

PASIG EMEA Event for Public Sector



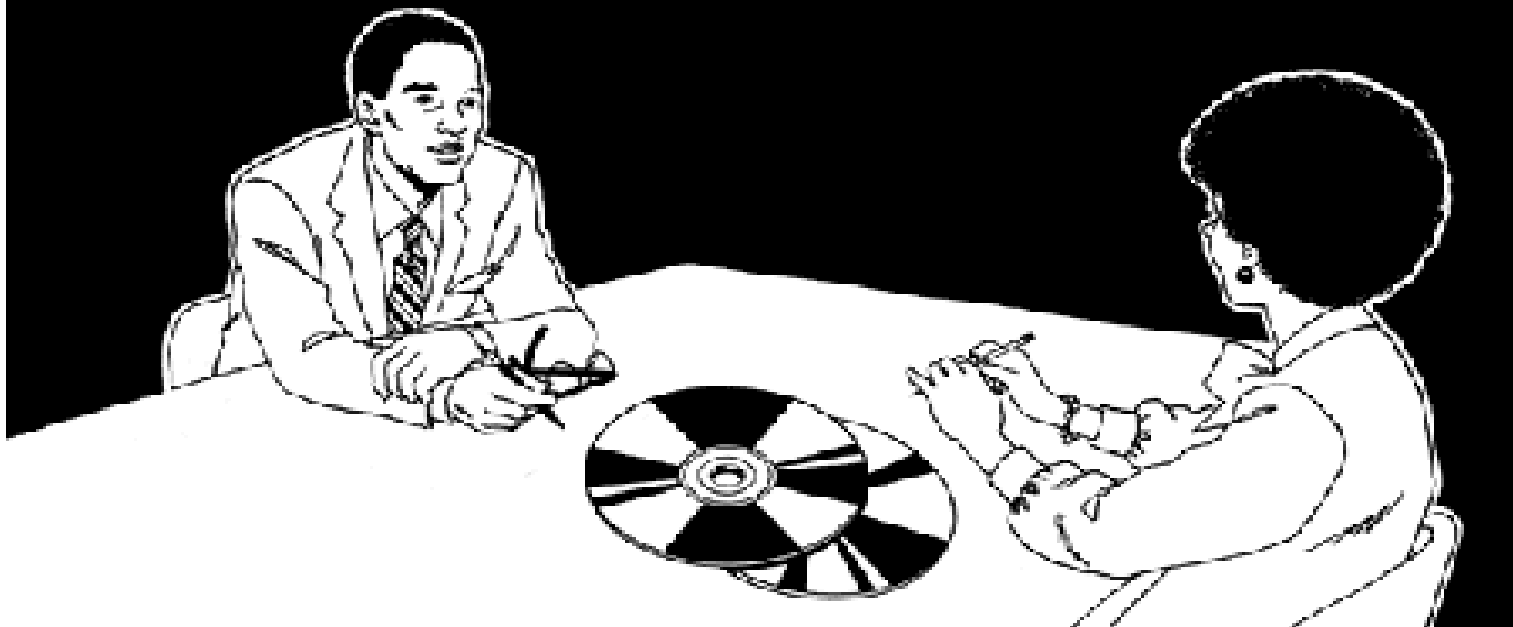
London, 4-5 April 2011

**Steve Knight, Programme Director Preservation Research and Consultancy
National Library of New Zealand**

Today

- Some introductory comments about the digital environment
- How we are looking to leverage the NDHA for the Government Digital Archive and the public record
- A brief look at the status of digital preservation globally
- Why a global, collaborative, interoperable approach matters

These mega-discs are great, Janice, but what we really need is something smaller... Something that can break easily. Make 'em so they rot out after ten or twelve years, too.



PARSE.Insight

June 2010

“The problem with floppy disks and the loss of NASA’s records of the first moon landing are two of the most striking examples of what can happen when digital preservation is not taken care of properly.”

Interview Dutch Ministry of Education, Culture & Science (OCW), 8 Feb 2010

IDC, 2008

In 2007, the amount of digital information created in a year surpassed, for the first time, the amount of storage to deal with it.

Of course we don't need to store all the bits created - like digital TV signals, phone-call routing information, or old email spam.

But if we wanted to, we couldn't.

IDC, 2008

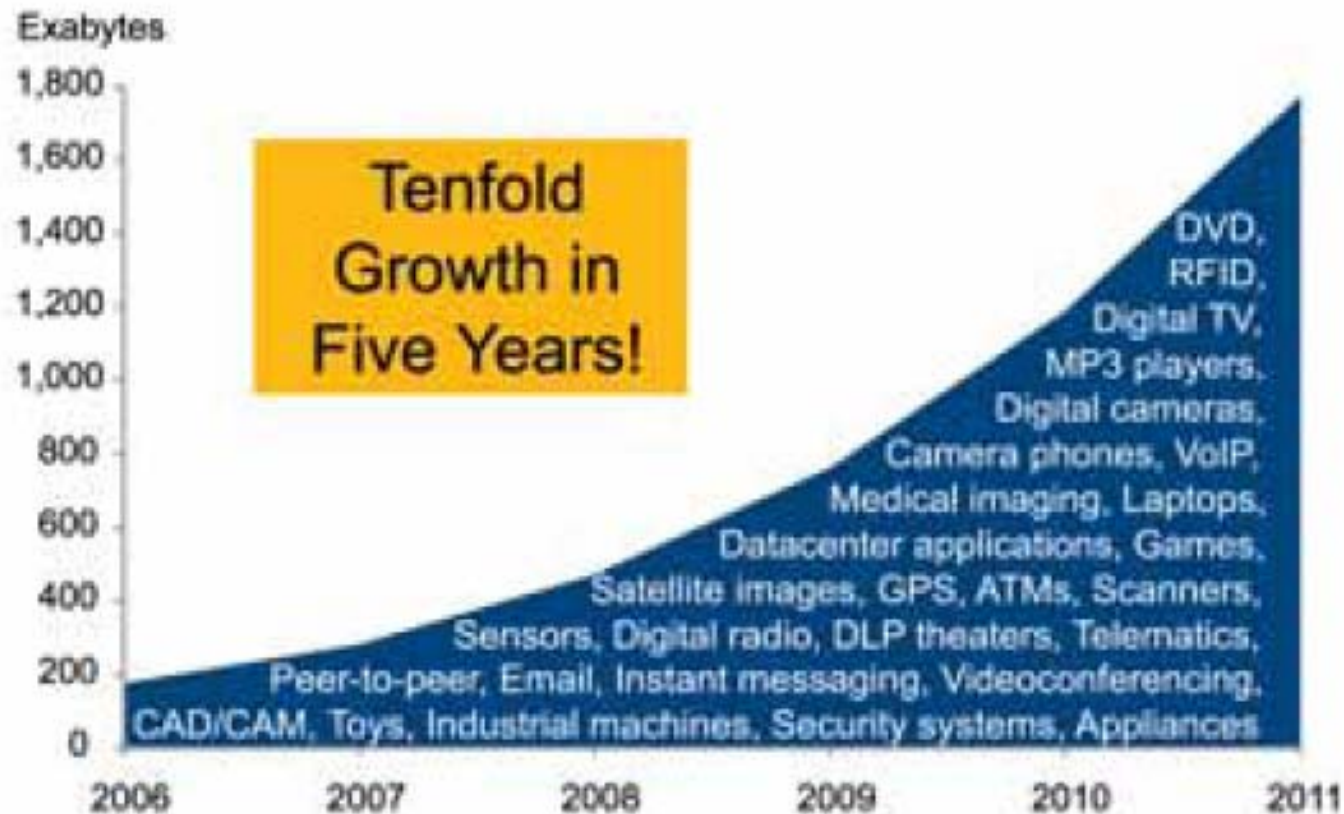
About the only growth rate that hasn't gone negative since the recession began is the creation of new digital information.

People are still taking pictures, making phone calls, sending emails, blogging, and putting up videos on YouTube.

Enterprises are still capturing daily transaction records, adding to their data warehouses.

Governments are still requiring more information be kept and protected, forcing the migration to digital TV, and taking surveillance photos of their citizens.

Digital Information Created, Captured, Replicated Worldwide

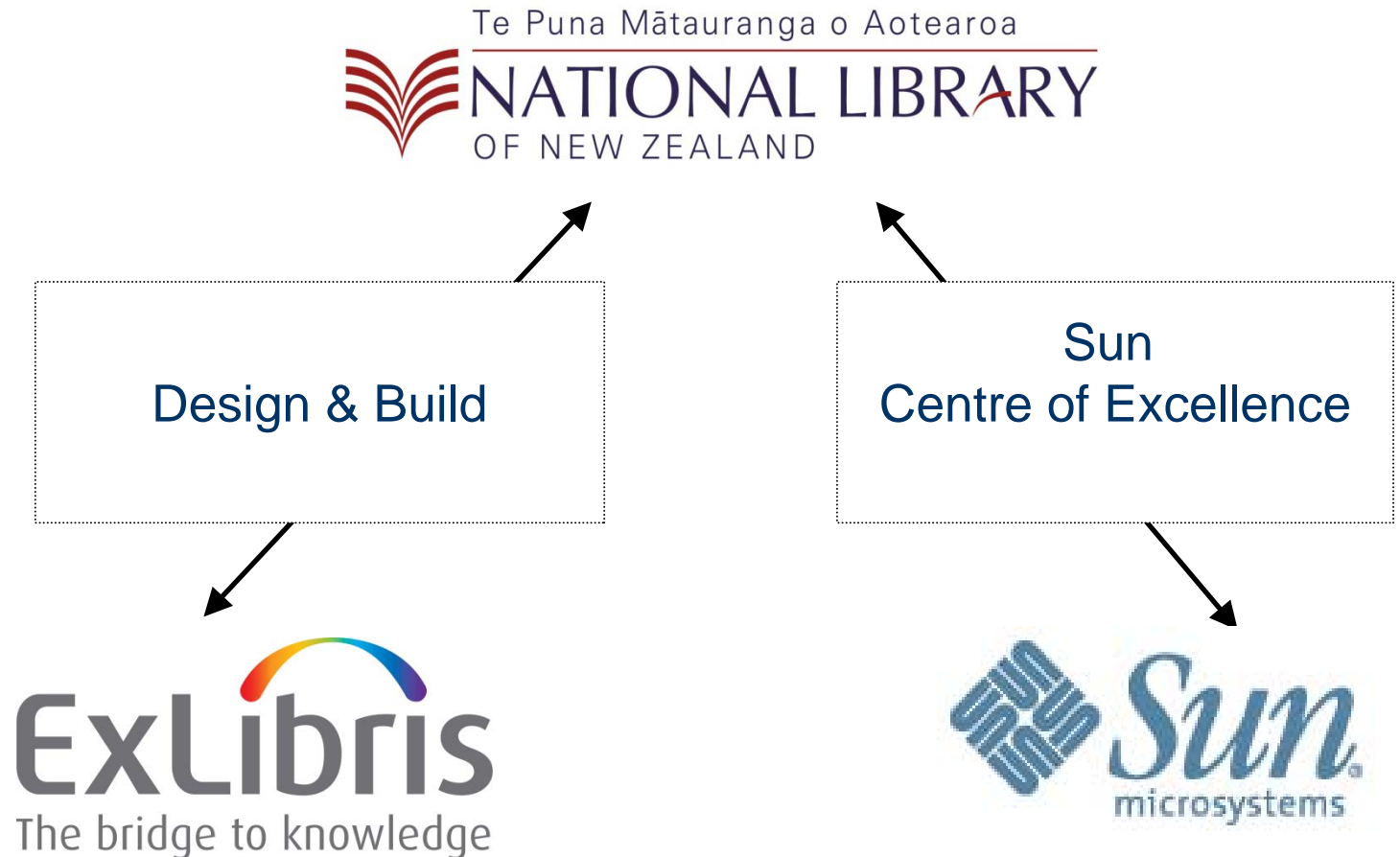


Not the trends part yet

Working with Archives
New Zealand to leverage the
NDHA for the public record

Partnership and collaboration

The NDHA Programme is a successful partnership with our software and hardware vendors resulting in one of the most fully featured digital preservation systems available



Rosetta functionality v2.0

From producer management → workflow automation → delivery, audit trails & reporting → format registry, preservation risk management, planning and action

- User management
- Producer management
- Deposit 1
- Deposit 2
- Validation stack
- Intellectual Entity (IE) data model
- Submission Information Package (SIP) submission
- SIP processing
- Deposit registration
- Technical analyst
- Workbench
- Consolidated appraisal workbench
- Rosetta transformers
- Deposit Application Programme Interface (API)
- Audit & provenance
- Process management
- User management API
- Permanent repository
- Format Registry
- Preservation planning
- Delivery
- Meditor
- Reports
- Back office configuration

The public record

Public Records Act 2005

A government
archiving point of
view

"through the systematic creation and preservation of public archives and local authority archives, to enhance the accessibility of records that are relevant to the historical and cultural heritage of New Zealand and to New Zealanders' sense of their national identity ."

Government Digital Archive (GDA)

A government
archiving point of
view

GDA – leveraging the learnings from the NDHA

A shared vision of how government digital information should be preserved

NZ\$12.6 million – 1 July 2010 to 31 June 2013

Control transfers of archival materials from government departments

Manage re-use of content by the creating department

Provide general access to government archives online

Manage preservation processes

Challenges 1

A government
archiving point of
view

Conceptual – how is content defined across the two organisations and how should it be preserved

Practical – managing the system for NLNZ and ANZ specific requirements as they become clearer

Rosetta enhancements:

- enhanced bulk updating
- enhanced support for consortium management
- ITP in local libraries
- enhanced support for Plug-ins
- enhanced functionality for deposit arrangements
- enhanced functionality for ingest workflows
- enhanced support for delivery rules
- enhanced functionality for logon and identity verification.

Clear benefit to the wider digital preservation community as we work to include increased support for archival practice in the Rosetta system.

Challenges 2

It's not just about the technology

Appraisal of very large transfers from government departments

Providing simple, secure access for departments to their own content

Updating primary collection management systems to incorporate digital preservation workflows

Migration of current corpus to Rosetta – approx 70TB

In a small country, continuing to extrapolate outwards for a national approach to digital preservation – primary research, data sets etc.

Digital preservation as central to a national knowledge infrastructure.

This is the challenge and the opportunity for Archives New Zealand and the National Library of New Zealand.

Trends – or things that maybe
haven't advanced as much as we (I)
might have liked or expected them to

High-Energy Physics' two questions:

"What is the world made of?"
"What holds it together?"

Unique, costly, non-reproducible data!
Pushing energy and precision frontiers

Why should we preserve our data?

EO archives and datasets are invaluable:

- Analysing the state of the Earth, its environment and its variability over time requires a very large number of observations;
- It is impossible to go back in time and resample environmental data, therefore global and complete measurements need to be performed;
- The value of an environmental data set is impossible to estimate and it is impossible to foresee its potential future uses

PARSE.Insight

June 2010

PARSE.Insight

June 2010

Q: When would you invest effort into preservation ?

A: Whenever I have time to do so. This would not be a priority because there would be no recognition for the large effort involved

Digital preservation systems



BL DOM

SPAR_BNF

Repositories ???



Standards, products, tools & services

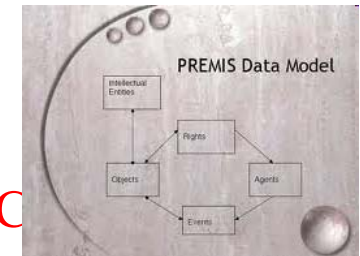


The **technical registry**
PRONOM

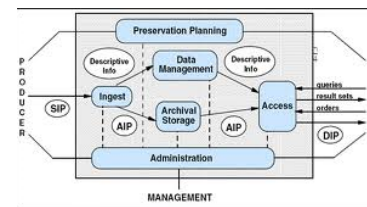
JHOVE2



Trustworthy Repositories Audit & Certification (TRAC)



Storage



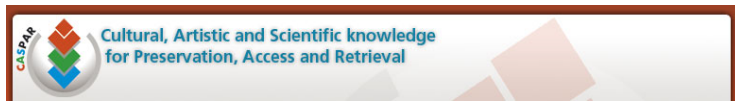
METS-ALTO

METS

NLNZ Metadata
Extract Tool

Oracle/MySQL

Projects, advisory, community



A Tower of Babel or a lingua franca?

What do we mean when we talk about digital preservation?

- Repositories
- Data archiving
- Digital archiving
- Life cycle
- Digital curation
- Data curation
- Digital preservation

- Standards
- Certification/Audit

Is the OAIS model still relevant or do we hold to it instead of developing more granular standards for digital preservation?

We need clarity and certainty about what we mean when we say digital preservation.

Why a global understanding and a collaborative, interoperable approach matters

What isn't under our protection

Every library is exclusionary ...every library conjures up its own dark ghost ...a shadow library of absences ...every library is by definition the result of choice ...and every choice excludes another, the choice not made.

Alberto Manguel 'The Library at Night'

Oh dear

It is not
an accident
that national
libraries,
archives, museums
are called memory
institutions

Destruction of the Library of Alexandria – date uncertain

Destruction of scientific and philosophical library in Cordoba – 10th century

Destruction of the Corvina Library in Buda – 1526

Destruction of the Fatimid Library in Cairo – 1806

Destruction of the libraries and archives of the Maya – Spanish Colonialists

Destruction of the libraries and archives of the Aztecs – Spanish colonialists

Destruction of the National and University Library of Bosnia and Herzegovina – 1992

Oh dear again

It is not
an accident
that national
libraries,
archives, museums
are called memory
institutions

Destruction of the Abkhazian Research Institute of History,
Language and Literature – 1992

Mayor of Orange, France removing material deemed to be not truly
French in support of far right National Front party – 1996

Patriot Act, United States requires libraries to hand over details of
their users – 2001

Burning Harry Potter in New Mexico – 2001

Destruction of National Library of Iraq including books that survived
the sacking by the Mongols in 1258 when the waters of the Tigris
were said to have run black with ink – 2003

Court ordered burning of books in Cuba - 2005

Dove World Outreach Centre, Florida – 2010, in the end refrains
from burning the Koran.

Memory institutions

It is not
an accident
that national
libraries,
archives, museums
are called memory
institutions

What is the impulse behind these acts?

What is it that is feared?

If it is not the memory held within these
institutions?

This is why interoperability matters.

This is why we need more collaboration.

This is why we need a shared understanding of
what digital preservation comprises.

Burning Harry Potter in New Mexico

Book burning is alive and well in the 21st century.



PARSE.Insight

June 2010



The most important reasons for
preservation are the ones we
do not see now

Thank you

Steve.Knight@dia.govt.nz