

## A FUTURE FOR OUR **DIGITAL MEMORY**

*permanent access to information in the Netherlands*

interim report - summary in English



Edited by: Inge Angevaare

Sector research by Petra Helwig (Nationaal Archief, government & archives), René van Horik (Data Archiving and Networked Services, research community), Annelies van Nispen (Digitaal Erfgoed Nederland, cultural heritage), and Thaesis (private sector) (sector reports published in Dutch at <http://www.ncdd.nl/activiteiten-natverkenning.php>).

Dutch report available at <http://www.ncdd.nl/activiteiten-natverkenning.php>

English summary available at <http://www.ncdd.nl/en/activiteiten-natverkenning.php>

This interim report will be discussed during the national conference *A future for our digital memory*, to be held at the KB auditorium on 18 September; see (Dutch only) <http://www.ncdd.nl/werkconferentie2009.php>.

After the conference a final document including the strategy of the NCDD will be published.

Queries, comments or additions can be sent to the NCDD's coordinator, [inge.angevaare@kb.nl](mailto:inge.angevaare@kb.nl)

Dutch report 1 July 2009; English summary 1 September 2009



**NCDD**

Nationale  
Coalitie  
Digitale  
Duurzaamheid

NCDD – Netherlands Coalition for Digital Preservation

PO Box 90407

2509 LK The Hague, Netherlands

I <http://www.ncdd.nl/en/index.php>

E [info@ncdd.nl](mailto:info@ncdd.nl)

This report was funded by the Dutch Ministry of Education, Culture and Science.

## **A future for our digital memory: permanent access to information in the Netherlands *interim report***

*This is an expanded English-language version of the Executive summary and Conclusions of the Interim Report, highlighting some relevant particulars of the Dutch situation from the main body of the report (especially from chapter 4, 'The (Dutch) public sector and digital information in practice') while paying less attention to more general aspects of the permanent access challenge which are well documented elsewhere in the English language.<sup>1</sup> Summary prepared by Inge Angevaere, NCDD coordinator and editor of the Dutch report.*

### **1. The Dutch National Digital Preservation Survey**

Digital information has become a vital component of the knowledge economy in the 21st century. The public sector invests heavily in the creation and acquisition of digital information and has recognised the need to preserve essential information for use by future generations.

Digital information requires special curation which cannot be provided just anywhere; where such curation is unavailable, important records may be lost forever. This led to the establishment in 2007 of the Netherlands Coalition for Digital Preservation (NCDD), a cross-sectoral, bottom-up collaborative initiative in which eleven national organisations with responsibilities for safekeeping information joined forces to address the challenge of permanent access. NCDD membership includes:

- 3TU.Datacentre (a project of the universities of technology of Delft, Eindhoven, Twente)
- Data Archiving and Networked Services (DANS), the national digital archive for research data in the humanities and social sciences;
- Koninklijke Bibliotheek, National Library of the Netherlands (KB);
- Ministry of the Interior and Kingdom Relations (BZK);
- Nationaal Archief, National Archives of the Netherlands (NA);
- Netherlands Institute for Sound and Vision (NIBG);
- Netherlands Organisation for Scientific Research (NWO);
- Royal Netherlands Academy of Arts and Sciences (KNAW);
- SURFfoundation (innovation in higher education and research);
  
- Digital Heritage Netherlands (DEN), an advisory body for ICT in the cultural heritage sector, associate member;
- Statistics Netherlands (CBS), associate member.

The coalition's mission is to facilitate the establishment of both a technical and an organisational national infrastructure for permanent access to digital information in the Netherlands; the coalition aims at reaching results within five years.

In order to underpin its strategy, the NCDD decided to first build a detailed picture of the current situation in the public sector in the Netherlands. Can institutions or domains be identified which have successfully risen to the challenge of digital preservation and permanent access? Which categories of

data are in danger of being lost? How can the risks be managed? This so-called *National Digital Preservation Survey* was funded by the Ministry of Education, Culture and Science.

After some preliminary consultancy work it was decided that the survey would best be carried out by researchers with both knowledge of the issues involved in digital preservation *and* of the three sectors, which were identified as: scholarly communications, government & archives, and culture & heritage. A team of three researchers was recruited from among NCDD member staff, with the NCDD coordinator leading the project. The initial objective, to conduct a statistically relevant quantitative survey, had to be abandoned early in the project. The field to be surveyed was vast and varied, and some of the target groups were quite unfamiliar with the specifics of digital preservation, making on-line surveys unproductive. Therefore, the research team decided on a methodology of (some seventy) semi-structured interviews with knowledgeable stakeholders, adding relevant information from both Dutch and foreign published sources.<sup>2</sup> Five interviews were held with major private sector parties to establish whether the private sector has best practices to offer for the public sector to emulate.

The project ran from 1 January 2009 until 1 July 2009, when this interim report was presented to the Ministry of Education, Culture and Science. To follow up on the report, the NCDD is organising a national conference to discuss the results of the survey and debate on ways forward with stakeholders in the community. The conference is scheduled for 18 September in The Hague (discussions will be in Dutch). The survey and the conference will both feed into strategic discussions of the NCDD Board which will ultimately lead to an NCDD strategy for the coming years, which is expected at the end of 2009.

## **2. The three sectors: reasons for preserving information and division of responsibilities**

The sectors identified each have their own reasons for ensuring long-term access to information. *Government* agencies are accountable to society for their actions; in addition they wish to preserve the cultural heritage. The Public Records Act ('Archiefwet') provides the legal framework for long-term preservation of information, be it on paper or digital. Good *scholarly* practice dictates that research results must be verifiable. The digital age has, moreover, opened up new possibilities to reuse a single set of data in other research contexts. *Cultural heritage* institutions have an inherent remit for long-term preservation. The *private sector* is less interested in long-term preservation: information is kept no longer than is strictly necessary from a legal or litigation point of view.

When it comes to dividing responsibility for long-term preservation of digital information, the government sector operates within the clear framework of the Public Records Act. Cultural heritage institutions' responsibility goes without saying. However, securing long-term access can be hampered by the sector's fragmentation (see para. 5). No-one has as yet assumed responsibility for safekeeping the vast quantities of audiovisual content that emerge on the internet, on Web 2.0. Within the *research community* there is little clarity as to who is responsible for long-term preservation, or indeed, who is the owner of, e.g., digital research data. Liz Lyon's study *Dealing with Data* is referred to as guidance.<sup>3</sup>

### 3. Preserving digital information: requirements and guidelines

Having established why the public sector must to preserve information, and (to some extent) who is responsible, the next question is: what makes digital information special? Very briefly: digital information in and of itself is nothing but a row of ones and zeros, which derives its meaning from the very specific combination of hardware and software on which it was created. As both of these evolve continually, the time will soon come when a (new) computer can no longer play (old) information. This risk is being managed by placing the information under constant surveillance and adapt it to new hardware and new software (migration) or developing computers that can emulate their predecessors (emulation). Managing the risks requires an technically advanced environment in combination with a sustained research & development effort.

At the organisational level, measures must also be taken. Digital information requires cradle-to-grave monitoring. Producers, (temporary or long-term) data managers and users of digital information must all work in unison to make sure continuity is not disrupted. This organisational dimension was the main incentive for the NCDD's establishment.

Leading organisations in the Netherlands are well aware of instruments that have been developed worldwide to provide guidelines for digital data management and assess the quality of existing facilities, such as TRAC, DRAMBORA, DAF and the nestor *Kriterienkatalog*. In the present study, the researchers based their assessment on three tools which were developed in the Netherlands, two of which in the Dutch language:

- Government sector: The Ministry of the Interior and the Ministry of Education, Culture and Science jointly developed the programme 'Information in Order' ('Informatie op orde'), which includes a *Baseline information management* which defines rules for managing information all through the lifecycle. It stresses the importance of acting with a long-term view from creation onwards. The Baseline was inspired by ISO 15489 and applies to all national government institutions.
- Scholarly communications: Data Archiving and Networked Services (DANS) developed a *Data Seal of Approval* for research data, a comparatively accessible set of guidelines for producers, data managers and users to ensure long-term access. The DSA is a voluntary instrument which is still under development. Responsibility for the DSA has recently been transferred to an international editorial board.
- Culture & heritage: Digital Heritage Netherlands (DEN), a national body set up to promote the use of ICT in the cultural heritage sector, added a digital preservation chapter to its set of IT guidelines 'The Basis'. The guidelines explicitly refer to TRAC for a full set of requirements.

All three sets of guidelines emphasize that permanent access can only be achieved by organisations which have an *express long-term mission* and have designed their organisation accordingly.

## 4. The Dutch public sector and digital preservation in practice

Having established some requirements of long-term access and the guidelines which are applied in the Netherlands, it is now time to look at the real world: how important is digital information to the sectors and how does day-to-day practice compare to requirements and guidelines?

### 4.1. *The importance of digital information to the sector*

The research community was the first sector to discover the potential of digital communications: take-up was quick and widespread. Science, technology and medicine led the way and soon established vast, world-wide data networks. The social sciences followed suit, and, more hesitantly, the humanities. But for a few exceptions modern research is quite unthinkable without digital communications.

The other two sectors, government & archives and culture & heritage, are still in the middle of the transition from analogue to digital. Although computers are widely used by government agencies, the ensuing official *record* more often than not is still a print-out, validated by a 'wet' signature. The need to secure authenticity is cited as the major stumbling block for going fully digital. Within cultural heritage institutions the physical object still takes pride of place – digital reproductions serving predominantly as on-line advertisements for the physical collection. All respondents, however, estimate that the impact of digital information will increase substantially in the near future.

### 4.2. *Should we preserve everything?*

The respondents in the Dutch study agree wholeheartedly that permanently preserving every digital object that is being created or acquired is too expensive and undesirable. The present generation must make informed choices as to what will matter to future generations and what will not. The government sector has inherited the most elaborate selection mechanism from the analogue era (mandated by the Public Records Act), but it needs to be adapted to the digital age. Scholarly communications show a varied picture: research *publications* largely follow the lines developed in the analogue age, with publishers and peer review acting as selection agents. When it comes to *research data*, however, no selection criteria have as yet become available.

### 4.3. *Facilities for long-term preservation of digital objects*

In the Netherlands, nine organisations have so far emerged which make it their core business to ensure permanent access to digital objects. Four organisations have a clear national mission:

- The Koninklijke Bibliotheek (National Library of the Netherlands) in 2003 established its 'e-Depot' to preserve the national digital deposit collection. The KB soon realised that *national* deposit collections hardly fit the age of internet, as academic publishers have offices all over the world and publish globally. The KB therefore opened up its e-Depot to scholarly publishers worldwide, provided they conclude archiving agreements with the KB.
- The Netherlands Institute for Sound and Vision (NIBG) developed an audiovisual archive together with Technicolor (Thomson) where the output of the Dutch public broadcasting companies is automatically harvested.
- Data Archiving and Networked Services (DANS) was established in 2005 by the Royal Netherlands Academy of Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO) to preserve research data in the social sciences and the humanities. DANS included previously established collections, especially in the social sciences (Figure 1).
- The Nationaal Archief (NA) will open up a digital archive in 2009.

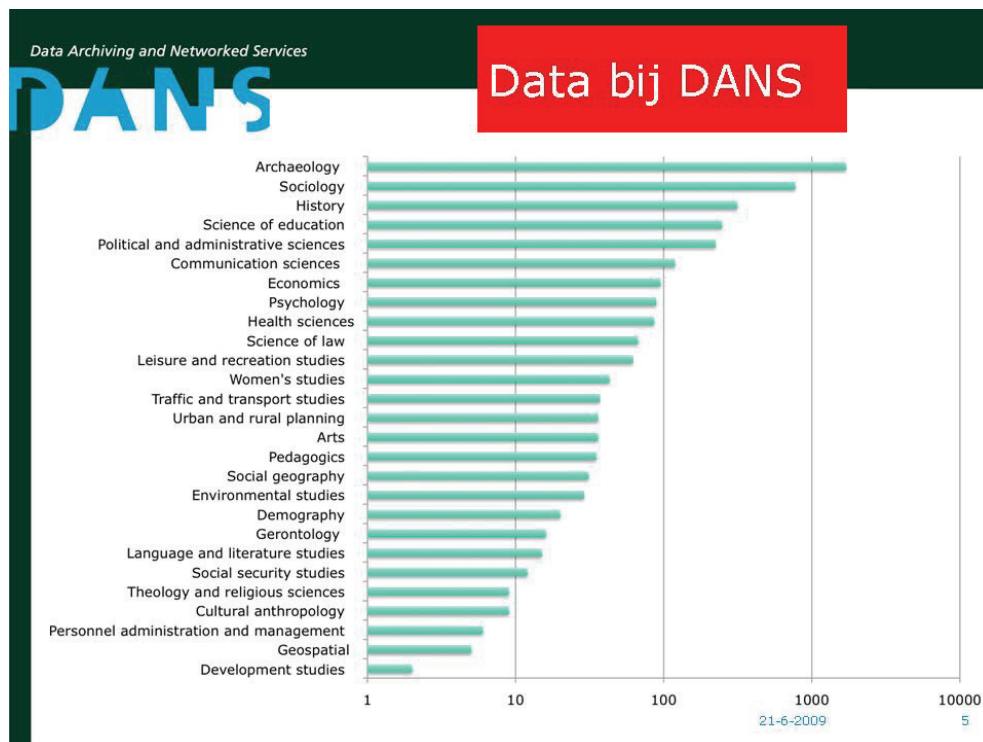


Figure 1. DANS digital data collections, June 2009.

KB, NBIG, DANS and NA meet the key requirement in the Dutch guidelines that their policy and organisation are specifically attuned to long-term preservation. The same can be said for a number of other organisations:

- The Judicial Information Service (JustID) developed a digital archive (CDD+) in which files from across the criminal justice system are brought together and made available to all parties concerned.
- The Rotterdam City Archives (Rotterdams Gemeentearchief) have adopted the technical solutions developed by the Nationaal Archief and will open up their digital archive in 2009 as well.
- The Dutch National Meteorological Institute (KNMI) has developed its KODAC system for meteorological data.
- The Max-Planck-Institute for Psycholinguistics preserves audio tapes of (nearly) extinct languages in its DoBeS archive.
- The Institute for Dutch Lexicology in 2004 established the Dutch HLT Agency of the Dutch Language Union ('Nederlandse Taalunie').

These nine organisations can be said to have reached the stage where *bitstream preservation* has been achieved; specific *preservation strategies* such as migration or emulation have not yet been implemented, in part because the archives contain relatively new material. Respondents indicated, however, that the moment when decisions have to be made with regard to migration or emulation is drawing near.

Other facilities for long-term archiving are in various stages of development:

- The International Institute for Social History (IISG) is building a system to archive mostly confidential social history data; the project runs until 2010.

- The Amsterdam City Archives are said to work in small steps towards long-term archiving facilities; the pace is determined by funding.
- The 3TU. Federation of the three universities of technology is involved in a two-year project to develop a data centre for technology.
- The Group of Archivists of the Province of Limburg (KVAL) has in 2009 published a strategy paper in which the establishment of a Limburg regional digital archive is foreseen.
- The Netherlands Architecture Institute (NAi) is moving towards a facility for long-term access to architects' collections.

Apart from these facilities which all share a specifically long-term goal, the study uncovered a wide range of (national and international) digital storage facilities with more or less short-term objectives, especially in the scholarly world. None of these facilities fulfill the requirement of a long-term strategy. Mention must be made in this respect of the institutional repositories of the Dutch universities. These have realised permanent access for the *publications* in their collections through a joint agreement with the Koninklijke Bibliotheek, which harvests the institutional repositories on a regular basis. As far as *research data* are concerned, however, there is sometimes an intent to secure permanent access (Wageningen, Utrecht), and back-up procedures have mostly been implemented, but implementation of long-term strategies is still a long way off.

The study concludes that content which comes within the purview of the nine organisations with explicit missions for long-term preservation, can be deemed to be relatively safe – provided these organisations show the necessary continuity both in terms of their organisation and their funding. However, when set against the quantities of digital data that are being created, the efforts described seem quite fragmentary, as will be shown in the next paragraph.

#### **4.4. A look at the creation phase**

To what extent are measures being taken to ensure long-term access at the vital creation phase which precedes any long-term archiving effort?

##### 4.4.1. The government

From 1988 on, the government's management of digital records has been the object of scrutiny by a number of oversight agencies, and their verdict was consistently devastating: the government is careless in its management of records and few agencies are aware of what goes on within their offices. The term 'dementia' was even used, both with regard to national government agencies and to local and regional authorities. Interviews in this study confirm that little has changed since 1988. Both nationally and locally, government agencies view data management as a very secondary aspect of their jobs thought to serve only cultural historians.

The national government has recently launched its 'Information in Order' campaign to improve the situation. Whether or not this programme has the desired effect is not yet clear; the Dutch Court of Audit ('Algemene Rekenkamer') is working on a first evaluation which is due to be published at the end of 2009.

##### 4.4.2. Cultural heritage

Interviews in the study corroborate the findings of Digital Heritage Netherlands (DEN) in various recent studies: the cultural heritage sector's main concern is still with physical collections. Digital records are mostly derivatives of physical collections: metadata or digital copies for online advertising. Where real (substitution) digitisation occurs, it takes up the available resources to such extent that planning beyond the phase of digitalisation is, for the moment, a bridge too far. Generally, awareness of digital preservation issues is low.

Some smaller and medium-sized cultural heritage organisations which *are* aware of the risks their digital collections run, look to the national institutions for help in securing archiving: the Koninklijke Bibliotheek, the Nationaal Archief, Sound and Vision.

Web archiving must also be mentioned here. The Koninklijke Bibliotheek has started harvesting a selection of Dutch websites and expects to archive about 45,000 individual sites by 2012. However, this cannot be but the tip of the iceberg. Web 2.0: Hyves, YouTube, blogs, Picasa, Twitter and all those other social sites where cultural heritage is emerging, is as yet uncharted territory when it comes to permanent access.

#### 4.4.3. Scholarly communications

When it comes to scholarly communications, a sharp dividing line can be distinguished between publications on the one hand and research data on the other. Scholarly publications largely continue to follow the routes inherited from the analogue era: via publishers to (national) libraries for permanent storage, with the difference that national deposit collections have been replaced by international archives such as the Koninklijke Bibliotheek and Portico.

Raw and processed research data, on the other hand, constitute a new category peculiar to the digital era, and both the magnitude of data flows and the variety of types of data boggle the mind. This makes it virtually impossible to speak in general terms of 'the' creation phase of digital objects, but some general trends have been observed<sup>4</sup> which apply to the Netherlands as well. In short: disciplines which can hardly function anymore without longitudinal data have readily joined forces to develop data networks and storage facilities, often on an international scale (social sciences, natural sciences); in disciplines where reuse of another researcher's data is a relatively new phenomenon, take-up is less pronounced and often hampered by a lack of motivation on the part of data producers.

#### **4.5. Impediments to long-term access to digital objects – and some inspiration for action**

The following is a list of impediments to long-term access which were identified by respondents in the survey. Each item in the list is followed by initiatives which have already been taken in the Netherlands to address some of the problems, and which may inspire action elsewhere:

1. *Lack of awareness.* Respondents report that despite efforts to turn the tide, the risks run by digital information are underestimated, both at the executive and at the work floor level.
  - \* Developing and publishing guidelines such as those mentioned in para. 3 (Data Seal of Approval, The Basis, Baseline) substantially contributes to raising awareness.
  - \* DANS, the KB, Statistics Netherlands and other academic institutions collaborate in the publication of the (printed) magazine 'e-data & research', which promotes sharing and reuse of data and digital collections among researchers in the humanities and social sciences.
2. *Shortage of knowledge and expertise on digital preservation.* Knowledge and expertise are concentrated within the four major national institutions (KB, NA, DANS and NBIG), which have set up websites to disseminate knowledge, but the information is often outdated as maintenance does not receive priority. Conferences and workshops are sometimes organised, but not in any systematic way.
  - \* In July 2009 SURFfoundation established a 'research data forum' to exchange knowledge and expertise among digital archives and repository managers.

- \* The Rotterdam City Archives are noted for their efforts to keep colleagues abreast of the development of their digital archive and for facilitating discussions in the government & archives sector about issues specific to the digital age.
  - \* The NCDD itself brings together on its website information and knowledge from all organisations concerned, but as yet on a modest scale.
3. *Inadequate information management throughout the lifecycle.* When information is mismanaged at the creation phase, ensuring permanent access becomes a virtual impossibility. The government sector is a case in point. Data sets of scientific projects are also often left in disarray when the researcher in question moves on to another project.
- \* The Government has developed the programme 'Information in Order' to further adequate data management by national government agencies. The Association of Dutch Municipalities (VNG) has indicated it wants to join the programme.
  - \* Chief Information Officers have recently been appointed to the Ministry of the Interior and to the Internal Revenue Service.
  - \* The 3TU.Datacentre is involved in a project entitled 'Value-added data and services', which aims to support researchers in managing their digital assets.
4. *Shortage of structural funding.* Digital preservation is a long-term effort which requires long-term funding. As elsewhere in the world,<sup>5</sup> both research and digitisation are often project funded, with little or no provision made for safekeeping the output once the project is finished. A notable example in the Netherlands is the €154 million audiovisual digitisation project *Images for the future*, which is expected to yield large quantities of digital files to substitute decaying originals; beyond 2014, no resources have as yet been made available for permanent access.
5. *Private, short-term concerns clash with public, long-term concerns.* In all sectors the concerns of the data producers lie mainly in the private domain and the short term – the long-term concern is not acutely felt, and any extra work that needs to be done to ensure long-term access is experienced as burdensome. Researchers' main goal is to publish their papers in reputed journals; managing their data for reuse by others is a very secondary concern.
- \* The 3TU.Datacentre's project is a best practice here too. The aim is to improve the quality of the research itself by providing data management services – while at the same time, almost imperceptibly, take the measures needed to facilitate long-term access.
  - \* The Netherlands Organisation for Scientific Research (NWO) has started making certain types of research funding ('NWO middelgroot') dependent upon a commitment to deposit research data with DANS.
  - \* DANS itself developed its online archiving system EASY to alleviate the burden on researchers when depositing their data.
  - \* The *Standard Evaluation Protocol 2003-2009* which was agreed by the Royal Netherlands Academy of Sciences (KNAW), the Netherlands Organisation for Scientific Research (NWO) and the Association of Universities in the Netherlands (VSNU) mentions

the possibility of including other types of scholarly output than publications in the process of rating the quality of an institution.

6. *Shortage of trustworthy storage capacity.* Smaller organisations which have insufficient means to develop their own repository, express interest in some type of shared long-term storage facility for their collections; one respondent spoke of a 'data hotel'.

- \* In June 2009 the National Archives and the eleven Regional Historical Centres (combinations of provincial archives and museums) formed a consortium to organise digital archive capacity for general use within the archival sector within the next five years.

- \* Municipal archivists in the province of Limburg recently decided to investigate the possibility of setting up a regional digital archive.

- \* Several respondents in the study suggested that regional intersectoral archives should be looked into rather than taking it for granted that national sectoral archives would come to be.

- \* The Judicial Information Service developed a joint storage facility while respecting each contributor's access regimes.

- \* Internationally, various projects have been set up to investigate the possibilities of storing a variety of content in one system (iRODS, TIPR, etc.).

7. *Insufficient manpower and/or IT capacity.* Even when the national infrastructure provides for a dedicated long-term archive for certain categories of content, the producing organisation may not have the resources to prepare its digital collections for permanent archiving. Statistics Netherlands reported such problems with regard to their digital publications and with regard to the data behind StatLine.

8. *Shortage of practical tools.* Organisations with digital collections ask for practical tools to manage their collections. One respondent spoke of a 'digital toolbox', and another emphasized that such tools must be inexpensive and easy to use.

- \* Research into such tools has been organised on an international scale: PLANETS (libraries and archives), CASPAR Preserves (research data) and PrestoPRIME (audiovisual content).

9. *Shortage of dedicated long-term preservation services.* A number of organisations with digital collections would prefer not to develop all the necessary expertise in-house, but be able to make use of services offered nationally or internationally. Such organisations have in the past applied to the four national archiving institutions (KB, NA, DANS, NIBG), but these often cannot meet the demand.

- \* The National Archives are in 2009 running pilot projects with various types of services: a) early transfer of records; b) acting as a trusted repository both for records to be archived *and* records to be disposed of (both projects with the Ministry of the Interior); c) acting as a trusted repository for Regional Historical Centres and other archives (pilot project with the Utrecht City Archives); and d) providing a service oriented architecture (pilot project with the Land Registry Office or 'Kadaster').

\* The Netherlands Institute for Sound and Vision (NIBG) in 2008 launched 'ProArchive', a service for metadating, archiving and providing access to audiovisual collections. So far, however, few contracts have been signed. Perhaps it is the price tag of digital preservation that scares cultural heritage organisations off. It may also be that NIBG needs to come up with a more differentiated supply of services (e.g., archiving without producing metadata).

10. *The sector is not organised according to the demands of the digital age.* Roles and responsibilities inherited from the age of analogue information sometimes hamper a smooth processing of data in the digital age, especially cradle-to-grave curation. This applies notably to research data, a relatively new phenomenon of unprecedented large-scale international proportions. So-called *enriched* publications may also present problems, as the publications are archived by the KB and the underlying research data by, e.g., DANS.

\* Inspiration for an international problem must be found on a European scale, e.g., the ESFRI European Roadmap for Research Infrastructures, the Alliance for Permanent Access, DARIAH, CESSDA, CLARIN and BBMRI.

\* A roadmap for permanent access to the records of science is being developed by the PARSE.insight project which originated within the Alliance for Permanent Access (APA).

\* In the Dutch context the SURFshare programme must be mentioned, in which DANS, the KB, the 3TU.Datacentre and other organisations cooperate to develop an infrastructure for enriched publications.

\* The government & archives sector as yet seems to be of two minds: on the one hand the *records continuum* principle is making headway in theory, while on the other the Public Records Act still draws a sharp dividing line between the legal regimes before and after transfer. Such a sharp division is really undesirable from a perspective of continuing curation. (Associations of) archivists are pleading for gradual changes to the Public Records Act. It is observed, however, that in and of itself such changes are not strictly necessary to support records management in the active use phase: the Rotterdam City Archives are presently developing one storage system for both active use and archival records while respecting the different legal regimes.

\* At the moment, administrative responsibility for government records is split between the Ministry of the Interior (active use phase) and the Ministry of Education, Culture and Science (archival phase). In their latest strategy document 'Archiving is a matter of looking ahead' BRAIN and KVAN, the leading national associations of archivists, are proposing that responsibility for the entire information lifecycle be placed under one single Ministry.

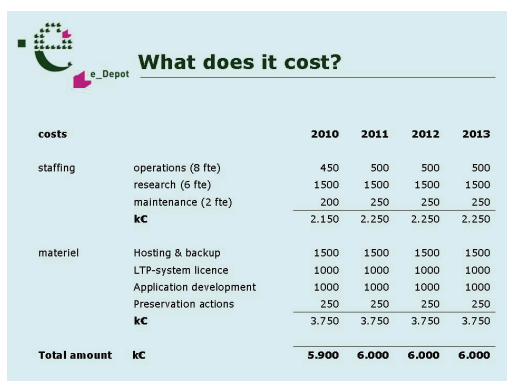
\* The CDD+ system developed by the Judicial Information Service provides a good example of bringing information from different stakeholders together in one system while respecting each stakeholder's legal rights and responsibilities.

## 5. Costs and funding

The question ‘What does it all cost?’ is asked in the Netherlands as often as it is elsewhere, but the Dutch study turned up no additional data to support costing long-term access to digital collections beyond what has already been reported in the English-language studies summarised in *Sustaining the digital investment* (especially Chapter 3, which includes data from work done by Beagrie et al.)<sup>6</sup> In short:

- So many variables are involved in costing permanent access, that it is quite impossible to draw comparisons between different organisations and studies.
- Scale matters: the costs per unit decrease as quantities increase; scale advantages go down if the number of suppliers or the variety and complexity of objects go up.
- As time goes on, the costs per unit per annum seem to taper off.
- Staff costs are a very substantial part of the total cost of ownership, much more so than physical storage media.
- Timing is of the essence: when proper measures are not taken at the creation stage, repairing the damage later in the lifecycle is prohibitively expensive.
- Acquisition and ingest seem to be the most expensive phases of the lifecycle.
- Although emulation is generally considered to be the most expensive preservation strategy, research by the KB has suggested that emulation may, in fact, be the cheaper solution in the long term. Emulation, however, does require substantial initial investments, while migration costs may be spread over time.<sup>7</sup>

As for more detailed information: so far only the Koninklijke Bibliotheek (KB) has published facts and figures about its e-Depot, which had been operational since 2003. In 2007 the KB estimated that operational costs and research & development amounted to some €3.655 million for about 10 million digital objects, mostly PDF's.<sup>8</sup> In 2009, the KB embarked on a programme to drastically expand the e-Depot's capacity in order to be able to archive the master files from the KB's digitisation projects and the Dutch Web Archive. All-in annual operating costs are now expected to go up to €6, as shown in Figure 2.



costs		2010	2011	2012	2013
staffing	operations (9 fte)	450	500	500	500
	research (6 fte)	1500	1500	1500	1500
	maintenance (2 fte)	200	250	250	250
	<b>k€</b>	2.150	2.250	2.250	2.250
materiel	Hosting & backup	1500	1500	1500	1500
	LTP-system licence	1000	1000	1000	1000
	Application development	1000	1000	1000	1000
	Preservation actions	250	250	250	250
	<b>k€</b>	3.750	3.750	3.750	3.750
<b>Total amount</b>	<b>k€</b>	<b>5.900</b>	<b>6.000</b>	<b>6.000</b>	<b>6.000</b>

Figure 2. Cost estimates for the KB e-Depot, 2010-2013.

Source: Ras, 2009.<sup>9</sup> Note: Material costs do go up, but the KB concludes multi-year contracts with suppliers in order to distribute costs more evenly over the years.

When it comes to funding, permanent access to public information is a government affair in the Netherlands – be it at the national, the regional or the local level. Recommendations from the US/UK Blue Ribbon Task Force to try and attract more private funding<sup>10</sup> hardly seem realistic in the Dutch context. The KB's e-Depot is explicitly funded by the Directorate of Higher, Vocational Education and Science of the Ministry of Education, Culture and Science, as is the DANS archive (through KNAW and

NWO). The National Archives are funded by the Directorate General of Culture and Media within the same Ministry.

Within the cultural sector, government policy has been aimed at decentralising responsibilities to lower levels of government, where funding at the scale needed for long-term archiving is not available. The sector as a whole is fragmented and resources are scarce. Where resources are available, these are directed mostly at safekeeping physical objects and at providing (short-term) online access to digital collections.

Project funding is a structural problem, as was mentioned in section 4.5 (bullet no. 4).

At the end of the chapter on costs and funding, the study touches upon the question whether the benefits can outweigh the costs – and fails to answer the question unequivocally. Perhaps the question should be rephrased: what requirements does society set for its digital memory? Or: how much is society prepared to pay for safeguarding its digital memory for future generations?

## 6. Prospects for (cross-sectoral) cooperation in the Netherlands

After briefly reviewing collaborative efforts in the US (NDIIPP), the UK (DPC) and Germany (nestor), the report states that respondents in the study are unanimous in their assessment that cooperation is conducive to securing a future for our digital memory – but opinions vary as what form such cooperation should take: from the most informal exchanges of information to building one national storage facility.

There can be no doubt that the digital age requires closer cooperation between partners *within* information chains: between producers, archives, and users.

In archaeology, a common interest and the Treaty of Valletta have led to successful cooperation between a variety of stakeholders: the Government Agency for Cultural Heritage (RCE), DANS, regional archaeological repositories and private companies have joined forces to build EDNA, an integrated information infrastructure for research data and publications.

The Judicial Information Service's CDD+ system is another example in which a shared problem (slow paper trails) can lead to effective cooperation.

When asked about the prospects for cross-sectoral cooperation, quite a number of respondents indicate that they are at present bothered by fragmentation in the field: 'DEN, DANS, NCDD ... they all do something, but what is it exactly that they do?' They expect that cooperation in the field of knowledge and expertise will yield short-term benefits for all.

Some respondents emphasize that their sector is so *different* from the other sectors, that cooperation will only yield occasional benefits. Others feel that the sectors are growing closer together. The question of authenticity, for example, used to be an archivist's problem only. Now libraries, too, have to make sure their migrated content is still representative of the original file. The records continuum theory developed by the archival community can inspire data management both in the research community and cultural heritage institutions. The research community, on the other hand, has made the best advances in building distributed networks and promoting interoperability by means of persistent identifiers which prevent internet content from being lost.

Research by OCLC, among others, has shown that users prefer their information resources to be as integrated as possible. They hardly care where the information originates: a library, an archive, a museum. Thus the ambitions for the *front office* are clear. The digital age presents new possibilities for scale advantages in the *back office*, without the user being adversely affected in any way.

## 7. Conclusions and recommendations

*The conclusions and recommendations in this interim report are of a preliminary nature. They will be discussed at the National Conference of 18 September 2009 and by the Executive Board of the NCDD. At the end of 2009 the NCDD will publish its policy choices.*

Digital information has become a vital component of the knowledge economy in the 21st century. The public sector invests heavily in the creation and acquisition of digital information and has recognised the need to preserve essential information for use by future generations.

Safeguarding permanent access to information in the digital age calls for specific measures to be taken:

- Digital information requires cradle-to-grave monitoring; at creation measures must be taken to enable permanent access. Careful monitoring and sustained research & development are needed throughout the lifecycle.
- It takes a sophisticated technical environment to ensure permanent access with concomitant costs and special expertise.

The National Digital Preservation Survey shows that long-term curation requirements are insufficiently met in the Netherlands:

- The number of trustworthy repositories is limited and their capacity insufficient.
- Substantial categories of digital information which are created or acquired do not come within the purview of any long-term repository. The reasons for this include lack of awareness and a fragmented division of roles and responsibilities from the analogue age which hampers cradle-to-grave curation.

If this situation persists, it will lead to loss of information, capital destruction and a weakening of the knowledge economy.

### A national infrastructure for permanent access

The members of the Netherlands Coalition for Digital Preservation wish to join forces to facilitate the establishment of a national infrastructure of repositories, services and tools, in which:

1. awareness of the long-term access issues is present throughout the lifecycle of digital information;
2. roles and responsibilities have been clearly defined to improve a smooth transfer from data creators to digital archives;
3. enough trustworthy repository capacity is available for storage and long-term management;
4. the parties involved share services and tools as efficiently as possible;
5. *structural* funding is available for those activities which secure permanent access;

6. dedicated selection mechanisms are available, attuned to the digital age;
7. stakeholders share knowledge and expertise – nationally and internationally.

Below are some more detailed recommendations to achieve the desired results. Where possible, parties who should take the initiative are identified.

### 1. Awareness

Respondents in the survey are unanimous in their assessment that awareness of the risks of digital data has not yet penetrated as deeply as it should.

- Make sure that issues of permanent access remain on the (political) agenda: locally, regionally, nationally, and within the various scholarly disciplines. Keep investing time and resources in information, communication and lobbying activities.  
→ NCDD members, NCDD office

### 2. Roles and responsibilities

Many producers of digital information do not recognise the long-term concern as their own: the immediate service to the public takes precedence or the publication in a reputable journal of a scholar's own research. Such conflicting interests can never be banned altogether, but the gap can be reduced:

- Give researchers a stake in long-term access to their research data, by making them formally responsible or rewarding them, financially or otherwise, for their efforts. *Funders* of scholarly research have a role to play here. They can make research grants contingent upon data deposition. In an international context, NCDD members should contribute to devising a system by which academic prestige can not only arise from publications but also from quality data sets.  
→ research funding organisations, universities, organisations for scientific research, umbrella organisations
- Make explicit arrangements with researchers as to who owns the data and who is responsible for various aspects of long-term curation (metadata, storage, etc.)  
→ universities, organisations for scientific research, umbrella organisations
- Make it compulsory that each research or digitisation project plan at least includes a provision on the need to provide long-term access to (selections of) the digital information, on who is responsible and/or how such access may be funded.  
→ NCDD member organisations, Ministry of Culture, Education and Science, the Netherlands Organisation for Scientific Research, the Royal Netherlands Academy of Sciences, universities
- Work on dedicated selection mechanisms for research data  
→ universities, research organisations, umbrella organisations, researchers
- Organisations which have permanent access in their remit should take the initiative to bridge the gap to producers of digital objects by providing data management support in the creation phase and embed long-term curation in short-term work.  
→ digital archives, libraries, data centres

- The new version of the *Standard Evaluation Protocol* should explicitly refer to the need to provide long-term access to the records of science.  
→ the Association of Dutch Universities, the Netherlands Organisation for Scientific Research, the Royal Netherlands Academy of Sciences
- Bring responsibility for the entire lifecycle of public records under one single ministry.  
→ the Cabinet
- Expand the national government programme 'Information in Order' to regional and local government agencies.  
→ local government agencies, Ministry of the Interior, Ministry of Education, Culture and Science
- Organise a conference to facilitate discussions on the responsibility for permanent access to Web 2.0 content  
→ NCDD office, NCDD member organisations, Ministry of Education, Culture and Science

### **3. Trustworthy repository capacity**

Long-term digital archives are expensive to develop and maintain. Sharing large-scale facilities might seem the most cost-efficient solution, but producer or user requirements may not dictate otherwise. Therefore, additional research is needed to determine how and where scale advantages can be achieved. Some respondents estimate that the best chances for sharing lie at the bottom of the information architecture, at the storage level. If true, this may also lead to marketing out storage to commercial suppliers.

Opinions differ as to the way in which central facilities should be organised: geographically (local, regional, provincial, national), by discipline, by (cultural) sector, and/or specialised according to file format.

- Make a more detailed inventory of storage capacity needs to inform policy.  
→ NCDD members through the respective umbrella organisations; coordination by the NCDD office
- Strive for scale advantages whenever possible to lower the costs  
→ all
- Align various current initiatives such as the plans developed by the consortium of the National Archives and the Regional Historical Centres, and the Limburg regional e-depot. Determine if such an infrastructure can also be used to accommodate digital collections of smaller cultural heritage institutions and libraries and/or if regional intersectoral facilities must also be built.  
→ all institutions involved; NCDD to coordinate the debate
- Find international support for developing facilities for internationally oriented academic disciplines, e.g., the European Alliance for Permanent Access, DARIAH, CESSDA, and CLARIN  
→ Royal Netherlands Academy of Sciences, the Organisation for Netherlands Research, the Association of Dutch Universities, universities, research institutions

#### 4. Services and tools

Medium and small-sized organisations especially ask for practical, easy to use services and tools to manage their digital collections. The guidelines for permanent access which have been developed so far (see paragraph 3) are quite high-level; they are built on the premise that organisations must be specifically designed to facilitate permanent access. Such guidelines do not meet the needs of institutions for which managing digital objects is only a secondary task and/or who cannot afford to contract third-party curation services.

- Develop a national, cross-sectoral supply of practical, easy-to-use services and tools which support smaller organisations in managing their digital objects (a 'digital toolbox'); consider establishing a separate, dedicated organisation to offer these services rather than asking one of the national institutions to do so.  
→ NCDD member organisations, coordinated by NCDD office
- Design a practical step-by-step plan which can help smaller organisations minimise the risks their digital collections run while they are still in their care.  
→ NCDD member organisations, coordinated by NCDD office
- Do not wait for the national organisations to develop initiatives to solve your problems, but take all the practical measures which you can take at your own level to reduce the risks (make an inventory of your collections, make selections for permanent access, choose your file formats wisely, make a back-up copy, produce metadata, etc.).  
→ smaller organisations, cultural heritage institutions, producers of digital objects
- Take a pragmatic attitude towards developing technical facilities, services and tools – a time horizon of five years is quite long in such a swiftly evolving field.  
→ all organisations with digital archives

#### 5. Structural funding

It is essential that funding organisations provide *structural* funding for developing an infrastructure for digital information, in order to prevent breaks in the required continuity.

- Make sure that organisations which have made substantial progress towards permanent access, such as the Koninklijke Bibliotheek, the Netherlands Organisation for Sound and Vision, Data Archiving and Networked Services, the Nationaal Archief, the Rotterdam City Archives, the International Institute for Social History, and the 3TU.Datacentre receive (additional) structural funding to keep their facilities up to date.  
→ Ministry of Education, Culture and Science, Ministry of the Interior, universities
- Consider additional up-front funding from endowments to which research and digitisation projects make contributions.  
→ NCDD members, Ministry of Education, Culture and Science, the Netherlands Organisation for Scientific Research, Royal Netherlands Academy of Sciences, universities

## 6. Selection

In order to contain maintenance costs, selection must be even more stringent than in the analogue age.

- Work actively on translating selection mechanisms from the analogue age into the digital age and to develop new selection mechanisms for content which has originated in the digital age (web 2.0, research data).  
→ NCDD member organisations, all sectors
- Strive for selection at creation and for automatic selection for digitally born content.  
→ cultural heritage institutions, libraries, digital archives
- Make a careful distinction between digital objects that really should remain available to future generations and objects which merely serve as on-line PR copies for physical collections. In the latter case, repeating digitisation if need be can be cheaper than permanent archiving.  
→ all sectors

## 7. Knowledge and expertise

Outside the national archiving organisations, respondents indicate that knowledge and expertise in the Netherlands are, at the moment, scattered and inaccessible. Websites are set up with enthusiasm, but are insufficiently maintained or are too specific; meetings are organised occasionally, but there is no sustained effort.

- Prioritise exchanges of knowledge and expertise in order to further innovation and development  
→ all institutions
- Put an end to fragmentation of knowledge and expertise on permanent access in the Netherlands. Set up a national knowledge centre where dissemination of knowledge heads the list of priorities  
→ NCDD member organisations, coordinated by NCDD office; cooperate with other knowledge centres such as ICTU (national government) and DEN (ICT in cultural institutions)
- Establish a national technology watch.  
→ all archiving institutions, coordinated by the NCDD.

## The role of the NCDD

It must be emphasized here that existing organisations are, in the first instance, responsible for actions within their own sector, especially when their remit includes national responsibilities within their sector (KB, National Archives, DANS, Sound and Vision). The NCDD is a forum where parties from various sectors sit at the table and align their initiatives. In addition, the NCDD proposes to play a role in establishing knowledge dissemination on a national scale and organise a lobby for the necessary resources.

Other parties than NCDD members must be involved in the infrastructure debate if it is to succeed: umbrella organisations of cultural heritage institutions and archiving institutions which are not (yet) members of the NCDD, such as the Judicial Information Service, the Royal Dutch Meteorological Institute, the Max-Planck-Institute for Psycholinguistics, and others.

---

<sup>1</sup> Original version: 'Toekomst voor ons digitaal geheugen: duurzame toegang tot informatie in Nederland', Nationale Coalitie Digitale Duurzaamheid, juli 2009, <http://www.ncdd.nl/activiteiten-natverkenning.php>, retrieved 30 August 2009.

<sup>2</sup> A full list of the literature cited (both in Dutch and in foreign languages) is provided by Appendix 1 of the Dutch-language version at [http://www.ncdd.nl/documents/NCDDToekomst2009\\_000.pdf](http://www.ncdd.nl/documents/NCDDToekomst2009_000.pdf), pp. 69-79, retrieved 30 August 2009.

<sup>3</sup> Dealing with data: roles, rights, responsibilities and relationships [2007], Consultancy report by Liz Lyon, UKOLN, <[http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dealing\\_with\\_data\\_report-final.pdf](http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dealing_with_data_report-final.pdf)>, retrieved 4 June 2009.

<sup>4</sup> Among others *The UK Research Data Service Feasibility Study, Final Report* (2008b), 8 December 2008, <<http://www.ukrds.ac.uk/HEFCE%20UKRDS%20Final%20Report%20V%201.1.doc>>; *To Share or Not to Share: Publication and Quality Assurance of Research Data Outputs* (2008), Report commissioned by the Research Information Network (RIN), <<http://www.rin.ac.uk/data-publication>>; *Sustaining the Digital Investment: Issues and Challenges of Economically Sustainable Digital Preservation* (2008), Interim Report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access, <[http://brtf.sdsc.edu/biblio/BRTF\\_Interim\\_Report.pdf](http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf)>; all retrieved January 2009.

<sup>5</sup> See, a.o., the report *Sustaining the digital investment*, op. cit.

<sup>6</sup> *Sustaining the digital investment*, op. cit. n. 3, and Beagrie, Neil; Julia Chruszcz and Brian Lavoie (2008): *Keeping research data safe: a cost model and guidance for UK universities*, <<http://www.jisc.ac.uk/media/documents/publications/keepingresearchdatasafe0408.pdf>>, retrieved 12 January 2009.

<sup>7</sup> Oltmans, Erik en Nanda Kol (2005), 'Emulation and migration: a comparison in terms of costs', DCC/DPC Seminar on Costs Models, British Library, July 26, 2005, <<http://www.dpconline.org/docs/events/050726oltmans.pdf>>, retrieved 21 May 2009.

<sup>8</sup> *Addressing the uncertain future of preserving the past: towards a robust strategy for digital archiving and preservation* (2007), RAND technical report in opdracht van de Koninklijke Bibliotheek, door Stijn Hoorens, Jeff Rothenberg, Constantijn van Oranje, Martijn van der Mandele en Ruth Levitt, <[http://www.kb.nl/hrd/dd/dd\\_links\\_en\\_publicaties/publicaties/rand\\_report\\_e-depot\\_TR510\\_3c\\_Cover.pdf](http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/rand_report_e-depot_TR510_3c_Cover.pdf)>, retrieved 9 February 2009.

<sup>9</sup> Ras, Marcel (2009), 'Long-term preservation of scientific publications in practice: The KB e-Depot', part of the presentation 'Policies, preconditions and costs' at the *Curating Research* Conference, The Hague, 17 april 2009, <<http://www.kb.nl/hrd/congressen/curatingresearch2009/presentations/beagrie-ras.pdf>>, retrieved 15 August 2009.

<sup>10</sup> *Sustaining the digital investment*, op. cit.