Best until … A National Infrastructure for Digital Preservation in the Netherlands

Barbara Sierman
KB, National Library of the Netherlands
PO Box 90407
2509 LK The Hague
+31 70 314 01 09
Barbara.Sierman@KB.nl

Marcel Ras
NCDD, National Coalition for Digital Preservation
PO Box 90407
2509 LK The Hague
+31 6 147 77 671
Marcel.Ras@ncdd.nl

ABSTRACT

This paper describes the developments in the Netherlands to establish a national Network for Digital Heritage. This network is based on three pillars: to make the digital heritage visible, usable and sustainably preserved. Three working programmes will have their own but integrated set of dedicated actions in order to create a national infrastructure in the Netherlands, based on an optimal use of existing facilities. In this paper the focus is on the activities related to the sustainable preservation of the Dutch national digital heritage.

General Terms

Infrastructure opportunities and challenges

Keywords

Digital Preservation, NCDD, national infrastructure

1. INTRODUCTION

From the very beginning, "collaboration in digital preservation" was a phrase used by many professionals in the field, "as no one institution can do digital preservation on its own". Voluntary collaboration between countries or partners in the same domain (libraries, data centers) found a more firm implementation in organizations like nestor (2003) [1], DPC (2002) [2], NDSA (2010)[3] and NCDD (2007)[1], often serving as a platform for training, knowledge exchange and the study of specific preservation-related issues.

With the growing amount of digital material and organizations involved in preserving it, to foster more collaboration between these organizations by establishing an infrastructure on a national level presents itself as a logical next step. The level of maturity in digital preservation, where now some basic principles are established, will also contribute to this development.

Some examples were already presented. At last year's iPRES the national digital repository of Ireland was discussed, that will host digital collections of a variety of Irish institutes [4]. Darryl Mead from the National Library of Scotland described the effort to create a national preservation infrastructure in Scotland [5]. And in Finland the national library, archives and museums already share an infrastructure. [6] This development is also reflected in Recommendation 3 in the Roadmap of the European 4C Project (Collaboration to Clarify the Costs of Curation), stating, "Develop scalable services and infrastructure", with the explicit benefit of enabling "the realisation of further cost reductions by improving efficiency of the workflows necessary to undertake digital curation" [7].

The Netherlands is no exception in this pattern and is now taking steps to create a national infrastructure for digital preservation.

1.1 The Dutch digital landscape

The term "national digital heritage" does not only cover the collections of the so-called cultural organisations, like archives libraries and museums. It also covers the scientific heritage, as collected by universities, research organisations and data centres. Therefore, the national digital heritage of the Netherlands is not collected in one place but is preserved by a set of national organisations together. Records of the public broadcasting, for example, are preserved by the Institute of Sound and Vision. Research data in the humanities and social sciences are preserved by Data Archiving and Networked Services (DANS) while research data from the technical universities are collected and preserved by 3TU, a collaboration of 3 technical universities. The National Library of the Netherlands is mandated by law to focus on the preservation of "publications" (without an explicit legal deposit law prescribing what should be part of the collection). The National Archive is responsible for governmental archives.

The organisations acted independently in the past but changing requirements in the digital age, both from a point of view of the users of digital material as well as from an efficiency perspective,

[1] On May 21st 2007 a group of organisations took the initiative to set up a coalition to address the problem of digital preservation in The Netherland in a collaborative way. This coalition of the willing became a foundation in 2008 with its mission to establish an infrastructure (organisational and technical) which guarantees long-term access to digital information in The Netherlands.

will require some stronger collaboration and a clear description of roles and responsibilities.

According to figures from the Enumerate data platform for the Netherlands [8], the national collection consists of 44 million museum objects, 845 kilometre of archives, 9 million publications and 1.5 million hours in audio-visual collections.

1.2 Access to the Dutch digital heritage

More and more collections belonging to archives, libraries, media, museums, and knowledge institutes are being digitised and made available online. Institutions are developing functional and technological facilities for making these collections available for digital access and use, simultaneously making the maintenance of these collections cost-effective and sustainable. These are often comprehensive programmes unique to the logic, solutions, and dilemmas that are common in that particular domain.

These are exciting times for archives, libraries, and museums. They are realising that, in the information society, their collections are goldmines. At the same time, the digital environment has made it impossible for them to continue overseeing the entire process of acquiring and managing their collections, and then making them available. For every work process, institutions are often using technology that is developed and managed by someone else.

Institutions that are charged with managing heritage collections and making them accessible are finding themselves in the position of having to redefine their roles. The questions they might ask themselves in this endeavour include:

- How do we reach new user groups? How do we engage them, and what services do we offer them?
- How can we carry out our mission while complying with copyright laws?
- What competencies do we need to be successful in a digital context?
- What are the costs and benefits of making collections available to the public?
- What facilities will we manage ourselves, what services will we purchase, and where will we link to other infrastructures?

The main challenge is to make the national digital heritage accessible for a wide range of users, anytime, anyplace. Developing sector-wide infrastructures and increasing their interconnection will help organisations to do so. By coordinating their IT strategies, parties can achieve benefits of scale and reuse existing building blocks. Making smart connections between collections will enable users to view, experience, and re-use each object in a much richer context. When it comes to digitisation, the major challenges facing the heritage sector relate to scaling up their facilities to be more effective and efficient and linking the collections together to facilitate use.

Being able to meet wide-ranging and constantly changing user demands will depend on having customised and flexible digital facilities. However, upscaling and standardisation are needed to lower costs, improve compatibility, and increase sustainability. This is why any facilities developed must be as reusable as possible.

2. NETWORK DIGITAL HERITAGE (NDE)

Initiated by the Ministry of Education, Culture and Science, the Network Digital Heritage (NDE) was set up in 2014. The participants in this network are national organizations with large digital collections and a mandate to preserve them, like the National Library (Koninklijke Bibliotheek or KB), the Institute of Sound and Vision (BenG), the Cultural Heritage Agency, the Royal Netherlands Academy of Arts and Sciences (KNAW), the National Archives (NA), the Cultural Heritage Agency of the Netherlands (RCE) together with other partners like, for example, the knowledge centre (DEN) and the National Coalition for Digital Preservation (NCDD). This Network Digital Heritage is a partnership that focuses on developing a system of national facilities and services for improving the visibility, usability, and sustainability of digital heritage.

The Network presented a National Strategy for digital Heritage in 2015 [9]. This strategy offers a perspective on developing a national, cross-sector infrastructure of digital heritage facilities. It contains objectives, starting points, and specific work programmes for a joint approach. The national strategy is the result of a one-year process. During this year dozens of professionals from the various sectors have contributed by engaging in working groups, attending meetings, and reviewing texts, including a public consultation.

A general principle is that no new facilities will be developed or new tools will be created, but that (in principal) existing facilities will be used or, if necessary adapted for better and broader use.

Another principle is that these efforts are focused on the user of this national heritage, now and in the future.

Implementing this strategy will require efforts at various levels. Individual institutions will develop an information policy and link their collections, knowledge, and facilities to a larger network. Assigning an active role to five sectorial organisations, so called “hubs”, will reinforce cooperation within sectors. The “hub organisations” have a track record in their domain and long-term sustainability. The “hub organisations” are: KB, BenG, NA, KNAW and RCE. They work within the network as a cross-domain partnership, open for other organisations to join. Commercial parties are explicitly not excluded from this network. Cooperation with industry organisations, user groups, governments, and international networks will be promoted.

The shared strategy must result in more facilities being connected, standardised, and jointly developed and managed in the coming years. This will require more cooperation and knowledge sharing between the various heritage sectors, governments, producers, knowledge institutions, intermediaries, and users. They are working on shared principles, standards, and new methods of knowledge sharing. Agreements and choices sometimes involve a degree of obligation to benefit interoperability or efficiency. This will ensure the development of an infrastructure that is helpful and stimulating for individuals, as well as for large and small institutions, businesses, and governments.

Cooperation will be based on existing sectorial facilities, responsibilities, and funding flows. Working from that foundation, the parties will seek out opportunities for linking and upscaling facilities, as well as for eliminating obstacles. A better understanding of user wishes, the need for a more efficient use of public funds, and the potential of the partnership will reinforce the parties’ readiness to change the existing situation.

Starting from existing facilities and services that have been established in recent years, the Network Digital Heritage defined three work programmes to put the shared strategy into practice. This should help to move from the current decentralized approach in which cultural heritage institutes organised preservation by themselves, towards a more shared approach. Clearly not a centralised approach as the Dutch government will not set up facilities on a must use basis. But helping and stimulating cultural

heritage organisations to make use from (existing) facilities on the basis of sharing.

3. THREE WORK PROGRAMMES

These Work Programmes are initiated to realize the goals set in 2015-2016. Their goals are summarized in the slogan "Zichtbaar, Bruikbaar, Houdbaar", translated as Making digital heritage visible (*Zichtbaar*), Making digital heritage usable (*Bruikbaar*) and Sustainable preservation of digital heritage (*Houdbaar*):

1. Work programme 1 (Visible): Making digital heritage visible (*Zichtbaar*). This should increase the visibility of collections, explore user demand, and promote the use and re-use of digital collections.
2. Work programme 2 (Usable): Making digital heritage usable (*Bruikbaar*): This should improve the possibilities for using collections by making them jointly accessible online, connecting and enriching data using lists of terms and thematic management, and developing targeted services.
3. Work programme 3 (Sustainable): Sustainable preservation of digital heritage (*Houdbaar*). For a preservationist this is the interesting part, although highly connected with the other working groups. The aim is to work on the cross-sector sharing, utilisation, and scaling up of facilities for sustainable preservation and access, while devoting attention to cost management and the division of duties. More details of this Work Programme 3 will be described in paragraph 5.

4. THE NATIONAL COALITION FOR DIGITAL PRESERVATION

The activities in this third work programme are based within the framework of the National Coalition for Digital Preservation, a partnership between the Dutch National Library, the Dutch National Archives, the Dutch Institute for Sound and Vision, Data Archive and Networked Services and several cultural heritage organisations. It is a member organisation, funded by the participating organisations above mentioned with additional funding from the Ministry of Education, Culture and Science.

The NCDD was established in 2008, as a national coalition designed to promote the preservation and the usability of digital materials comprising the cultural and scientific heritage of the Netherlands. NCDD is the national platform for exchange of knowledge and expertise and has a role in coordinating and facilitating the establishment of a national network in which long-term access to digital information, which is of crucial importance for science, culture and society is guaranteed.

A national survey on the state of affairs in digital preservation carried out in 2009 [10] gave a better understanding of the then-present status of digital preservation in the Netherlands. According to the outcomes of this NCDD survey, problems could be best addressed by developing a distributed national network for managing digital resources in the public sector. This infrastructure was understood to include not just storage facilities, but also a whole range of less tangible matters: a clear definition of roles and responsibilities, selection criteria, quality criteria, shared services, knowledge and expertise. The network should be based on collaboration between stakeholders, because the resources required by long-term digital preservation exceed the means of most individual institutions.

Following on the national survey, the NCDD in 2010 formulated a strategic agenda. This agenda consisted of a description of the major steps to be taken on a national level in the Netherlands in order to address the issues described in the 2009 survey. It was also thought necessary to create a sense of urgency towards policy makers on all levels, with the message that we had to act, and act on a national level, to ensure long-term access to digital information. Within the sense of urgency the focal point was the development towards a national infrastructure. Therefore NCDD and especially the partners within the NCDD took the lead in addressing the problem on a policy level, but also on a practical level. It was decided that under the umbrella of the NCDD coalition, the large heritage institutes in The Netherlands would work out a "collaborative model", setting up collaborative facilities or share facilities where possible, which in reality would not always be the case.

In 2013 NCDD made it part of its strategy to work on this collaborative model that should result in a distributed national infrastructure [11]. The first results are becoming available now (spring 2015). A roadmap for certification of Dutch digital repositories has been shaped, workflows for ingest of various types of born-digital materials are described and a scenario for a distributed infrastructure for permanent access has been laid out. This national distributed infrastructure will be based on a reference model developed by the NCDD in which all elements as services are laid out. Services can be everything from storage to preservation watch. The basic starting point is that infrastructures are in place, services are developed and facilities are already shared. But these facilities need to be scaled up, standardised and offered to more and different organizations, sometimes in different domains. These "service seekers" should be enabled to find the best services for their needs and have the professional skills to make the right judgements.

The next steps will be worked out in Work Programme 3 of the NDE, where the current situation will be turned into a networked future.

The efforts of the partners in the NCDD have led to an important result, namely that the preservation issues are addressed on a governmental level, and will be addressed in the goals set in the National Strategy. NCDD is commissioned to lead Work Programme 3 on sustainable preservation in 2015 and 2016.

5. SUSTAINABLE DIGITAL HERITAGE

The objective of Work Programme 3 is to create, through cross-domain collaboration, a shared infrastructure that guarantees sustainable access to digital information. The assumption is that this cooperation will lead to increased effectiveness, greater efficiency and cost reductions. As already described, some of the activities in this work programme have been started and scheduled within the NCDD.

The work towards this goal is being done along three lines:

1. Better utilisation and upscaling of facilities
2. Cost management
3. Roles and responsibilities in digital collection development.

5.1 Better utilisation and upscaling of facilities.

Two examples will be explained here in more detail: storage facilities and persistent identifiers. Apart from these, some other (smaller) projects will be started with regards to, for example,

participating in investigations towards a software repository for tailor made software used in research or art projects, national collaboration in file format research and preservation watch.

Storage facility for permanent access.

This facility will be especially focused on small organisations, which currently have no or hardly any professional facility for permanent storage of digital material. Research showed that this is the case in several specialized cultural and research domains like digital photography, digital art, humanities and architecture. During the 2 years of the program, an inventory of existing storage facilities in cultural heritage organisations in the Netherlands will be created. Based on this list, a small set of representative organisations in the above mentioned domains will be connected to these existing facilities. Apart from making use of facilities of their colleagues in the cultural heritage sector, there will also be the opportunity to make use of commercial partners. A programme for training staff will increase their knowledge of digital preservation. This will make staff more professional and enable them to either manage their digital collections or to outsource this task. Suggested models of service level agreements for different levels of preservation will be designed in collaboration, so that staff will well understand the terms and consequences. This project has a strong connection with the projects in Roles and Responsibilities and Cost Management.

Persistent identifiers

Facilities for assigning persistent identifiers to digital objects need to be implemented and this is highly related to Work Programmes 1(Visible) and 2 (Usable). In close collaboration with existing organisations distributing persistent identifiers (like DataCite Netherlands, Institute of Sound and Vision, DANS and 3TU), existing facilities and their use will be inventoried and a nationwide model will be designed. The goal is to make these facilities affordable and within reach of small and medium cultural heritage organisations as a web service offered by professional providers. A sound overview of the related costs for this facility is needed to estimate operational budgets. There is no intention to restrict the use to a limited set of identifiers.

5.2 Cost management

With regards to costs of preservation, valuable work was done in the Collaboration to Clarify the Costs of Curation (4C) project.[2] The tool they developed, the Cost Exchange Tool (CCeX) will be used to collect cost figures from the main part of organisations in the Netherlands with a preservation mandate, especially the above mentioned "hubs". This will require a different way of reporting from the financial administrations. Within two years it is planned that the main players in the Netherlands will have entered their key figures in CCeX, based on which a benchmark is planned. Training and communications will help to prepare organisations and share experiences.

5.3 Roles and Responsibilities

The activities under this theme aim to achieve a better-integrated way of and collaboration around selection, maintenance and providing access to collections.

Current collection policies are often a continuation of existing collecting policies, established in a physical world. Building digital collections requires evaluating these traditional collection policies, as digital objects poses the boundaries of what, for example, a "publication" is. There is a serious risk that certain digital objects belonging to the Dutch national heritage are not collected at all and that other digital objects are collected by more than one organisation. Getting an overview of "who is collecting what" will lower the risks of gaps or duplicated preservation activities. One of the projects will be related to web archiving, which is done by various organisations in the Netherlands, currently without an overview of the results.

Also the interconnectedness of digital objects will require more streamlining of preservation policies between organisations. A few pilot projects will be undertaken to support a new way of thinking, like archiving Interactive Media Assets and preservation of "enhanced publications" or "digital objects in digital context".

Setting up a Dutch national infrastructure in which facilities will be shared and offered by various organisations, will require a certain level of openness and trust. It will become more important to be open about the various preservation approaches and tune in with other organisations. Sound and published preservation policies will contribute to this openness. Training in preservation planning and watch will be developed to support organisations in developing their own preservation policies. This work will be based on the results from the European project SCAPE, where a Catalogue of Preservation Policy Elements [12] was created, as well as an overview of existing Published Preservation Policies [13].

To establish trust in digital preservation, a set of certifications are available, combined in the European Framework [14], consisting of the basic level Data Seal of Approval, the self-assessment level of the German standard DIN 31644 and the highest level of ISO 16363, the TDR. In the Netherlands a new tool, the Scoremodel, [15] was developed by DEN especially for small cultural heritage organisations, which is a starting point.

Some organisations in the Netherlands already acquired a DSA certificate, like DANS, 3TU. But many of the larger organisations have not. A roadmap has been developed with the aim to get the larger organisations like the National Archive, the National Library and the Institute of Sound and Vision DSA certified before 2018, with other repositories soon to follow.

6. EXPECTED RESULTS AND BENEFITS

The presentation of a national strategy and the establishment of 3 Work Programmes are an important development, which brings many existing initiatives and plans together. This is a start of an integrated approach for access to and preservation of Dutch digital heritage. The timing is perfect as there is a growing community of professionals involved in digital preservation. Exemplary was an expert meeting organised by the NCDD in February 2015 to discuss this proposed infrastructure. On this occasion over eighty Dutch preservationists (and some Belgian colleagues) came together and discussed the national plans, sharing approaches, plans and doubts. The level of knowledge exchange and the willingness to collaborate were promising and proofs that we have made important steps forward. It is a fair promise for the next steps to be taken.

7. REFERENCES

[1] http://www.dnb.de/EN/Wir/Projekte/Abgeschlossen/nestor.html

[2] http://www.dpconline.org/about/dpc-history

[3] http://www.digitalpreservation.gov/ndsa/about.html

[2] http://4cproject.eu/

[4] Sharon Webb, Aileen O'Carroll The process of building a national trusted digital repository: solving the federation problem. Ipres 2014 Proceedings, https://phaidra.univie.ac.at/detail_object/o:378066

[5] Darryl Mead: Shaping a national consortium for digital preservation. iPRES 2014 Proceedings, https://phaidra.univie.ac.at/detail_object/o:378066

[6] Towards Preserving Cultural Heritage of Finland Heikki Helin, Kimmo Koivunen, Juha Lehtonen, Kuisma Lehtonen, 2012 NBN: http://nbn.depositolegale.it/urn:nbn:it:frd-9299

[7] Investing in Curation. A shared Path to Sustainability http://www.4cproject.eu/d5-1-draft-roadmap

[8] http://enumeratedataplatform.digibis.com/datasets

[9] http://www.den.nl/art/uploads/files/Publicaties/20150608_Nationale_strategie_digitaal_erfgoed_Engels.pdf

[10] http://www.ncdd.nl/documents/NCDDToekomstDEF2009.pdf

[11] http://www.ncdd.nl/documents/NCDDToekomst_2_Strategischeagenda.pdf

[12] http://wiki.opf-labs.org/display/SP/Catalogue+of+Preservation+Policy+Elements

[13] http://wiki.opf-labs.org/display/SP/Published+Preservation+Policies

[14] European Framework, see http://www.trusteddigitalrepository.eu/Welcome.html

[15] DEN Scoremodel Digitale Duurzaamheid http://www.den.nl/standaard/383/