

Executive Summary

As the largest public archive in the Netherlands, the 200 year old Nationaal Archief plays an important role in safeguarding the Dutch memory. Its collections document the activities of central government and record the evolution of the Dutch state.

Archives transferred by the government to the Nationaal Archief increasingly comprise digital material. Fast technological developments, causing hardware and software to become obsolete within just a few years, present a challenge in keeping these records authentic and accessible through time. The Nationaal Archief is building a Digital Depot to cater for digital records and plans to share its acquired knowledge and facilities with other government institutions and other Dutch archives. It will thus provide a “smart hard disk” for storing and accessing Dutch history.

In the Digital Depot, which combines an automated system and skilled staff, the focus has been on the ingest functions and safe storage. Preservation functions such as keeping an authentic version of the ingested material accessible are not yet in place. The Nationaal Archief aims to use solutions developed in Planets to provide this preservation component.

This case study focuses on the way the Nationaal Archief envisages implementing two emulation approaches to address these challenges: emulation for rendering and emulation as an intermediate step in migration. Though much work remains to be done, for the Nationaal Archief, Planets has meant a big step forward.

Authors

Petra Helwig, The Nationaal Archief of the Netherlands
Bill Roberts, The Nationaal Archief of the Netherlands
Emily Nimmo, The Humanities Advanced Technology
and Information Institute

This case study is part of a series of case studies on the application of Planets in major European libraries and archives.

They are all available via the Planets website, www.planets-project.eu

The Nationaal Archief and its collection

As the largest public archive in the Netherlands, the 200 year old Nationaal Archief plays an important role in safeguarding the Dutch memory. Its collections document the activities of central government and record the evolution of the Dutch state. From military conflicts, the political constitution, economic policy and developments in agriculture, trade, industry, healthcare, voyages of discovery to acquisition of overseas dominions and public works such as dykes, roads, bridges and churches. Through these collections it becomes possible to discover more about the lives of Dutch citizens, diplomats, civil servants and politicians and those who have come into contact with the Dutch government such as refugees and Asian monarchs. The Nationaal Archief also holds private archives of people or organisations that have played an important role in the political or social history of the Netherlands; or that illustrate aspects of Dutch society, such as the Royal Dutch Football Association (KNVB) and the Netherlands Olympic Committee (NOC*NSF).

New challenges and opportunities

In line with society as a whole, the activities of Dutch government are increasingly becoming digital. Agencies communicate, create and store documents online internally and externally. Archives transferred by the government may exist now only in a digital format ("born digital"). This presents new opportunities to the Nationaal Archief as digital material can be accessed independently of physical location. Materials can be passed over much earlier than the statutory maximum of 20 years without making it any more difficult for agencies to access them. This adds to the transparency



The Nationaal Archief of The Netherlands

and openness of government and availability of information for citizens, policy makers and members of parliament.

Of course, these digital developments also present new challenges. The fast pace of technological development and change can cause hardware and software to become obsolete within just a few years. Accessing digital records once the software used to create, view and interpret them has become obsolete presents an issue. Transferring it to new formats may result in changes to the record that could affect how true they are to the original and in turn their archival and legal status. For institutions such as the Nationaal Archief, with a mandate to safeguard national heritage, these challenges, left unchecked, could ultimately lead to the loss or misinterpretation of history.

The Nationaal Archief recognises these problems cannot be solved in isolation. Expert knowledge and skills are needed and separate facilities for each institution will be very costly. The Nationaal Archief plans to share its acquired knowledge and facilities with other government institutions and other Dutch archives, and thus provide a "smart hard disk" for storing and accessing Dutch history.



«Emulation is a preservation strategy of world wide importance. Not only archives and libraries but also museums that keep digital artefacts will benefit from a trustworthy infrastructure for emulation. International projects like Planets bring together the required knowledge and experience to develop these solutions, which will in turn enable the Nationaal Archief to offer reliable services to its partners.»

Jacqueline Slats, Manager, Digital Innovation

The Digital Depot

Dutch law requires the Nationaal Archief to preserve the paper and digital records transferred to them “forever”. The Nationaal Archief has developed the Digital Depot to cater for digital records – both born-digital and a huge amount of archival material that has been digitised. The Depot combines an automated system and skilled staff to maintain, preserve and provide access to them. Pilot projects to ingest material transferred by various government agencies are underway and, in November 2009, the system was officially put into service with the transfer of a set of records from the Ministry of the Interior in the presence of the minister. It will become fully operational in 2010.



Guusje ter Horst,
Minister of the Interior

The system enables government agencies to transfer their digital records to the Nationaal Archief in an automated, controlled and structured way. In turn, the depot will be available as a “shared service” to other regional Dutch archival institutions - the “Regional Historic Centres” - for preservation of electronic records transferred to them.

To date, the focus has been on the ingest functions and safe storage. Preservation functions such as keeping an authentic version of the ingested material accessible – are not yet in place. NANETH aims to use solutions developed in Planets to provide this component.



The Digital Depot

The need for emulation in preservation

Different digital objects have different preservation requirements, depending on why the record is being preserved, how long it needs to be preserved, and the context of the record. According to Dutch archival regulations, what has to be preserved depends on the requirements of the working process of which the record was a part. It is not always necessary to preserve every aspect of the record.

From 2000–2003 the Nationaal Archief carried out an extensive research project, known as “The Dutch Testbed”, into the feasibility of different preservation strategies for different types of digital objects. These included: email, text, spreadsheets and databases as they were expected to make up the bulk of transferred material.

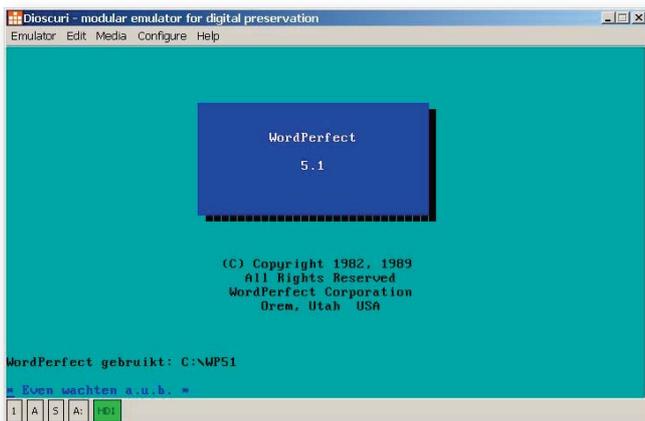
Where it is most important from an archival perspective to preserve the content of the digital object such as e-mail and texts, the Dutch Testbed concluded migration to be the most effective strategy. In other circumstances, e.g. where there is a significant interactive element, and the nature of that interaction is important from an archival perspective, or where there is a need to preserve the original “look and feel” of the object, it may be necessary to maintain the ability to run the original software application.

Emulation provides a way to keep obsolete software running and offers the most reliable way for a reader to see – or hear – the record in its original form. The Testbed concluded emulation could thus play an important role in a comprehensive approach to preservation for e.g. spreadsheets and databases.

An overview of emulation

So what is emulation exactly? There are many types but the most common and most useful is “hardware emulation.” This means creating software which runs on one computer (the “host system”) and makes it behave exactly like another – usually older – computer (the “target system”). Software which runs in the target system can be run inside the emulator in the host system. This includes both operating systems and applications and supporting software components such as device drivers.

For example, 20 years ago, WordPerfect 5.1 was the most popular word processing package. WordPerfect files from that period cannot be read by Microsoft Word or most other present day software packages without the risk of considerable changes to the “original” object. We can use an emulator for a 286 PC and install DOS and WordPerfect on the emulated computer. Then we can open a WordPerfect file in the original application and run it in the emulated environment, allowing us to work with the file in the same way as the creator did, 20 years ago.

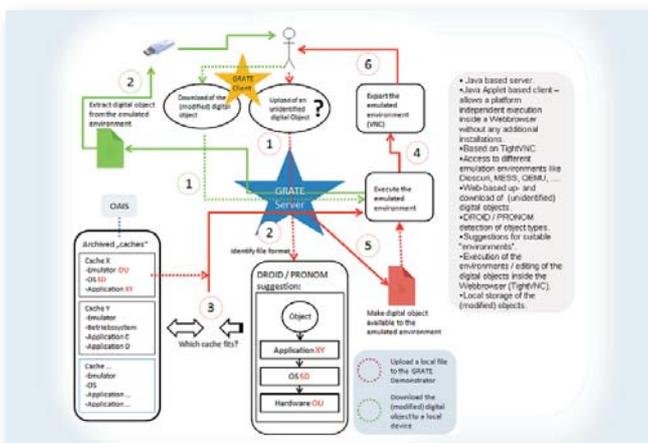


Word Perfect running in Dioscuri

In 2005 the Nationaal Archief and the Koninklijke Bibliotheek (Dutch Royal Library) set up a joint project to create Dioscuri, a modular emulator designed for digital preservation. Dioscuri was further developed within Planets and forms one of the emulation services in the Planets suite of preservation tools.

Installing and configuring an emulator can be complex for an average computer user. Until recently the only way to give a user access to an emulated object was to set up the emulator on-site in the reading room of the archive.

Planets has tackled these issues using the GRATE framework, developed by partner the Albert Ludwig University of Freiburg. The GRATE service provides remote access via the web to a selection of centrally-hosted emulators. Users can connect to the remote emulator from their home or office using their internet browser, without needing to know anything about how to install or start the emulator.



The GRATE framework

Digital objects can be uploaded to GRATE, which transfers the file to the emulator, typically by wrapping it up inside a virtual floppy disk which can be "inserted" in the emulated computer. Any outputs from work carried out inside the emulator can be saved back to this virtual floppy disk and then extracted for storage or download.

Emulation keeps options open

Most institutions agree that, even if they plan to adopt migration as their primary preservation strategy, it is important to preserve the original files. Indeed this was one of the recommendations of the Dutch Testbed project. It implies that there must be a continuing means by which to render the files in the future.

While, with progress, more and more reader software and viewers become available for older file formats, we cannot be sure these will continue to be available especially for less commonly used file formats.

Most format migration processes are not reversible with 100 percent accuracy. Migrations are designed to preserve the features of a record that are considered important for its designated purpose at the time of migration, but other aspects may be lost or altered. Such decisions cannot predict future historical interest with certainty and aspects of the record may be lost. Future technological capabilities may also permit processing or exploitation of the digital object in a new way. By retaining the original, we maximise future preservation and access options. Emulation provides a means by which to do this. There are similarities with developments in the field of archaeology: In the 19th century, the typical approach was to dig up an artefact, maybe record where it was found and put it in a museum. The artefact itself was considered to possess all the "significant properties". In digital terms it is comparable with migration where significant properties are retained but it is not possible to go back to the original. In the 20th century, archaeologists realised the site where artefacts were discovered may also provide useful contextual information. A more recent approach is to dig, describe and take pictures but leave the artefact where it is and close the pit again, or to use scanning techniques. In this way, future archaeologists will be able to extract more and more information – significant properties – that we cannot yet think of. Retaining the original files and using emulation provides access to the object in the original environment and makes it possible to keep preservation and access options open – it is a contemporary approach to "digital archaeology".



Archaeological site, Kent, UK

Emulation at the Nationaal Archief

The Nationaal Archief has two important tasks to fulfil in relation to digital material: to give readers access to the files it has stored and to ensure this material remains accessible indefinitely. It envisages implementing two emulation approaches to address these challenges: emulation for rendering and emulation as an intermediate step in migration.

Emulation for the reader (rendering)

Database backed applications used to aid government decision-making are a good example of types of material where emulation may be the best way to render the digital object properly. In some cases databases can be preserved by migrating the contents to a standard format, for example using the SIARD software developed by Planets partner, the Swiss Federal Archive. In other cases, it may also be necessary to preserve the application used to work with the data, to reproduce the information as it was originally used. Emulation is a good approach for preserving the ability to run an application over the long term. The Nationaal Archief is considering how the GRATE service could be used to provide emulator access to selected records, either for visitors to the archive reading room, or remotely over the web.

Emulation as an intermediate step in migration

The Nationaal Archief also sees a role for emulation as an intermediate step in migration. While migration can offer the most appropriate preservation strategy for objects such as text documents, emulation can play an important step in the process.

Research undertaken in the Dutch Testbed has shown that when migrating to another file format, results improve when the original bitstream is saved in the latest available version of the software, before migration to another file format. For example a WordPerfect 4.2 file should be opened and saved in WordPerfect 10 first, before being migrated to Microsoft Word. Later versions of the original software often run only in the original hardware environment and emulation is the only available option to re-create these environments.

This approach is also useful if the digital object has gone through a long “unmanaged” period since its creation. For example: the donation of a personal archive collection of an individual of cultural or historical importance. Such collections can contain digital material in obsolete formats that have not benefitted from a managed process and are inaccessible with present day software. Emulation can be used to run the original software application before the files are converted to a more accessible format.

As an example, the Nationaal Archief has a number of storage media (hard disks, DVD's, floppy disks), which were transferred to it alongside paper records. These disks contain files in a range of old formats, such as WordPerfect. Emulation could be a means to re-create the original context as a preliminary step in preparing these files for managed ingest in the Digital Depot.



Examples from a software collection

What's next?

For the Nationaal Archief, projects like Planets are vital. They offer a platform to bring knowledge together, to work with colleagues to achieve results that no institution could reach by working on its own. Planets has produced a suite of working software tools backed up by the understanding of when and how they can be used.

As the Nationaal Archief further develops its Digital Depot, the emulation tools and knowledge gained through Planets will be part of the preservation strategy for the growing digital collection. There are still issues to investigate: for example how best to establish a software archive to maintain old operating systems, software applications and supporting components; and how to manage the IPR and licensing issues around using old software in new ways.

The emulation tools we have will need further development and new emulators will be needed: we look to the Open Planets Foundation and other EU projects to maintain the momentum that Planets has established.

Conclusions

The Nationaal Archief of the Netherlands holds collections of national and international importance which are increasingly digital in format. The fast pace of technological development and change can cause hardware and software to become obsolete within just a few years. Accessing digital records once the software used to create, view and interpret them has become obsolete presents an issue which, as the Nationaal Archief recognises, cannot be solved in isolation.

Therefore, for the Nationaal Archief, projects like Planets are vital. Planets has produced a suite of working software tools backed up by the understanding of when and how they can be used. The Nationaal Archief aims to use solutions developed in Planets to provide the preservation-component of its Digital Depot and envisages implementing two emulation approaches to address these challenges: emulation for rendering and emulation as an intermediate step in migration. Though much work remains to be done, Planets has meant a big step forward.



«Nationaal Archief has always been a strong advocate of emulation as a preservation strategy. In Planets, substantial progress has been made towards a practical tool that can potentially play an important role in our digital repository.»

Remco Verdegem, Programme Manager, Infrastructure

Acknowledgements

The authors would like to thank all those who have contributed to this case study.



Planets (Preservation and Long-term Access through NETworked Services) is a four-year, €15 million project, co-funded by the European Commission under the Information Society Technologies (IST) priority of the 6th framework Programme (IST-033789).

The project has developed a suite of tools and services to support preservation of digital content for the long-term. Planets tools make it possible to define digital preservation goals and policies; understand the characteristics of a collection; build, evaluate and execute preservation plans, convert objects into up-to-date and accessible formats and run software on legacy operating systems. It offers an automated solution to support informed decision-making and justify actions taken.

Planets is coordinated by the British Library and has been delivered by a Consortium of 16 national archives, libraries, research institutions and leading IT companies.

Further Information

For more information about Planets visit: <http://www.planets-project.eu>

You can email your questions to us at:
info@planets-project.eu



The Open Planets Foundation (OPF) builds on the investment made in the Planets project. It will sustain the results of this investment and further develop and coordinate development of the capabilities that its members require. It will provide services, knowledge, methods and tools to its members and the broader community.

For further information visit: <http://www.openplanetsfoundation.org>