Strategies and cooperative approaches
of long-term preservation in the European Union

The Alliance for Permanent Access

Hans Jansen
Director R&D, National Library of the Netherlands
Frankfurt, 20 April 2007

Digital media have become the dominant way in which we create, shape and exchange information. Government, businesses, research organisations and memory institutions as well as individuals have become completely dependent on digital information.
As you all know, this dependence on digital information presents us with a number of major risks because of the many unresolved challenges in the long-term management, access and preservation of this information.

There is a growing realisation that the answer to these challenges calls for coordinated approaches on both national and international level. Digital preservation is too big an issue for individual institutions or even sectors to address on their own. The required effort is simply not feasible and international exchange and collaboration are therefore essential.

Over the last five years we have witnessed a series of workshops and conferences organised under different EU-presidencies, devoted to long-term preservation of digital information. During the first years the main focus was on the libraries and archives sector. This initial focus on the heritage sector is obvious, since these so-called memory institutions were the first to be confronted with digital and digitised objects.

In Europe several organisations, like the British Library and the Koninklijke Bibliotheek, the national library of the Netherlands, have taken the lead and put in a large effort in raising awareness and working on solutions that may be relevant to long-term preservation and access.

In 2004 the KB organised the conference *Permanent Access to the Records of Science* within the framework of the Dutch presidency of the European Union. In contrast with its

predecessors this conference did not concentrate merely on the cultural heritage sector, but tried to bring together the worlds of culture and science.

This link is obvious. The rate at which the scientific world has become digital also poses a threat to the access of research information. We all know that science depends entirely on the knowledge gained in the past to further advance. Therefore, the prospect of losing the digital records of science is very alarming.

The conference at the KB enabled government representatives, experts from the cultural heritage and science sectors, and the European Commission to discuss the issue of permanent access. The participants agreed to a series of conclusions and recommendations. And in order to remain focused, the KB was asked to set up a high level Task Force Permanent Access.

This Task Force started its work in 2005. It argued that addressing the digital challenges would require the creation of a sustainable organisational infrastructure for permanent access within Europe.
In order to achieve this, the Task Force developed a proposal for a long-term R&D programme in the field of digital preservation. This programme was submitted to the European Commission.

In addition, the Task Force recommended the formation of a strong European alliance, consisting of parties from the scientific, library and publishers' world. This alliance should become the driving force in the efforts to realise a European organisational infrastructure.

This was the end of 2005. Since then, a lot has happened.

As set out in the i2010 Strategy of the European Commission, digital preservation is a matter of concern to each country in Europe. Within several European countries national preservation coalitions have been established: the Digital Preservation Coalition in the UK, NESTOR here in Germany, and PIN, (Pérennisation des Informations Numériques) in France. In the Netherlands the Dutch National Coalition for Digital Preservation will start its work next month. Later on this afternoon we will be further informed about these and other national initiatives.

Next to these national initiatives, a number of European projects supported by the Commission were set up:

- the PLANETS project, by far the largest R&D project on digital preservation solutions worldwide;

- the DPE project (Digital Preservation Europe), which concentrates on fostering collaboration and synergies between many existing national initiatives across the Europe;

- and CASPAR, which is focused on a number of technical challenges, like the implementation and further extension of the OAIS reference model

So, during the last few years numerous projects have been set up and carried out at the European level. This development is of course very valuable. But we all know that projects by definition have a limited focus and time frame.

The big challenge for the coming years will therefore be to establish a viable and sustainable European infrastructure for access to the records of science, which means research data as well as scientific publications. The FP7 Work Programme Research Infrastructures shows that the Commission shares this view and is willing to support proposals.

The conference flyer we all have received sets out as one of the goals of this conference:

..to sound out the feasibility of a European alliance of the individual national alliances devoted to preventing digital data loss.

With this goal in mind I want to go back to the recommendation of the Task Force to create a European alliance as the driving force in the efforts to realise a European organisational infrastructure. However, the Alliance does not take national alliances as its preferred starting point. I'll come to that in a minute.

At this moment we have a substantial number of founding partners for the European Alliance, which are all major stakeholders from the worlds of science, libraries and archives, and publishing.

Among the confirmed members are: The European Organization for Nuclear Research (CERN), European Space Agency (ESA), European Science Foundation, CCLRC (now the

Science and Technology Facilities Council) in the UK, JISC, the national libraries of the UK and the Netherlands, the STM Publishers Association, the National Archives of Sweden, and the national preservation coalitions of the UK, Germany, France. The Dutch coalition will follow soon.

What does the Alliance want to achieve within Europe?

In its position paper it is stated that:

"The Alliance is developing a shared vision and framework for a sustainable virtual infrastructure for permanent access to scientific information." This will include all disciplines from physical, biological, or environmental sciences, to social sciences and humanities.

The Alliance advocates a pragmatic approach both to the idea of an infrastructure that serves all scientific fields. And to the way to accomplish this.

The basic notion is the assumption that the repository infrastructure will be different for each scientific "community", such as particle physics, astronomy and space science, life sciences, earth and environmental sciences, or social sciences.

The repositories will be part and parcel of organisations that exist within a particular community. In particle physics for example CERN will definitely be one of the main repositories.

But all these communities must agree on certain standards to make their repositories interoperable.

The repositories will also benefit from a number of common facilities: from common R&D activities or a framework that offers technical tools, to a single accreditation body for guaranteeing the quality of companies or organisations that verify whether a repository meets the requirements.

The repositories themselves are physical, but the overall infrastructure is virtual: there is no need for a central governing body, though some kind of competence centre to support the common goals makes sense.

So how does the Alliance want to achieve these goals?

The key to the approach of the Alliance is the engagement with the various scientific "communities" to help them build their part of the infrastructure.

The practical experience of the Alliance members is very important here. The Alliance can also establish the connections to a variety of parties, including policy makers and funding bodies. This way, communities will be assisted in some crucial preservation activities, like:

- identifying which digital information to preserve and for how long;
- identifying key repositories;
- establishing a scheme for metadata suitable for the community;
- organising test beds, etc.

The pragmatic approach of the Alliance is also shown by the choice to start with three or four communities that are already reasonably well-organised.

The Alliance believes that this will generate sufficient momentum for others to follow. And in developing common tools for this small number of initial communities, it is easy to check for their wider suitability.

In want to finish by stressing one final point.

The actions that are needed within Europe are so closely connected to the process of science that the stakeholders directly involved in that process must be in the driver's seat.

But next to that it is clear that ensuring the preservation and accessibility of data and documents produced in science is an important policy objective both on a national and European level.

Therefore the Alliance is very willing to engage in discussions on the implementation of such policies with both national governments and the European Commission.

I hope this presentation has convinced you not just of the feasibility of a European Alliance, but of the urgent need for such an Alliance as the driving force behind the creation of a European infrastructure for the permanent access to the records of science.