossible task, so choices have to be made on what is to be digitised and when.

Organisational challenges: A 'digitise once, distribute widely' strategy can benefit all the organisations involved. Duplication of effort - digitising the same works or collections several times - must be avoided. This can only be achieved by a sustained co-ordination effort at national and European level. The latter can also enhance European added-value, where the final result is more than the sum of the parts.

---

[3]   Ibid.
[4]   Survey by the IST Presto project which ended in October 2002. http://presto.joanneum.ac.at/index.asp

New ways of working are necessary to make digitisation happen. Partnerships with or sponsoring by the private sector could contribute, but this type of partnership is not yet well developed. Furthermore, investments in digitisation must be accompanied by organisational changes within the institutions concerned, including upgrading the skills of the staff involved.

Technical challenges: A key technical challenge is the need to improve digitisation techniques in order to make digitisation - for audiovisual material as well as books - more cost-efficient and affordable. For digitising written texts improved automatic book and document feeding equipment is necessary, as well as higher performance optical and intelligent character recognition for languages other than English.

> *English-language optical character recognition (OCR) systems have been tested and run on most print fonts over the past 10 to 15 years, and have been enhanced by automatic or semi-automatic spelling correction algorithms. Similar systems for other languages are less advanced, which leads to higher costs and lower quality results.*

Legal challenges: Digitisation presupposes making a copy, which can be problematic in view of intellectual property rights (IPR). Directive 2001/29/EC on the harmonisation of certain aspects of copyright and related rights in the information society[5] foresees an exception for specific acts of reproduction by publicly accessible libraries, educational establishments, museums or archives. The exception is however not mandatory and has led to different implementations in the Member States. The limited use that can be legally made of the resulting digital copies is a further disincentive for digitisation.

5. ONLINE ACCESSIBILITY

The traditional model of library services based on lending of the physical items they own is not easily translatable to the digital environment. Under current EU-law and international agreements, material resulting from digitisation can only be made available online if it is in the public domain[6] or with the explicit consent of the rightholders. Therefore a European digital library will in principle be focused on public domain material.

> *For literature, this means that only works from the early 1900s or before are available without copyright restrictions, depending on the year of death of the author. But even if works are out of copyright, the situation is not always straightforward. There may be rights attached to the different editions of a work that is itself no longer protected by copyrights, for example rights to introductions, covers and typography.*

An online library offering works beyond public domain material is not possible without a substantial change in the copyright legislation, or agreements, on a case by case basis, with the rightholders.

In this context, the clarification and transparency of the copyright status of works is very relevant. In some cases, the costs of establishing the IPR-status of a work will be higher than

---

[5]   European Parliament and Council directive of 22 May 2001, OJ L 167, 22.6.2001, p. 10.
[6]   The term public domain is used here in a narrow sense, referring to information resources which can be freely accessed and used by all, for example because copyrights have expired.

its digitisation and bringing it online. This is particularly true for so called 'orphan works' – films or books for which it is impossible or very difficult to determine who are the rightholders. These elements are highlighted in the staff working paper and the online consultation accompanying this document.

Questions of online accessibility are not limited to intellectual property rights. Putting material online does not mean it can be found easily by the user, still less that it can be searched and used. Appropriate services allowing the user to discover and work with the content are necessary. This implies structured and quality description of the content, both the collections and the items in them, and support for its use (e.g. annotation).

6. PRESERVING DIGITAL CONTENT

Making a digital copy of a book or a film does not necessarily guarantee its long-term survival. All digital material – digitised works as well as 'born digital' material – has to be maintained in order to keep it available for use. Therefore digitisation without a proper preservation strategy can turn into a wasted investment.

Digital preservation is a vital problem for the information society, where the supply of information is growing exponentially and where content is becoming more and more dynamic. At present, there is little experience with digital preservation, the legal framework is evolving, resources are scarce and the outcome of preservation efforts is uncertain. The problem deserves to be urgently addressed both by politicians and by the institutions most concerned. Its impact goes far beyond the realm of libraries and archives and concerns all organisations producing digital information and interested in maintaining its availability.

There are different causes for the loss of digital content. A first reason is the succession of generations of hardware that can render files unreadable.

> *To mark the 900th anniversary of the Domesday Book in 1985, a new multimedia edition was compiled. In 2002, it looked as if the disc had become unreadable as computers capable of reading the format had become rare. To save it, a system was developed capable of accessing the discs using emulation techniques. Interestingly, while there are difficulties accessing digital data from 1986, the original Domesday Book, now over 900 years old, can still be consulted.*

The rapid succession and obsolescence of computer programmes is another factor. Unless data are migrated to current programs or care is taken to preserve the original source code, retrieval of information may become very costly, if not impossible. This is particularly true of 'closed' data formats, for which the source code is not publicly known. The limited lifetime of digital storage devices, for example CD-ROMs is another reason for the loss of digital content.

Libraries and archives have started tackling the issues of preservation in the digital age at a limited scale. Some collaboration across borders exists, but overall action in Europe is fragmented. Within the individual Member States there is, in general, no clear policy on digital preservation. Where national preservation plans do exist they tend to concentrate on safeguarding analogue material which is at risk, rather than venturing into the area of digital materials.

Most progress is made in the area of legal deposit, with Member States introducing obligatory deposit for digital materials to ensure that the material is collected and stored by the relevant institutions. However, at present the scope of the schemes varies widely from country to country. For example, sometimes they cover dynamic online material, sometimes not.

The basic challenges for digital preservation are similar to the ones for digitisation:

Financial challenges: The real costs of long-term digital preservation are not clear. They depend on factors such as the number of migrations needed over time. It is, however, obvious that due to the limited resources available, choices have to be made as to which material should be preserved.

Organisational challenges: Choices are necessary, but who decides and who is responsible for preserving what? In an area where some basic questions are far from being answered, there is a risk of widely divergent approaches and duplication of efforts. European added-value can be found in ensuring complementarities and an exchange of good practices. Preserving digital information also requires new ways of working. This includes upgrading the skills of staff as well as more collaboration between public and private players.

Technical challenges: So far, limited research has been done on digital preservation. A major challenge is to improve its cost-efficiency and affordability. Advancing understanding of how to preserve high volumes of rapidly changing distributed information is another essential area to be addressed. Progress on related topics, such as tools for automatic analysis and for indexing will contribute to ensuring that information remains accessible and available for re-use.

Legal challenges: As digital preservation depends on copying and migration, it has to be considered in the light of IPR legislation. Other challenges are related to the legal deposit of digital material: The diverging speed and scope of the legal measures adopted by the Member States could lead to a patchwork of different rules affecting content producers with cross-border activities. At the intersection between legal deposit and IPRs, the introduction of technological protection measures to prevent copying, or of digital rights management systems restricting the access to digital material raises a whole set of new issues. Legal deposit schemes may fall short of their useful purpose if unprotected copies are not made available by those who produce the information.

7.   THE EUROPEAN RESPONSE

Organising and funding the digitisation of cultural collections and digital preservation is primarily a responsibility of the Member States. There are, however, areas where considerable European added value can be achieved, and where work at European level has started some years ago. This work will now be stepped up and complemented by new activities.

7.1.   Initiatives under way

Co-ordination at European level

The importance of digitising Europe's cultural heritage was already recognised by the Commission in the *e*Europe action plan. This gave rise, in 2001, to the Lund Principles and the corresponding Lund action plan, and to the creation of a National Representatives Group

on digitisation. As for digital preservation, a Council resolution of 2002 underlined the importance of this issue and called on the Member States to address it.[7]

For the audiovisual sector, the draft Parliament and Council recommendation on the cinematographic heritage and the competitiveness of the related industrial activities[8] is particularly relevant. It calls on the Member States, amongst other things, to allow copying for preservation purposes.

Co-funding

EU research funding has resulted in a portfolio of projects aiming at making Europe's cultural heritage more accessible through the use of new technologies. Projects supporting co-operation between Europe's national libraries have developed from exchange of cataloguing records to The European Library project (TEL). TEL has now been launched as an operational service providing a single gateway to the collections of Europe's national libraries.

While only a few projects under the research programmes have started tackling the issue of digital preservation, several of the projects specifically address digitisation. One example, in which digitisation is used as a means to preserve analogue material, is PRESTOSPACE.

> *PRESTOSPACE (2004-2007, 9 MEUR co-funding) is developing a toolkit which audiovisual archives can use to digitise audiovisual material. The tools will lead to more affordable and better quality digitisation in view of converting and preserving analogue audiovisual material.*

7.2. New or reinforced initiatives

Strategic discussion with stakeholders and reinforced co-ordination

On the basis of the questions in Annexe II to this Communication, there will be an online consultation on organisational, legal and technical challenges for digitisation, online accessibility of content and digital preservation. The results will be an important input for a possible Recommendation on digitisation and digital preservation. They will feed into other processes, such as the implementation of the Community RTD programmes and the revision of the copyrights framework. A High Level Expert Group on digital libraries will advise the Commission on how to best address the identified challenges at European level.

The Commission will increase its efforts to stimulate European co-operation on digitisation and digital preservation. The aim is to give a new impetus to the digitisation process in the Member States, to avoid duplication of effort, and to encourage take-up of good practices. The update of the Lund action-plan, scheduled under the UK presidency, can have an important function in this respect, dealing with a number of technical aspects. The use of quantitative indicators within the Lund-context would help to measure progress throughout Europe.

---

[7]   Council Resolution of 25 June 2002, Preserving tomorrow's memory: preserving digital content for future generations, OJ C 162, 6.7.2002, p. 4. The issue is also touched upon in the Commission proposal of 18.2.2005 for a Council recommendation on priority actions to increase cooperation in the field of archives of Europe, COM(2005) 53 final.

[8]   Agreed between Parliament and Council in first reading, but not yet formally adopted.

Special attention will be devoted to the work of the national and deposit libraries, predominant actors and committed to contributing to a European digital library.

Co-funding

At Community level, the Research and *e*Content*plus* programmes will be mobilised towards actions with a European interest and scale on digitisation, digital preservation and accessibility of cultural content. The Regional Funds already co-fund digitisation initiatives in some of the Member States and could further contribute to digitisation.

Applied digitisation of cultural heritage has become an objective of cooperation projects co-financed under the "Culture 2000" programme in an increasing number of cases; within the "Culture 2007" programme as proposed by the Commission, co-funding will be available for cooperation projects improving the transnational circulation of cultural works and products, including through digitisation and online access.

Research

Under the research programmes, the Commission will take the following **specific initiatives** in the area of cultural content, next to ongoing generic content-related research (search engines, language technologies, etc):

The latest call of the Information Society Technologies programme, closed on 21 September 2005, has earmarked 36 MEUR for research on new technologies for searching and retrieving cultural content (22 MEUR) and for digital preservation (14 MEUR).

Under the seventh Framework Programme, the Commission proposes to expand its contribution to research in the areas of digitisation, access to digital cultural content and digital preservation. Technological progress in this area is proposed, amongst others, to be stimulated through a network of Centres of Competence, which could become the cornerstone of European digitisation and preservation. The Community would contribute to these Centres by co-financing projects.

> *The Centres, selected through calls for proposals, will house the different competencies – technical and legal – needed to achieve excellence for digitisation and preservation processes. They will integrate and build on existing know-how in technology companies, universities, cultural institutions, and other relevant organisations. They will:*
>
> *- do research and support advanced technical solutions, testing and benchmarking them in practice on high volumes of resources;*
>
> *- promote the adaptation and transfer of these solutions to cultural institutions;*
>
> *- transfer knowledge to and support the development of skills in the user institutions.*

eContentplus

Within the *e*Content*plus* programme, 60 MEUR will be available in the period 2005-2008 for projects improving the accessibility and usability of European cultural and scientific content. Achieving interoperability between national digital collections and services (e.g. through

common standards) and facilitating access and use of the material in a multilingual context will be core objectives.

In summary:

Strategic discussion with stakeholders and reinforced co-ordination

- Online consultation on digitisation and digital preservation (Commission 2005)

- Establishment of High Level Expert Group (Commission 2005)

- Proposal for a Recommendation on digitisation and digital preservation (Commission 2006)

- Input for relevant initiatives such as the review of the copyright rules (2006) and the implementation of FP7 (2007)

- Update Lund action plan (including quantitative indicators) and discussion in Council on commitment MS to digitisation and digital preservation (MS/Commission 2005)

- Catalyse and stimulate work of the national libraries (Commission/national libraries 2005-2007)

Co-funding

- 36 MEUR in call 5 FP6 (2005)

- Increased research effort on digitisation, digital preservation and access to cultural content in FP7 (2007)

- Centres of Competence on digitisation and preservation in FP7 (2007)

- 60 MEUR in eContentplus (2005-2008)

---

8.   CONCLUSION

Information technologies have the potential to make Europe's cultural and scientific heritage visible and available for present and future use. This requires a common effort at the level of the Member States - whose efforts on digitisation and digital preservation are at present very much fragmented - and at Community level. A successful response to the challenges involved can speed up digitisation, increase the accessibility of information and ensure the long term preservation of digital content. The Commission invites the European Parliament and the Council to endorse the approach presented and Member States to join the efforts at Union level to make digital libraries a reality across Europe.