From Oais to DPS to NDHA

Steve Knight; Associate Director National Digital Library; National Library of New Zealand

Background

In New Zealand, the issue of preserving the nation's digital cultural heritage, past, present and future is addressed by legislation and central government policy. The National Library of New Zealand (Te Puna Mātauranga o Aotearoa[1]) Act 2003 requires the National Library to collect, preserve, protect and make accessible digital collections, along with traditional paper collections, in ways that ensure current and future access to New Zealand's documentary heritage.

As well as being in line with the Library's New Generation National Library[2] strategy the National Digital Heritage Archive (NDHA) also fits with government strategy in the digital arena, in particular the Digital Strategy 2.0[3] and the Digital Content Strategy[4].

This abstract briefly describes the National Library's response to this legislative mandate for digital preservation (development of the NDHA) and the challenges of implementing and sustaining a digital preservation programme. It is important to be clear that digital preservation is a larger concern than the software and hardware elements that support it, thus the title of this abstract – From OAIS to DPS to NDHA.

National Digital Heritage Archive

The National Digital Heritage Archive (NDHA) Programme was established in July 2004 with a view to ensuring that the Library has the infrastructure, technology and organisational structure and work processes to preserve and provide access to the digital heritage collections it is responsible for.

The National Library requires a system that secures the integrity and authenticity and therefore trustworthiness of digital material deposited with the National Library.

> *"A National Library is a place where a nation nourishes its memory and exerts its imagination – where it connects with its past and invents its future."[5]*

Due to be completed in 2009 the NDHA is being developed and implemented in partnership with Ex Libris Group and Sun Microsystems as a commercially viable solution to digital preservation. It has been developed in line with general access and preservation trends and is designed to meet the needs of a range of institutions now and in the future. An international Peer Review Group consisting of recognised leaders and innovators in digital preservation provides independent advisory services for the partnership.

The Digital Preservation Problem

Digital preservation is too often characterised as a technological problem similar to the following:
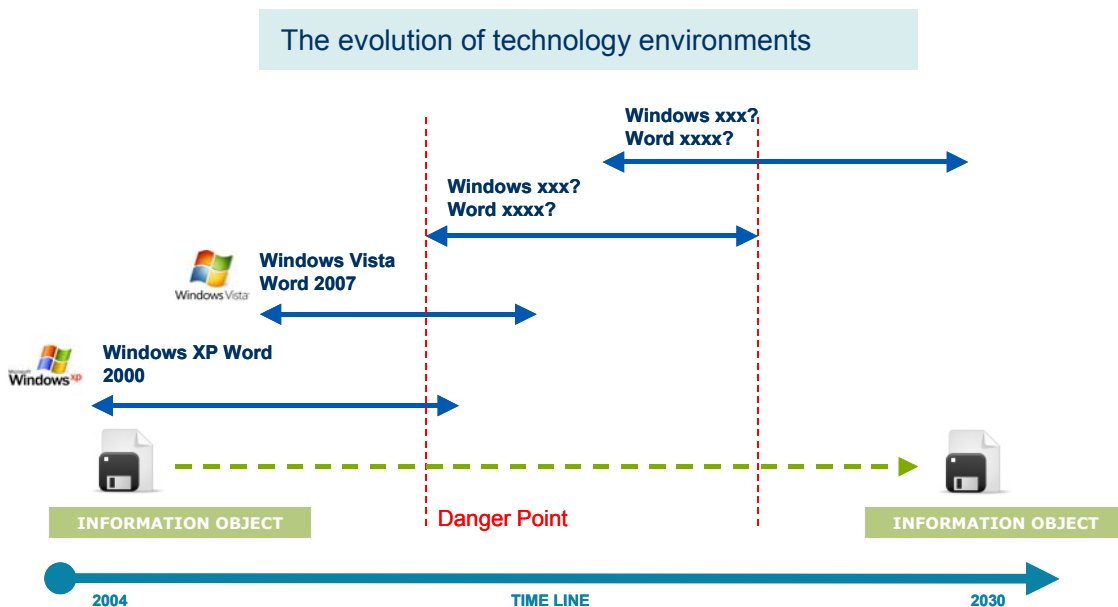
Concurrent to software and hardware development, organisational readiness to integrate the new system is being addressed. Digital preservation must be seamlessly integrated into the organisation's overall business and technology infrastructure including:

- existing collection management systems
- deposit applications development
- new and existing resource discovery and delivery systems and access products reporting systems
- common services
- data migration.

The Digital Preservation Problem

Digital preservation is too often characterised as a technological problem similar to the following:

However, the challenge of digital preservation is much more germane to social, cultural and psychological issues than it is to issues of technology. This was clearly articulated **by Garrett and Waters when they noted that:**

> *'the problem of preserving digital information for the future is not only, or even primarily, a problem of fine tuning a narrow set of technical variables. It is not a clearly defined problem ... rather, it is a grander problem of organizing ourselves over time and as a society to maneuver effectively in a digital landscape. It is a problem of building ... the various systematic supports ... that will enable us to tame the anxieties and move our cultural records naturally and confidently into the future.'*[6]

Why do we talk about the notion of a Trusted Digital Repository? Surely we are attempting to enter into some form of pact with the future, with researchers, geanealogists, scholars, students, those requiring material for evidential purposes, a plethora of users and usages that we an only guess at from out current vantage points.

This notion that any ''preservation enbiroment manages communication from the past while communicating with the future' has has been recently canvassed by Reagan Moore in the context of a potential 'theory of digital preservation' and an articulation of the necessary requirements for sustaining this 'validation of communication from the past.'

Current National Library of New Zealand work is focused on these issures, in particular provenance, context, authenticity and integrity.

Some Small Steps Towards a Viable Dialogue with the Future

In order to pursue this dialogue with the future there are some things that we, as a community, need to address. These include language, products/tools and services, quality assurance and confidence, standards and best practice and a cohesive managed approach to the challenges ahead, both research and practical.

1   What do we mean when we talk about digital preservation? Where is the agreed definition of what digital preservation comprises at a granular level, ie what are the business and functional requirements for digital preservation that will provide us as practitioners and vendors as suppliers with the knowledge we need to do digital preservation?

When we talk about repositories, data archiving, digital archiving, life cycles, digital curation, data curation or digital preservation, does the language we use assist or hinder us in our understanding of digital preservation?

2   What are the economic models for sustaining our digital preservation activities? Recent research notes that 'in many institutions and enterprises systemic challenges create barriers for sustainable digital access and preservation'[7] including:

- inadequacy of funding models to address long-term access and preservation needs
- confusion and/or lack of alignment between stakeholders, roles, and responsibilities with respect to digital access and preservation
- inadequate institutional, enterprise, and/or community incentives to support the collaboration needed to reinforce sustainable economic models
- complacency that current practices are good enough
- fear that digital access and preservation is too big to take on.

What are we not doing that allows these barriers to remain? Is it that we are in the early phases of development of a new discipline and organisational requirements to support digital preservation will emerge in time? Or are we not succeeding in getting the message across that digital preservation is the biggest challenge facing our institutions now and on an ongoing basis?

3   What is it about the current products, tools and services that we use for validating our digital preservation work practices that gives us confidence that they are doing what they should? For example, there are several tools for characterising, validating, extracting data from and managing file formats. These tools are used almost blithely in our digital preservation workflows even though we know that there are problems with them. What does this say about the authenticity and integrity of the objects within our preservation repositories?

Concern about formats is concern about risk management. We need a comprehensive management approach, strategies that identify the risk of format problems/obsolescence and strategies that mitigate the risk of format problems/obsolescence. We need to be able to identify specific files that are most at risk and we need ready access to detailed, accurate information describing file formats. We need centralised registries to support format management and we need an agreed set of risk grading criteria for formats to aid in preservation risk management and planning.

This is a real world problem right now and will be an ongoing problem for those preservation repositories that do not define their selection by format. A solution would generate real value-add to the whole community and surely there must be benefit in standardising the tools we use for identification, validation, extraction etc?

4   Where do we look to for advice on standards and best practice? There is an increasing array of digital preservation projects, models and practices – OAIS, PREMIS, NARA, PLANETS, CASPAR, NDIIPP, SHAMAN, DURASPACE, HathiTrust – but how do we know what to trust?

Similarly where do we look for certification and audit of our systems, repositories, organisational capability, sustainability? While effort has been put into the development of tools such as Drambora[8] and TRAC[9] it is still not clear yet whether these will be effective mechanisms for monitoring our digital preservation activities. Work to be undertaken in 2009 by the Center for Research Libraries[10] in **the US should add to our knowledge in this space.**

Conclusion

This abstract has attempted to provide a brief description of the work and central concerns of a small national library in addressing digital preservation. There is no other issue facing our institution at the moment which will have as deep an impact on our ongoing activities.

How we configure our resources to respond to technological change and innovation, how we respond to citizen's created content and its impact on our collection and description processes, how we define, resource and pursue our internal research agenda (understanding the web, science data sets etc) will all need to be approached through the lens of digital preservation.

Content (ie digital preservation) systems will be our core operational systems and digital preservation will be increasingly seen as fundamental to the notion of a national knowledge infrastructure.

While some of the issues raised above may seem negative or intractable, quality assurance of products, tools and services, a professional services market (commercial or otherwise), coordinated national/international approaches are merely next steps and what we may be seeing is the exigencies of practice overtaking the research agenda within a still emerging digital preservation discipline.

References

[1]The Māori name of the National Library means "well-spring of knowledge".

[2]New Generation National Library Strategic Direction to 2017. 2007. http://www.natlib.govt.nz/catalogues/library-documents/strategic-directions-to-2017.

[3]http://www.digitalstrategy.govt.nz/upload/Documents/Digital%20Strategy%202.0%20FINAL.pdf

[4]http://www.digitalstrategy.govt.nz/Resources/New-Zealand-Digital-Content-Strategy/NZ-Digital-Content-Strategy---Html/

[5]Ryckmans, Pierre. 1996. Perplexities of an electronically illiterate old man. Quad-rant, September 1996, No 329.

[6]Garrett, J. & Waters, D. (eds). (1996). *Preserving digital information: Report of the Task Force on Archiving of Digital Information.* Washington, DC: Commission on Preservation and Access and the Research Libraries Group. http://www.oclc.org/programs/ourwork/past/digpresstudy/final-report.pdf

[7]Moore, Reagan. 2008. *Towards a theory of digital preservation.* The International Journal of Digital Curation, V3 (1) 2008. http://www.ijdc.net/index.php/ijdc/article/viewFile/63/42

[8]Blue Ribbon Task Force on Sustainable Digital preservation and Access. 2008. *Sustaining the digital investment: Issues and challenges of economically sustainable digital preservation.* .

[9]http://www.repositoryaudit.eu/

[10]http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=

[11]http://www.crl.edu/content.asp?l1=13&l2=58&l3=181