ENABLING

PERSISTENT AND SUSTAINABLE

DIGITAL CULTURAL HERITAGE

IN EUROPE

EUROPEAN ACTIONS FOR SUSTAINABILITY AND PRESERVATION

The "Firenze agenda" working group
on long-term digital memory preservation

2. POSITION PAPER

By Hans Hofman, Nationaal Archief, The Netherlands

ABSTRACT

…

Introduction

The diversity of European cultures is most apparent in the area of cultural heritage and accumulated in many information resources. By bringing these sources together in a virtual environment their visibility, availability, and accessibility to a broad audience will be increased tremendously. It is therefore relevant to think about how to achieve an added value at a European level. Achieving that will have a significant impact on European co-operation within the domain of a digital cultural area by making actions more coherent, structured and visible. It regards many different areas such as digitisation, contextualisation, creation of digital repositories and digital libraries and the development of a common cultural knowledge infrastructure. This shared environment -to-be will allow easy and equal access to cultural heritage for all European citizens.

In this report, produced under the Netherlands 2004 Presidency of the EU, an overview will be given of the developments so far and the current situation with respect to digital preservation or persistence of digital information resources. During the last decade, many initiatives and projects in this field have been carried out and are still being conducted at the moment, funded at national and European level. The results are mostly reports or guidelines, sometimes tools or prototypes. These projects were and are based upon action lines defined by the European Commission with the goal to stimulate thought and to promote experiences with permanent access to digital information, application of new technologies etc. A certain level of maturity has already been reached. It is therefore opportune to rethink the objectives: where do we stand and where do we want to go, taking into account the long term sustainability of both the resources *and* the services.

By taking stock of what happened so far and analysing the current situation, it should be possible to define how to proceed, achieve and define targets regarding the long term preservation that will contribute to the common objective of a common digital cultural area or a, referring to the leitmotiv introduced by the Netherlands Presidency, a 'continuum of digital heritage'.

In the following paragraphs a discussion of persistence related issues will be undertaken to identify relevant activities and to put them into perspective.

A vision

In the Fall of 2003, the Dutch national representative for the "Lund Action Plan on Coordination of Digitisation Programmes and Policies" presented in Parma a vision on possible next steps after the Lund Action Plan[1]. The vision envisages a 'European Continuum of Digital Heritage' that can be accessed any time from any place by all European citizens. Existing barriers that we all know so perfectly well are relieved. Such a continuum should support the mobility of knowledge and information, the exchange of cultural ideas and traditions, and the visibility of the variety of European cultural traditions. The multidimensionality and variety of cultures within the EU will be brought together in the virtual space of the World Wide Web. It will open up new opportunities and new connections, and it will help a better understanding of each other.

Although much work is done already, it will require a lot of effort and time to achieve the implementation of the interesting vision. How do existing initiatives fit into the big picture? What is needed to get a better co-ordination and who will be responsible for what? Many different aspects have to be taken into account in getting there, including

- persistence
- interoperability
- roles and responsibilities
- funding mechanisms
- standardisation.

It will be necessary to set priorities and ensure consistency between the activities undertaken under each of them. The results should not only be concepts, but also practical tools. Investments are only worthwhile if continuous availability of the digitised and digital-born resources within the continuum is guaranteed. However, the very nature of these resources is and will always be fragile and there are still no adequate strategies for their long term maintenance. Therefore one of the priorities has to be on persistence and sustainability. This interest area is also very dependent on international collaboration, because of the challenges to overcome. As stated also in the vision by the Dutch National Representative, the European level will be best suited to co-ordinate the efforts in this area. The interest in coherence and consistency is most significant there, but it can also promote and enhance the cross-domain and cross-organisational collaboration best. Important criteria for users, such as trust and reliability find their ultimate foundation in continuity and persistence. That should be shown not only in the availability of resources, but also in understandable and reliable access and search services. In the end, the European citizen should benefit from it. The persistence of cultural

---

[1] The Lund Action Plan (http://www.cordis.lu/ist/directorate_e/digicult/lund_ap_browse.htm) is expected to finish in the end of 2005.

heritage as societal memory touches the identity of individuals, the member states as well as the European community as a whole.

Current situation

So far preservation or persistence has not been a really big issue in the co-ordination of digitisation activities in Europe. The Lund Principles on Coordination of Digitisation Programmes and Policies mention among many other issues the need "to increase awareness of long-term preservation issues" and the need for "guidelines for digital preservation and content longevity". The Lund Action Plan has ten objectives of which 4b ("Sustainable access to content") sets the objective: "ensure digitised cultural and scientific content is available over time". As such its focus is on analogue material and does not include the growing amount of digital-born resources. The MINERVA project supports the co-ordination efforts of Lund and helped to improve the digitisation process and the management of it, e.g. through benchmarking instruments and handbooks. Yet the approach regarding long term preservation has been fragmented, practically invisible, and primarily focussed on analogue resources.

Under the Spanish EU Presidency in 2002, a Council Resolution[2] on "Preserving tomorrow's memory – preserving digital memory for future generations" also underlined the importance of preservation of digital-born assets and set out a number of measures for the Commission and Member States to investigate. Another example of attention that was given to the subject is the workshop on the preservation of digital memory held on 12 December 2002 in Copenhagen under the Danish Presidency. But like these two examples, most activities still have a rather incidental character.

A special position in all this is taken by the experts group that was established under the Italian EU Presidency during the conference 'The Future of Digital Memory and Cultural Heritage' in Florence in autumn 2003. This initiative clusters knowledge and projects around several long term preservation objectives laid down in the so-called Firenze agenda. It is a further development based upon the Council Resolution of 25 June 2002. The Firenze Agenda objectives include raising awareness, establishing co-operation, exchanging experiences and best practices, developing a research agenda and training programmes, and developing long-term preservation policies. The projects involved in the agenda include ERPANET, MINERVA, PRESTOSPACE, DigiCULT, and DELOS. It also makes clear connection to UNESCO work on digital preservation.

The Firenze Agenda was submitted for endorsement in Parma on 19 November 2003 to the official meeting of the National Representatives Group (NRG) for Lund, as an extension of the current Lund action plan, and in order to make a visible and practical claim with respect to digital preservation. The NRG then established as one of its five priorities for the second semester 2003 to "carry on the activity on preservation of digital memory". On the same day the NRG adopted the "Charter of Parma", a strategic document promoting all the NRG activities, but this document does not address the issue of digital preservation.

---

[2] 25 June 2002 (2002/C162/02)

So the interest in the issue does exist, but has not yet had a significant practical impact, such as on the proposals for projects or activities related to digital objects and services. So far, work at the European level on digital preservation is mainly taking place in the area of building or increasing awareness and of supporting exchange of information and experiences, like ERPANET.

Most European projects are rather limited in scope and scale, though useful, like PRESTOSPACE (focusing on the preservation of digital audio-visual contents). Many other projects are not specifically focusing on digital preservation but may include it (such as the research being done in the field of digital libraries (DELOS2)), or are monitoring the latest developments in technologies including preservation (DigiCULT).

The DLM-Forum (a European Economic Interest Group) is working on what is called a new version of the 'black book', describing the situation in the archives of the member states and including a chapter addressing long term preservation of electronic archives, also with respect to archives in the enlarged European Union. Its recommendations will be translated in a list of concrete actions, one of them addressing digital signatures and indicating that standards and specifications will be developed. An operational model, however, still has to be defined, and will be limited to archives, archival records and government administrations.

The BRICKS-project takes a much broader view. It covers the sectors of libraries, museums and archives (although it does not seem to have a strong commitment from these sectors) and therefore surely may give an impulse to the realisation of the "European Continuum of Digital Heritage". Sustainability is one of its main areas of work, however this is interpreted mostly as the "management of the practical outcomes of the project in order to make it a self-sustaining, profitable European asset". More is needed for the actual issue of digital preservation and persistence. Nevertheless, the project is positioned by the European Commission as an important project in this area.

The Sixth Framework (2002-2006) includes six research areas in the cultural heritage domain, of which "Digital Libraries" may be best serving projects enabling the "European Continuum" leitmotiv. In general, however, digital preservation of cultural heritage is not identified as important separate issue in this European research framework.

In October 2003 the General Conference of UNESCO adopted the "Charter on the Preservation of the Digital Heritage". It includes sections on "The digital heritage as a common heritage" and "Guarding against loss of heritage". The section called "Measures required" contains the following four very relevant articles:

- Developing strategies and policies: *"The co-operation of holders of copyright and related rights, and other stakeholders, (…) will facilitate this."*
- Selecting what should be kept: *"Born digital materials should clearly be given priority."*

- Protecting the digital heritage: *"Member states need appropriate legal and institutional frameworks (…). Access (…) without causing prejudice to their normal exploitation. Legal and technical frameworks for authenticity."*
- Preserving cultural heritage: *"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."*

In the final section on "Responsibilities" the Charter claims "it is necessary to reinforce international co-operation".

Next, or rather preliminary and parallel to the international and European level, much work on digital preservation is done at the national and at institutional level. At the national level, initiatives like the Digital Preservation Coalition (DPC), the Digital Curation Centre (DCC) in the United Kingdom and NESTOR in Germany are good examples of the fact that the issue of digital preservation is taken up seriously in these countries. Both DPC and NESTOR are network building initiatives, bringing together stakeholders and setting a framework for co-operation and exchange of information and experiences. The recently established DCC intends to support UK institutions to store, manage and preserve digital data to ensure their enhancement and their continuing long-term use.

At the institutional level the Koninklijke Biblitoheek in the Netherlands with its

"e-depot" and several National Archives (e.g. in Denmark, Sweden, UK) have actually implemented initial versions of facilities that support longer term preservation. The recent Dutch "Testbed Digitale Bewaring", a joint venture of the Ministry of the Interior and the National Archives, conducted experiments to identify the best current options for maintaining records in a more persistent and usable way.

Research in the past has offered some important reference models that also support sustainability, such as the Open Archival Information System (OAIS), or some possible approaches and practical guidelines from projects as CEDARS, but it is still a long way to really establishing an infrastructure that is persistent and includes both preservation and access to digital resources, either digitised or digital-born.

Software suppliers provide software applications and tools that only accommodate part of the requirements, if it is even clear yet what these requirements are in principal! There is for example a lack of standards, but at the same time there are too many. It is a widely accepted notion that persistence starts at the moment of creation of digital objects and even before in the design stage of the systems that enable it. That requires a kind of chain management of the digital objects from their creation to wherever they will be maintained. It means the perspectives will be no longer limited to the immediate interest of the responsible party (e.g. the publisher, the records creating organisation or the custodial body), but should be based upon a shared vision and broader interest. All key players in the chain are involved and should subscribe to the same objective that goes beyond there own direct interest. Such a change entails not only a new paradigm, but also takes a long time to implement and institutionalise.

So, as a preliminary conclusion, the current situation still seems far from working towards a much desired shared European digital cultural area. If anything already exists, it is a patchwork of initiatives that may overlap, are not necessarily in line with or at best coincidentally complement each other. The existing networks are already of some use but have not yet reached all relevant parties such as private sector companies or certain management levels. Yet, they raise awareness of what is happening in different places and support the circulation of practical experiences and new insights. The community or communities involved are still rather small though and limited to direct interest parties: the convinced talking to the convinced. Expansion is needed to get the attention and the support of all levels of users and creators of digital information, decision makers and of funding agencies. Long term thinking is not a built in human characteristic and therefore inciting and suitable approaches will be needed to attract the necessary attention. Possible triggers may be awareness of the value of digital resources for society and economy and of the consequences and risks if they are lost because of negligence.

As indicated earlier, persistence is not limited to the cultural heritage sector. The cultural heritage characteristic is only one possible link in a whole chain of digital information. Thinking strategically about persistence will always need a broad perspective including connections with all different kind of domains. Libraries for instance are connected to and dependent on publishers, while archives are for an important part embedded in government practice. As a matter of fact, producers of digital information are more often outside then inside the cultural heritage domain. Even, most of the times they won't be aware that the digital things they produce could come to belong to the cultural heritage domain after a certain amount of time and therefore should be carefully kept after! This situation requires special attention to prevent loss of information and incompatibility when resources are crossing boundaries between domains. The emergence of electronic service delivery in and outside government agencies, publishing companies etc. through the World Wide Web entails even closer connections between producers and memory organisations. This growing interconnectivity means a considerable extension of the persistence issues involved, of the need to collaborate with other partners and even of the re-engineering of business processes. That certainly is a challenge, because in the past those aspects were always rather limited and much easier to control. Many organisations also may fear a loss of identity when aware of their incapability of controlling the whole information chain. Strategies are needed to influence (or to find common ground with) other organisations that may not have a similar direct interest in long term access and preservation. In the end this might even entail a re-arrangement of responsibilities.

In general, however, traditional structures and existing division of responsibilities are very persistent, thus a potential obstacle to new and innovative approaches in meeting the challenges facing us. Most co-ordinated work is still very much sector bound. Examples are The European Library (TEL) and DELOS. The activities initiated by the DLM Forum are very

much from the archival perspective, i.e. the traditional mandate of preserving records appraised as having archival value.

A possible example of crossing boundaries may be the fact that archives are trying to influence or to get grip on what is happening in the government organisations in order to ensure the survival of digital records for cultural heritage. That may turn out to be a rather difficult task. Unless there is some common interest in preserving their records, the primary interest in those organisations, however, will be with their immediate responsibilities. The main pillar for transfer of records therefore is still archival legislation.

A project like ERPANET and also the Firenze Agenda initiative more or less encompass scope beyond sectors. They bring together people from all interested different sectors (archives, museums, libraries, industry, consultancy, private companies).

Reactions from European experts in the field of preservation and persistence show that institutions and people still struggle with many preservation issues and lack a framework for positioning those in a useful approach. Based on the discussions in the ERPANET workshops and seminars the observation can be made that organisations need direction for implementing adequate digital information and records management including preservation strategies. Digital preservation, unlike analogue information, has to be taken into account at the moment of developing information and records management policies and in designing supporting infrastructures and systems.

To enable exchange and cross-domain searching and use of information resources, focus on interoperability will be crucial. Interoperability in itself is very dependent on persistence. In cases where metadata formats or ontologies are not maintained and kept consistent through time, this will undermine the viability, quality and usability of the cultural area. Consequently not only a continuous management with a long term view on digital objects and services is needed, but also the underlying preservation actions in order to keep the cultural heritage area vital and the resources available. These challenges of digital preservation and persistence are so huge and complicated that it is impossible for most organisations and countries to solve them on their own. Only very large organisations, such as national libraries, will be capable to work on approaches that are significant on much larger scale. In fact, most organisations and national situations require a concerted approach in combination with clear intermediate structures: at European level a co-ordinating framework must be developed and agreed upon, while the actual work on access and preservation must be done at the country and institutional level.

Cross-sector and cross-border collaboration is hesitantly emerging, creating its own dynamics. The transition from rather closed and distinct communities with specific mandates, to open, connected and cross-sector interwoven communities is not easy. In the end it will be the ultimate stakeholder, the user, who is confronted with the results, and therefore has to be

taken into account in this decision-making process as well as in the design and implementation of the envisioned digital cultural area.

A broad range of issues, not exclusive for or necessarily limited to persistence rise from this situation:

- How to avoid fragmentation without restricting flexibility of institutions/ allowing for own solutions?
- Different levels of expertise/conditions: how to align them?
- Collaboration. There are different sectors, how to bridge differences between them and even within sectors? Need to involve IT-industry or other private sectors with an interest in digital preservation (both as suppliers and as stakeholders). They should become active partners in doing research, setting standards, and developing tools.
- The need to understand the requirements of the users
- The need for cultural change within institutions.
- The need for (other) expertise and therefore of continuing training.
- Will there be a need for redefining roles (e.g. of national institutions)?
- Need for standards.
- Funding (more specifically how to organise it).

All this raises the question whether a more structured approach is desirable to focus the rather limited expertise in the area of preservation and curation in a more cost-effective way. Both at the international (UNESCO) and the European level several resolutions were adopted upon and established indicating that preservation is an important issue, that it should help maintain our cultural heritage, so future generations will also be able to know and use it. There is however a big gap between those rather ephemeral though politically appealing statements and every day practice where institutions and projects are trying to achieve some concrete results. That gap has to be bridged, if we want to be successful in building a persistent memory that can be used throughout Europe and on an ongoing basis. A conclusion is that more focus and co-ordination are needed and that the scope has to be broadened to both digitised and digital-born resources. How to achieve that?

Options for how to proceed

The previous sections show that persistence, including maintenance, preservation and curation, is not yet really integrated into all activities with respect to creating a new digital order, but is defined and seen mostly as a separate issue that is dealt with separately. The consequence of this way of thinking is not only that it will be considered as an issue only for the immediate interested parties, such as memory organisations, but more importantly in the end will be a liability or risk for the overall sustainability of access to digital resources and services. It is the responsibility of all direct in indirect stakeholders, national governments and inter-governmental cooperation, to prevent this from happening. It is required to make persistence an inextricable aspect of all developments and activities, either structural or on project basis, with respect to creating, managing and making available digital resources. Only then, there is a chance to establish a digital environment that proves not only to be reliable and successful for users, but also to be cost-effective and sustainable.

As indicated already the current Lund Action Plan does not really address preservation and persistence, nor does it include digital-born sources, an increasingly important area. A possible successor to the Lund Action Plan when the current one finishes in 2005 does need a higher level of ambition with a broader and more flexible scope. It should at least imply the inclusion of persistence as a natural aspect of anything that deals with digital resources and digitisation practice, even when an analogue alternative will always be available. That also includes the services through which these resources will be made available, the knowledge needed to create and manage them, the infrastructure that supports access, exchange and maintenance, the metadata structures for retrievability, reliability, and usability, as well as the preservation and curation of analogue objects.

Other things to be supported by a future coordination initiative are the preparation and development of common preservation policies, fostering communication across Europe, promoting a European division of labour and expertise, coordination of project efforts, and organising funding as well as identifying sponsors or champions and include more effectively the private and commercial sector.

The current situation with respect to managing digital objects and, as a consequence, the activities and expertise on digital preservation in each of the European countries is rather different. Some have just begun, while others have already gone a long way and built up substantial experience, both with respect to digitised and to digital-born sources. In some countries and organisations a core infrastructure may be available to manage, maintain and preserve both these kinds of digital objects; in others thinking about it still has to start. Some practical tools exist, e.g. to harvest resources or to store documents in a sustainable format and with their proper metadata, but far more are needed. Preferably such tools should be

open source, so they will be easily available throughout Europe at affordable cost and without copyright restrictions.

What preliminary conclusions may be drawn from all this?

1) Persistence is still defined and seen too much as a separate issue, only of importance for certain institutions responsible for long term preservation. As a consequence it becomes a liability not only to initiatives aiming at building a European digital cultural memory, but also in general to any organisation that manages its intellectual capital and information assets in a digital form.

2) In addressing the issue of persistence the scope of activities should not be limited to digitising analogue objects, but should encompass all digital objects both digitised and digital-born.

3) Despite the resolutions and charters decided upon by European Council of ministers, the General Assembly of UNESCO and the NRG, and despite the fact that they have raised a lot of awareness, the consequences have not yet been integrated or formulated into concrete action plans nor have they taken it beyond the level of a stand-alone topic. As long as the practical integration of persistence into our daily economic, social, cultural and policy issues is not achieved, it will be difficult to raise it to make it politically appealing and interesting for funding.

4) The current situation shows that most of the initiatives are sector bound or focusing on a certain type of objects (e.g. audio or video files), on a temporary basis, or in an experimental phase. As a result, efforts are fragmented, their impact too limited and their final contribution to cost-effective solutions or approaches unclear. The challenge will be, and this is where the European added value comes in, how to build a critical mass of convergent activities, and of collaborating institutions and people, in such way that it will become a self-sustainable mechanism/entity.

5) There is a great need for coordination: preparation and development of common preservation policies, fostering communication across Europe, promoting a European division of labour and expertise, fusing of project efforts, organising funding as well as identifying sponsors or champions and include more effectively the private and commercial sector.

6)   There is also an ongoing need for more practical coordination with respect to the exchange of information and experiences, practical tools, practical 'how to'-guidelines, identifying and/or establishing common standards.

6)   The interests and role of the user need to be clearly identified and defined, and should be included in any further action plan.

How to organise

If the above conclusions are shared, there is especially a need for better European coordination. This requires some structuring mechanisms. Based upon the objective of a European Area of digital heritage, it is clear that continuing the current situation will require an enormous amount of effort and time and still will lead to a very fragmented picture with respect to access and persistence of European digital cultural resources. The opposite, a centralised European approach, will also be difficult for political and organisational reasons. It is not desirable that any European body will determine the national or institutional priorities.

Furthermore, it will be very difficult to organise and monitor all activities that will be needed to achieve the objectives. The ultimate responsibility for persistence of digital objects lies with the institutions that created and/or maintain them, but the range of problems to be tackled is such that it will be difficult if not impossible for institutions to solve them on their own. The solution may be found in a more mediate approach, that tries to combine both views, the individual and the centralised one. Coordination will be necessary and may be achieved by using the 'model' of the Digital Preservation Coalition in the UK, that brings together the key players in the area of digital preservation and tries to support, co-ordinate, and encourage related activities. In Germany the recent NESTOR project fulfils a similar role. It could be a model for every country in Europe. However, the focus on preservation should be inclusive and not exclusive as is the case now in the UK and Germany. Persistence has to be embedded in the creation and management of the resources as well as in the development and provision of services that allow people to use them. This should not preclude initiatives specifically focused on digital preservation though. They still will be necessary in order to make progress in this area.

Priorities lay in the field of
a)   developing practical tools such as metadata extraction tools, ingest tools, conversion-to-open-standard tools;
b)   identification (and may be development) of standards that support persistence as well as interoperability of digital objects;
c)   promoting the establishment of national networks, so the basis will be strengthened and broadened;
d)   creating an open 'market place' where organisations and people can discuss ideas, upload and find (if possible open source) practical tools, identify interesting partner institutions or projects and so on;
e)   establishing funding models.

One aspect that deserves attention is the connection with other international developments, since preservation is not solitary European issue. What for instance should Europe do by itself and what can be learned from research and developments elsewhere? In

the USA and Australia for instance many efforts in the same area are going on and should be taken into account when developing programmes or research agendas. Collaboration with leading projects in those countries has to be sought to increase the impact and synergy of research. The DELOS-NSF research agenda on digital preservation already provides an example. What ways of co-operation are possible or desirable and should this be coordinated?

Recommendations

1. Integrate the concept of persistence as an ongoing and natural aspect of the digital cultural heritage continuum and, as a consequence, make sure the integration of aspect of all activities contributing to that European continuum. Persistence should be a criterion for all future projects.

2. Given the different situations in institutions and countries with respect to digital preservation it is recommended to develop a framework with different levels of ambition, that are consistent among each other, and will serve as a structured mechanism to stimulate and coordinate efforts to achieve the objectives under 1.

3. Each country should develop a DPC-like networking body to stimulate, coordinate and support national and cross-sector initiatives in the area of digital preservation and curation.

4. A high-level political body should be established for co-ordination issues at a European level. The current NRG for the Lund Action Plan consisting of representatives of ministries of culture may be well the right body when its mandates and responsibilities are better determined. One of its possible tasks may also be to identify when co-ordination with other international political developments will be necessary.

5. Collaborative and concerted activities or focus areas with respect to persistence and digital preservation will still be necessary. Priority areas are the development of practical tools (e.g. for metadata extraction, ingest, support of preservation strategies) if possible as open source, identification or further development of standards (e.g. metadata, storage or file formats, functional requirements for software applications) and training courses for professional development. Those collaborative activities may take place at European level, especially for supporting and stimulating exchange of

information, knowledge-sharing, and practical experiences and for organising of training seminars and courses (e.g. together with academic institutions).

6. Promoting a European division of labour and expertise, stimulate the organisation of funding or economic models and attarct more effectively the private and commercial sector.

ì