Conference of Directors of National Libraries Annual Meeting

Lyon, France 19 August 2014

Workshop on 'Preservation of digital heritage: regional and international cooperation as a solution'

Background

CDNL requested a workshop be prepared for CDNL members at their annual meeting in Lyon on 19 August 2014. The workshop was to address the idea of 'Preservation of digital heritage: regional and international cooperation as a solution'.[1]

The 2013 IFLA Trend Report posed the question 'How will we access, use and benefit from information in an increasingly hyper-connected world?'

The CDNL Workshop posed the question 'What role should national libraries play in ensuring the long-term sustainability of digital assets so that we can access, use and benefit from information in an increasingly hyper-connected world?'

The workshop builds on prior CDNL work on the Global Digital Library which notes the need to 'promote the development of digital tools, products and services and common standards … **with a particular focus on critical issues for national libraries such as long-term digital preservation**'.

Recommendations

It is recommended that CDNL:
- note the contents of this write-up of the 'Digital heritage: regional and international cooperation as a solution' workshop at Lyon on 19 August 2014
- agree to set up a working group of CDNL members to determine whether CDNL wish to commit to a global, national libraries based cooperation regarding digital preservation
- agree to set up a working group of CDNL members to define a roadmap for addressing digital preservation pursuant to commitment in line with the previous recommendation.

---

[1] Workshop and subsequent report were prepared by Steve Knight (National Library of New Zealand) and David Pearson (National Library of Australia).

IFLA Trend Report 2013

The IFLA Trend Report takes a broad approach in assessing the impact of new technological developments on society as a whole. The 2013 report identified five high level trends shaping the information society, spanning access to education, privacy, civic engagement and transformation. These five trends were:

- New technologies will expand access to information for some people and limit access for others
- Online education will democratise and disrupt education around the world.
- The boundaries of privacy and data protection will be redefined.
- Hyper-connected societies will listen to and empower new voices and new groups.
- The global information environment will be transformed by new technologies.

The 2013 report also explicitly acknowledged the need for digital preservation stating that 'preserving our digital heritage is a priority for libraries and archives around the world' and noted the importance of 'identifying and capturing digital content of historical or cultural significance in the information deluge'.

The challenge of digital preservation

Digital preservation combines policies, strategies and actions to ensure access to reformatted and born digital content regardless of the challenges of media failure and technological change. The goal of digital preservation is the accurate rendering of authenticated content over time.[2]

While based on the same concepts and ethical principles as traditional conservation practice, digital preservation is a new challenge as our digital collections continue to grow in both volume and complexity.

One of the signature challenges of digital preservation is that it is an ongoing activity which may require innovative solutions to address. Individual institutions may not have the capacity to do it and may need to look at a consortium approach to fulfil their mandates. Equally, some institutions may be able to do something but sustaining a long term programme may be problematic. Perhaps outsourcing models for resources and skills will need to be considered. Which raises questions of jurisdiction the extent to which we are prepared to lose authority and control?

In addition, there are challenges to our business models, our collecting strategies, and our workforce planning - a fundamental problem is access to technology or having the skills to access technologies. Without these skills, discussions about collecting or preservation or access become moot.

Perhaps the greatest challenge now and increasingly in the future is the sheer cost of storing and managing large-scale digital collections. And if our digital collections are not large-scale now, they will be in 10, 20, 50 years' time.

---

[2] ALA Definitions of Digital Preservation – Medium Definition, 2007. http://www.ala.org/alcts/resources/preserv/defdigpres0408

Workshop questions and discussion

The workshop agreed that:
- Preserving our digital heritage is a priority for libraries and archives around the world?
- Failure to look for trusted means and methods of digital preservation will certainly exact a stiff, long-term cultural penalty?[3]
- There is a resourcing problem – eg system librarians are generally not the right people to deal with digital preservation issues.

The workshop also agreed that digital preservation is as much a social, cultural and psychological exercise as it is a technical one. Technology is the means to ensuring the long term safekeeping of digital objects into the future while maintaining their renderability and their meaning.

However, we also need to be aware that in some disciplines the technologies are going to be key to understanding and providing access to digital collections, particularly in the context of 'big data' and in making our large digital collections computationally available for researchers. New insights can be gathered from the outcomes of previous research.

This raises the question of the role of the library and national libraries especially in the collecting, managing and preserving of the outputs of science and research. Are we the natural home for this activity and what capability do we require in order undertake this activity?

It was also noted that the cultural can also be seen as 'physiological', eg in the preservation of indigenous practices which often also includes traditional perceptions related to sharing and restricted access.

People or machines

The workshop also addressed the question of whether digital preservation is a mechanical process, eg characterised by storage management, fixity assurance and periodic technology refreshment. Or, is it more a social, cultural and psychological process, eg characterised by provenance, context, chain of custody and ensuring understanding of change over time?

This led to a discussion of how do we determine whether an object is what it purports to be (and remains uncorrupted and free of unauthorised and undocumented change) and can we leave our digital collections 'safe' in the hands of technology.

---

[3] Garrett, J. and Waters, D. 1996. Preserving Digital Information: Report of the Task Force on Archiving of Digital Information. Commission on Preservation and Access and The Research Libraries Group. Page 4. http://www.clir.org/pubs/reports/pub63watersgarrett.pdf

The workshop agreed that we need a mixture of machine and human mediated, curatorial oversight to provide assurance that a viable chain of custody has been provided for our digital collections over time?

Digital preservation practitioners have an interest in the technologies that impact on our ability to understand, manage and provide access to our digital collections. For example, we are interested in the idea of formats, the container an object arrives in (eg an image might arrive in TIFF or png format). But we are also interested in what is required to ensure that the content within the container does not change over time. Digital practitioners (or the digital preservation programme) are therefore a hybrid resource within our institutions, requiring both technological and curatorial skills and focus.

Break out session

The workshop broke up into six groups to determine the top 3 factors related to implementing a digital preservation programme.

Group	Factors
Group 1	1. Lack of legislative support for digital collecting and digital preservation 2. Adequate funding
Group 2	1. Noted role of digital preservation in many countries 2. Need to identify essential national networks 3. Need to make digital preservation a priority for the public and lawmakers
Group 3	1. Need to be able to use and re-use digital content collected today in a hundred years time and beyond 2. Public involvement in what digital materials need to be preserved
Group 4	1. Cooperative guidelines for what needs to be collected 2. Develop tools collaboratively including with the private sector who need to be a strong actor in this space
Group 5	1. Issues with big data, managing complexity – cloud, internet – need more technology and skilled staff 2. Non-linear evolution has the potential to kill libraries – ie the sheer quantity, variety and complexity of material over time 3. potential for 'world library infrastructure' to provide local, national competitive advantage
Group 6	1. Policy development 2. Advisory and support services 3. Digital collecting policies and content access policies for access by citizens

Cooperation as a solution

Having agreed that:

- Preserving our digital heritage is a priority for libraries and archives around the world
- Failure to look for trusted means and methods of digital preservation will certainly exact a stiff, long-term cultural penalty
- We need curatorial oversight to provide assurance regarding the authenticity and integrity of our digital collections over time?
- Currently we are not addressing the problem satisfactorily (thus the title of the workshop),

the question that remained was 'what role should national libraries play in ensuring the long-term sustainability of digital assets so that we can access, use and benefit from information in an increasingly hyper-connected world?'

It was noted that the way CDNL is currently addressing digital preservation is unsatisfactory and lacks collaborative intent and practice.

It was also agreed that it is the function of national libraries to do this work but that most do not have the requisite resources and the question was raised regarding the role of more developed libraries?

Considerations for CDNL

The workshop agreed that cooperation will be one of the keys to success in the digital preservation domain.

This leaves the question of what activities does CDNL wish to pursue to ensure accurate rendering of authenticated content over time? What does it look like for each CDNL institution to be a node in a global network of authoritative digital preservation practitioners and repositories?

Some considerations might include:

- Legal, strategic and policy frameworks – we need to get this right – we need to take the lead
- People resources – how do we train or select the right people
- working together to develop the right tools and processes
- Develop more robust networking, sharing and cooperation mechanisms in this domain
- Develop shared, open responses to the technology world as it evolves – cloud, big data, robustness, scalability.

Appendix One – Cooperation – What Can we Do Together?

Category	Description	Cost (H, M, L)
What do we need in our digital preservation programmes?	Do we know enough or know what is important? Ingest • Pre-conditioning • Authenticity pre-ingest • Tracking handling • Bulk workflows	M
Shared strategy and policy	Are we so different that we each need our own suite of tools and workflows? • Digital preservation • Digitisation • Digital collecting • Legal issues – Legal Deposit • Resourcing – the shape and cost of digital storage • Resourcing – the lack of high quality specialists in the domain • Critical infrastructure decisions	L
Tools development and enhancement	Ensuring that the tools we use are fit for purpose: • JHOVE • PRONOM/DROID • NLNZ Metadata Extract Tool • NSLA Digital Preservation Technical Registry	M
Preservation standards	What do compliance / conformance mean for national libraries? • PREMIS	L
Workforce planning	Resourcing / staff exchanges: • Strategic • Technical • Operational	L
Certification and audit	Should national libraries settle on one? • ISO 16363:2012 • NDSA levels of digital preservation • DRAMBORA • Data Seal of Approval • DIN 31644 – Nestor • ED3 • A new one was presented at IS&T in 2014	M
Research and practice	Bridging the gap for national libraries:	H

	• Repository analysis • Data mining (interrogating large data sets) • Visualisation • Preserving email systems • Preserving databases	
TREASURES	Supporting the development of sustainable global digital technical registry for digital preservation (currently a Horizon 2020 bid)	M
PERSIST	Supporting the UNESCO Digital Roadmap to bring industry and cultural heritage together to support the digital preservation agenda	M