SHAPING COURSE. The development of the strategy for the e-Depot of the Koninklijke Bibliotheek, National Library of the Netherlands, in a national and international context

Els van Eijck van Heslinga, Koninklijke Bibliotheek, National Library of the Netherlands, The Hague, The Netherlands

els.vaneijck@kb.nl

Abstract

Operating an international archive for digital information resources by a relatively small organization as the Koninklijke Bibliotheek (KB) asks for a firm foundation of its policy. Continuing research and development efforts to secure permanent preservation and access to electronic information on a national and international scale require substantial financial, technical and staffing commitments. Cooperation in sharing resources and knowledge is called for.

In the national context of the Netherlands the KB recently took the initiative to build the National Coalition for Digital Preservation. All relevant sectors of Dutch society are represented. Its aim is to develop a national infrastructure and to assign clear responsibilities.

In the European arena the KB, together with the British Library, started the European Alliance for Permanent Access to the Records of Science. The Alliance works towards the widespread and joint responsibility for maintaining the records of science by all major stakeholders, including national governments and the EU.

In the global context the KB developed the concept of the 'Safe Places Network'. This Network will include a limited number of institutions with certified digital archives for scientific publications, sharing responsibility for complete, world-wide coverage and allocate tasks accordingly.

Introduction

Our digital information is worldwide under immediate threat from deterioration, damage or obsolescence as media decay and hardware platforms evolve. The question is who is doing what, when and how, to tackle the problem. I am quite sure that most of you, the participants of this conference, are thinking in terms of technical solutions. But my hypothesis today will be that the technical angle of incidence is just one part of our answer. I will argue that we should broaden our perspective and start acting also in other directions.

The incentive for making out a case for this statement stems from the new phase our electronic depot, the so-called e-Depot, is entering. So for this presentation I depart from my own library, the Koninklijke Bibliotheek, National Library of the Netherlands[1].I do realize that my view, this specific lookout from our library, is biased. After all the Netherlands is a small country, based in Western Europe and our library is also small, compared to many others in the Western hemisphere.

---

[1] http://www.kb.nl

The first stage

As always with new developments, the first stage of the history of the permanent preservation movement is a somewhat diffuse picture. The first institutions to realize that there could be a problem, were libraries, archives and space research institutions. Libraries and archives are committed to the long term access to published material and to source material. In the beginning of the 90's the first digital publications and sources started to come in. And for instance NASA discovered that their data from the first landing on the moon could not be retrieved. In this first stage institutions, national governments and international organizations committed themselves mainly to possible technical solutions.

At the KB a small unit for Research & Development began –together with IBM and supported by project money from the European Commission- to build a long-term storage for e-journals. Our operational e-Depot started in 2003.[2] The mission is to ensure permanent access to digital information resources in a national and international context. Directly linked with the operational e-Depot is a R&D team, whose main target it is to update the system and to create and maintain intrinsically international standards. We participate in national and international projects, always keeping an eye on technical developments in the software and hardware. The KB considers the test-audit of its e-Depot by the American Centre for Research Libraries in 2006 with a positive result as a firm step in the process of international certification of trusted digital repositories.[3]

In line with the international nature of information provision, specifically for the records of science, the KB opened up its e-Depot to international publishers worldwide from the start, next to archiving and preserving all output with a Dutch imprint.

We contracted twelve international academic publishers to store their digital publications and digitized back files in our e-Depot. Today the e-Depot numbers about ten million articles. We estimate to ingest 400.000 articles per year in the coming years, the 'normal' output in the western world. But perhaps this will be more, if we may believe the expectations. Especially in Asia the number of new academic journals will grow rapidly. [4]

However, the term 'digital information resources' is not identical with e-journals only. Already Springer and Elsevier, the two biggest publishers in the world, began to publish e-books. These involve not only reference works, bibliographies and reports, but also monographs. Furthermore we have to take into account that researchers are already communicating in more informal virtual ways and that scholarly dissemination in future possibly, partly, will no longer be the prerogative of official publishers. And –next to the academic output- we have to consider also the heritage from other worlds, for example websites and digitized material from cultural collections. The KB has decided on a strategy concerning the .nl domain and Dutch websites. We will start in 2008 with the collecting of about 3000 Dutch websites per year, adopting a strict selection-policy.

In our e-Depot we also have to preserve millions of masters, coming from national projects like the Dutch journals-project to digitize all Dutch newspapers, or the Memory of the Netherlands with many cultural heritage collections, or the digitized records and minutes of Dutch Parliament since 1814, all accessible via website-gateways.

---

[2] https://www.kb.nl/dnp/e-depot/e-depot-en.html. See also Erik Oltmans and Hilde van Wijngaarden, The KB e-Depot digital archiving policy (Library Hi Tech, Vol.24 No 4, 2006) Available at: www. Emeraldinsight.com/0737-8831.htm.
[3] Draft KB Final Report March 21, 2007/Part of test of RLG/NARA Audit Checklist fot the Certification of Trustworthy Digital Repositories
[4] ISI Journal Citation Reports Science Edition 1995-2005, Thomson ISI.

In this first phase of developing digital repositories, working out operational solutions went hand in hand with the setting up of an organizational concept. At the KB we decided that in essence ingesting electronic documents is not different from the existing deposit workflow in the paper-world. That is the reason why the operational e-Depot is just another branch of the Acquisitions & Processing Division. The R&D staff for Digital Preservation belongs to the R&D Division, but is housed nearby their colleagues working for the operational e-Depot. We are convinced that the close alliance between the operational tasks and R&D efforts will always be necessary.

Financial sustainability was also part of the agenda in the first years. While building up experience we worked out a balance between operational costs and costs for Research & Development. We are however aware of the fact that our international scope asks for a rethinking of our funding model. This holds true for the structural maintenance of our international e-Depot and for capital outlay in the area of R & D.

The second phase

So far, so good. We now live in 2007. If we look back, the digital world has evolved very fast and it will continue to do that. I just took our library as an example of an organization that developed a working repository in an early stage: a natural part of its public responsibility as the national library of the Netherlands. Others did so too. One of the characteristics of a period of first developments is that many organizations and institutions joined in this area, each in its own pace and without much consultation of each other[5]. But in recent years we see a growing number of initiatives to work together, specifically in projects focused on tool-building. An excellent example is PLANETS, Preservation and Long-term Access through Networked Service, a European project where archives, libraries, research institutes and technology vendors work together[6].

I think that this first phase is now slowly moving on into stage two, at least for the runners-up in our field. Next to the continuing efforts to develop new tools and standards, in this new phase we should actively start working toward a worldwide concept or policy concerning the responsibility for long term preservation of our digital information resources: a new kind of balance. I would like to quote the Lord Speaker (who happens to be a lady) in the British House : "order, order in the House!"

The question at stake here - can we work towards a clear system to determine who is responsible for what in the digital world of archiving and preservation - can only be answered if we try to establish a number of things. The first action could be a practical one, specifically targeted towards academic publications. It is to be expected that only a limited number of institutions around the world will have permanent archiving of these records of science prominent in their mission. We suggest that these institutions form a global network of so-called 'Safe Places' to ensure a more systematic and concerted approach to digital preservation of scientific information. These Network members should answer to the following requirements:

- The institution has a proven performance in the field of long-term preservation an access. Its digital preservation policy, procedures and standards are well documented. Certification is obligatory.

---

[5] See Anne R.Kenney and others, September 2006, "E-Journal Archiving Metes and Bounds.:A Survey of the Landscape, at http://www.clir.org/pubs/abstract/pub138abst.html. Also: forthcoming in *Alexandria*: Ingeborg Verheul and Hilde van Wijngaarden, "Measuring Progress: recent developments in digital preservation in the library community".

[6] http://www.planets-project.eu. Accessed 6/10/2007.

- The institution must be capable of guaranteeing long-term funding and a stable political environment.
- The institution must have clear disclosure policies with regard to the contents of the archive and its selection criteria.
- The institution must be committed to ongoing research and development in the field of long-term preservation and must be willing to collaborate and share information with other Network participants.

Of course the implementation of the Safe Places Network will be a gradual, step-by-step process. The Koninklijke Bibliotheek and the American organization Portico started cooperation in this context this spring. We invite others to join us in developing a common view on the way we should progress in the field of the long-term preservation of academic publications. We would very much appreciate specifically institutions in the non Western world to contact us.

The background for the second action is provided by the thought that digital archiving and preservation in general is a universal problem. Worldwide there are several stakeholders we can identify. The first group involves researchers, their universities and worldwide communities, and other large research-institutions. International publishers are important players, as are national libraries and archives, specialized research-libraries, national governments, international organizations, the business world, commercial developers and vendors of systems and last but not least giants like Google and Yahoo.

Let's look at each of these groups of stakeholders more closely. First the researchers, the producers of the world's records of science and their organizations. Thanks to technological possibilities, the means by which researchers can disseminate knowledge have increased enormously, from dynamic webpages to visualisations and programs of all kind. Overall they still publish in official journals. But institutional repositories of university-libraries now also preserve results of research. And in their new e-communities researchers produce all kinds of 'unofficial publications'. We could ask ourselves whether libraries should reconsider the definition of an 'official publication'? One thing we can say for sure is that individual researchers in general have no awareness of the problems we encounter when dealing with these developments. Once an electronic article or book is published, researchers trust that someone will take care of those publications. In general one could say they don't realize that the archiving and preservation of digital objects is fundamentally different from keeping a printed publication safe.[7]

We turn to the publishing world. Worldwide there exist almost 35.000 academic journals, of which 21.000 are only electronic. And – as mentioned above- there is a growing number of other publications only electronically available. [8]Also for publishers the world is changing dramatically. With the new opportunities of electronic publishing they are also confronted with the shifting of roles. Karen Hunter, Senior Vice-President of Elsevier, has put it loud and clear in presentations in 2006 and 2007[9]. Researchers, libraries and publishers have become competitors in many areas – or, if not competitors, actors in overlapping roles. As one of the areas where roles have changed, Hunter mentions archiving of publications. Elsevier

---

[7] Stijn Hoorens,Jeff Rothenberg, Constantijn van Oranje, Martijn van der Mandele,Ruth Levitt, Adresssing the uncertain future of preserving the past. Towards a robust strategy for digital archiving and preservation. Report RAND Europe 2007, forthcoming.

[8] Kurt de Belder, Saskia de Vries, Disseminatie van wetenschap onder invloed van nieuwe digitale ontwikkelingen en de rol van de universiteit, Internal report Leiden University 2007, p 2.

[9] Karen Hunter, " The End of Printed Journals: (in)frequently asked Questions" , Paper delivered at the University of Oklahoma Conference " Printed Resources and Digital Information": The Future of Coexistence", March 2-3, 2006, Oklahoma City. Karen Hunter, " Publishers and Librarians: Competitors or Collaborators?", Digital Libraries Symposium Seattle, January 20, 2007.

considers this a shared responsibility of the publishers, the library and third parties as Portico, the American initiative dedicated to long term preservation.

The world of libraries then. Just one look at the program of the IFLA (International Federation of Library Association) of this year at Durban, is enough to conclude that digital archiving and preservation as such is still a non-item in the international library-world in general. The only session organized about this specific subject "Digital Archiving and Preservation Update" was indicated as an informal session. Next to excellent and to-the-point presentations there were also speakers about 'digital archiving', who actually talked about their digitization projects, not about the archiving and preservation of the digitized objects. Apparently we still have some problems with the terms we use.

But let's not be too negative. Of course there is a lot going on and progress has been made, as is shown in the RLG DigiNews article "A Digital Decade: Where Have We Been and Where Are We Going in Digital Preservation?"[10] . Again: the focus of these publications is on technical, organizational and funding issues.

Other stakeholders are national governments and international organizations. In a growing number of western countries the deposit laws for national libraries are being actualized, now also including e-publications. In the European arena the European Commission is actively pushing forward developments, in the i2010 Strategy clearly establishing that digital preservation is a matter of concern to each country in Europe. And UNESCO published the Unesco Charter on the Preservation of the Digital Heritage[11]. Business enterprises also wrestle with their often very big data sets.When looking at what we call our commercial partners, like IBM, ExLibris/SUN and Tessella, one can detect a growing tendency toward more standardized systems. And last but not least we have active stakeholders like Google or Yahoo who are diving into the field of digitization of millions of older publications in library collections. With those actions they become our colleagues by providing access to academic digital resources. This kind of publications represent major sources for researchers in many fields, especially in the humanities. How Google and Yahoo plan to archive and preserve this massive body of knowledge is yet unclear.

The question arises how we bring about coordinated approaches, involving as many public and private stakeholders as we can. At the national level the KB took the initiative to set up the National Coalition for Digital Preservation[12] in the Netherlands, following other European examples like the British Digital Preservation Coalition[13]. Partners are amongst others the National Academy of Sciences, the National Archives, and the Central Institute for Statistics of the Netherlands. One of our first actions will be the exploration of the present situation in four fields: science, culture, government services and private enterprises. Who preserves what in the Netherlands and how are they doing that? A strategy will then be developed for a national infrastructure, and partners can be identified to take responsibility for certain activities.

In the European context the KB initiated, together with the British Library, the European Alliance for Permanent Access to the Records of Science[14]. Partners are amongst others the European Science Foundation, the European Space Agency, CERN (the European Organization for Nuclear Research), the Centre National d'Etudes Spatiales, the International Association of STM Publishers, the British Library, the Deutsche Bibliothek, the National

---

[10] NancyY.McGovern, "A Digital Decade: Where Have We Been and Where Are We Going in Digital Preservation?", April 15,2007,RGL DigiNews, Article 2; http://www.rlg.org; accessed 24/4/2007.

[11] See http://portal.unesco.org.

[12] http://www.ncdd.nl

[13] http://www.dpconline.org

[14] http://www.alliancepermanentaccess.eu

Archives of Sweden, the Digital Preservation Coalition etc. The Alliance wants to develop a shared vision and framework for a sustainable virtual infrastructure for permanent access to scientific information.

Recently the Alliance put forward a proposal to the European Commission called PARSE.Insight (Insight into issues of Permanent Access to the Records of Science in Europe). In short the strategic objectives are the support of policy development, including international cooperation for e-Infrastructures, in particular improved multidisciplinary ICT based infrastructures.

Intiatives as mentioned above are aiming at bringing order in our world of archiving and preservation of digital information resources. The strategy of the KB e-Depot is dedicated to this goal in the coming years. However, our actions do not cover the whole world, or all kinds of digital material and neither do they involve all stakeholders. And we do realize that the most important elements in all this are communication and the sharing of knowledge, while the most difficult elements in our worldwide communication are language and culture. So perhaps we should increase our efforts, next to what we are doing already. It would help if there would be a global overview of the different actions going on in different parts of the world, all using the same concept, as far as possible. That means we should all start in our own countries, then expand towards our continents. So let's shape our course together, mapping the world in a new way!

Author Biography

Els van Eijck van Heslinga started her career as a maritime historian at the University of Leiden in 1976. PhD in 1988. She joined the National Maritime Museum in Amsterdam as head of Collections and deputy-Director in 1990. In 1999 she went to the Koninklijke Bibliotheek/National Library at The Hague as deputy-Director. From 2004 till 2007 she was general director of the Central Office of Leiden University. As of April 2007 she is responsible for the program development of the International e-Depot at the Koninklijke Bibliotheek at The Hague.