Progress in digital preservation: an Australian experience

Nitya L. Karmakar

University of Western Sydney, School of Management, Sydney, Australia
Today’s content

- Paradigm Shift in Digital preservation
- The Preservation Challenge
- Australia’s response
- Australasian Digital Record keeping Initiative (ADRI)
- Uniform Australian Approach
- Key issues
- Public Record Office Victoria
- Concluding Remarks
Paradigm Shift in Digital preservation

The information revolution has radically changed the concepts of our work and long term preservation of all records. Corporate information is now shared via email, websites, shared folders and relational databases. Documents are saved in digital forms in several media. The paper based preservation has been replaced by digital preservation.
The preservation challenge

- Need reliable and authentic records accessible for many decades
- Global competitiveness, efficiency and time demands ➔ electronic systems & formats
- Digital solutions ➔ unstable, changing
The known problem

- Rapid obsolescence of
  - data formats
  - digital media, densities etc.
  - hardware
  - software – operating systems & applications
- Physical instability of digital media
Australia’s track record in Digital Preservation

- Library community:
  - NLA – PADI (web resources)
- Spatial data community
  - Geoscience Australia
- Archival community
Alternatives

- Data migration
  - Media deterioration
  - Media obsolescence
  - Software & personnel
  - Time cycles
- Emulation
  - Presentation vs preservation
Convergence to XML?

- International standardisation vs proprietary
- Extensibility = adaptability
- Platform neutral
- Commodity status (?)
Australasian Digital Recordkeeping Initiative

- National Archives of Australia
- Archives New Zealand
- Public Record Office Victoria
- State Records NSW
- ACT Territory Records
- Archives Office of Tasmania
- Northern Territory Archives Service
- Queensland State Archives
- State Records South Australia
- State Records Office Western Australia
ADRI Framework

- Making & managing digital records
- Keeping digital records (both in agencies and in the archives)
- Transferring digital records to archives
- Using digital records
ADRI Framework: Making & managing digital records

Guidelines and tools for:
- creating accurate & reliable record
- classification and control metadata for records
- technology-specific issues for records (eg, email; encryption; web-based resources)
- functional requirements for recordkeeping systems
- model procurement specifications
- compliance assessment and/or self-diagnosis
- schemas for record genres

Standards for
- recordkeeping metadata
- methodologies for recordkeeping system design and implementation
ADRI Framework:
Keeping digital records
(both in agencies and in the archives)

Guidelines on:
■ preferred data formats for long-term preservation
■ XML-based approaches to digital preservation
■ migration paths for long-term digital records
■ preservation treatment of specific data formats (e.g., database-derived records; web-based materials)
■ process models for preservation of digital records in agency custody
■ dealing with changes in the machinery of government over time
■ process models & recommendations for retrieval of records from defunct systems or media

Standards for:
■ describing digital records
■ minimum requirements for preservable data formats
■ trustworthy digital repositories
■ maintenance of provenance and authenticity for digital records
ADRI Framework:
Transferring digital records to archives

Guidelines on:
- Preferred data formats & methods for transferring records to archival custody
- Methods for automatic transfer of recordkeeping metadata
- Maintenance of provenance and authenticity

Standards for:
- Transfer between government agencies and from agencies to archival custody
- Minimum authenticity requirements
- Transfer media
ADRI Framework: Using digital records

Guidelines on:
- Legal provisions affecting access to archives in digital form
- Expunging of sensitive data from public access copies
- Storage & presentation of preserved data formats
- Certification of provenance & authenticity
- Fraud prevention
- Data re-formatting and presentation (eg, databases)
- Searching & discovery mechanisms

Standards for
- Uniform resource discovery based on metadata sets (eg, AGLS)
- Thesaurus-based searching (eg, AGIFT)
Uniform Australasian approach?

- Public standards
  - ISO 14721:2003 (OAIS Blue Book)
  - ISO15489 – Records Management
  - XML rfc
- Data centric (records are data)
- Up-front (continuum)
  - Recordkeeping standards
  - Metadata standards
  - Genre schemas
- Comprehensive ‘whole-of-life’ cover
- Implementation approaches
Uniform Australasian approach
Public standards

- ISO 14721:2003 – Reference model for an open archival information system
Uniform Australasian approach
Public standards

- ISO 14721:2003 – OAIS
Uniform Australasian Approach
Public standards

  - Functional requirements for recordkeeping

- XML v.1.0: Extensible Mark-up Language (W3C – 1998+)
  - Standard Generalised Mark-up Language – ISO8879:1985
Uniform Australasian approach
Implementation-level approaches

- XML package wrapping
- Industry commodity platforms
- Open, documented formats
- Minimise migration requirements
- Provenance & authenticity mechanisms (checksums)
Uniform Australasian approach

What is not uniform?

- Access regimes
- Implementation details
- State of play
Issues to be resolved?

- Authenticity mechanisms
- Storage media (RAID-Redundant Array of Intelligent/Inexpensive Disks)
- Technology migrations
- Capacity planning
- Repository management
- Scalability
How long to keep digital records?

Australian Government agencies have legal responsibilities under the *Privacy Act 1988*, the *Freedom of Information Act 1982*, and the *Archives Act 1983* to ensure the ongoing maintenance and accessibility of their records.
How long to keep digital records? Contd..

Temporary value- the records can be disposed of at an identified time (eg ‘Destroy 7 years after action completed’). Temporary-value records can range in retention length from a very short period, such as one year, up to an extended period, such as ‘Destroy 130 years after date of birth (of subject)’. records.

Retain permanently in agency- the records have a long term business use in the agency, but are not considered to have archival value.
How long to keep digital records?

Contd..

Archival value- the records cannot be disposed of but instead will be retained in the custody of the National Archives indefinitely. Agencies nominate groups of records when they develop a record disposal authority. The National Archives decides which records meet the criteria for archival value.
How long to keep digital records? Contd..

As the official custodian of Commonwealth records (NAA) has a significant role to play in ensuring the preservation and accessibility of digital records well beyond their creation.
Digital Preservation Process

Three Distinct Processes:

**Quarantine** – Records checked for viruses and completeness against list provided by government agencies.

**Preservation** – Records preserved(normalised, or converted, into open formats) using the in-house developed software application, Xena( XML Electronic Normalising Archives),

**Storage**: Records deposited for long-term storage in a digital archive
Archival Preservation

• Investigating what we should do,
• Investigating the proposed solution,
• Metadata,
• Implementing the solution,
• Implications of the solution.
OBJECTIVES

- Don’t adopt preservation solutions which close off future options
- Stop thinking about preservation as an end: it is a means to future access
- Preservation comes at a price - and we shouldn’t be afraid to articulate it
- Influence decision-makers on their terms
Reasons to be cheerful

- Greater understanding of the issues
- Conceptual tools in place
- Cross-governmental responses
- e-government initiatives are an unmissable opportunity
- Heavy duty preservation projects
Reasons to lie awake at night
- Electronic Record Management (ERM)
  - Problem not yet fully internationalized
  - Handcrafted preservation will not do
  - Skill shortages
  - Archives slow to learn how to manage change projects - and to work with others
Where is Australia Heading?

- Joined-up strategy
- Organisation repositioned
- New skills
- Necessary conceptual tools
- Innovative funding
- Effective response to custody debate
Opportunities and Threats in the digital environment

Opportunities
- **Easier Access to records online**
- **Creation of multiple copies of records for back-up and recovery in case of disaster**

Threats
- **More vulnerable than records created in more traditional ways**
- **Lack of Enforcing Policy**
- **Inconsistent cyber law**
VERS, Developed by Public Record Office of Victoria (PROV),

VERS is a world-leading solution to the problem of Capturing, Managing & Preserving electronic records,

VERS is promoted globally and has been accepted and used by archival institutions, national and international Governments and local and global product vendors,

VERS MISSION

“To promote the understanding and effective adoption of VERS across Victorian Government- and beyond, leading to accurate, reliable and authentic records, supporting open and accountable Government and preserving digital heritage.”
VERS Digital Preservation Solution
End to End Process

Govt Agency
- Create Records
- Use / Manage Records
- Export Records

PROV
- Manage Records
- Store Records
- Import Records

Govt Agency / Public
- Search for Records
- View / Copy Records

transfer process
The VERS Program

- 1995-96: Research
- 1998-99: Proof of concept project ($0.5 Million)
- 2000: Published VERS Standard (PROS 99/007)
- 2000-02: Pilot Implementation ($4.8 Million)
- 2002-04: VERS Centre of Excellence ($2.7 Million)
- 2003-05: Digital Archive ($5.5 Million)
- **2004-09: Sustaining VERS adoption ($4.2 Million)**

Total $18.0 Million AUD
The VERS Standard

- Documents and all metadata are stored together in a single file (using XML)
- Defines minimum, maximum metadata for context
- Reduced set of long term preservation formats (Text, PDF & TIFF)
- Is system and application independent
- Is self-describing
- Uses digital signatures to ensure integrity and authenticity
- Can be extended in many different ways
- Is mandatory in Victorian Government
VERS Format

Archivable Record

- Long term format
- Encapsulation
- Self documenting
- Digital signature adds integrity
VERS Centre of Excellence

- Market Development
  - Standard, Vendors, National, International

- Building the Digital Archive
  - Design, Development & Implementation

- Agency Implementation Support
  - Training, Consultancy, Assessments

- Collaboration & Information Delivery
  - Government, Office of the Chief Information Officer, Business, Academia, Research
Electronic records scenario

- Internet
- Intranet
- Content Management
- EDMS
- eMinisterials
- Case Management Systems
- Workflow Systems
- QuickPlace (collaboration)
- RDBMS
- Line of Business Systems
- Email Systems

Information Technology

Information Management

Business Requirements

Office Applications

Shared Drives

PROV

VEOs

ERMS
Native Compliance

Operational

EDMS

Management of Document

Create a VEO

RM/Archiving

TRIM

Management of Record

Public Records

Transfer to PROV
Export Compliance (1)

Operational

Copy

EDMS
Management of Document

RM/Archiving

ERMS
Management of Record

Public Records

Transfer to PROV

VEO
Vendor Participation

- VERS Compliant
  - VERS@DOI, VERS@DOJ, Objective, WinVEO

- Certification Process
  - Hummingbird, Documentum, FileNet, Fuji-Xerox, C digital, Tower Software (TRIM), IBM
VERS Digital Preservation Solution
Services to Government

New government client services

- Streamlined and quicker transfer process
- Safe and secure storage for permanent electronic records
- Quick and easy online retrieval of records

Services available mid 2005
Electronic records from govt agencies

External Inbox
- FTP transfer
- Transfer through firewall

Internal Inbox
- Media transfer
  - Records are collated in Inbox.

Quarantine
- Virus scanning occurs in Quarantine.

VERS Compliance Support Area
- Non VEOs / exceptions
  - Manipulation of non-VEOs or exceptions.
  - VEOs
    - Records undergo archival and VERS validation.

Processing
- VEOs
  - Records are moved into digital repository for permanent storage.

VERS Digital Preservation Solution
Transfer Process
VERS Digital Preservation Solution
Public Services

New public client services
- Integration of archives@victoria into PROV website
- A single point of access to entire PROV collection
- Better searching capabilities
- Ability to save search results
- Online and FREE access to digital archives

- Services available mid 2005
VERS Digital Preservation Solution
Access – Current Homepage
VERS Digital Preservation Solution
Access – Current Entry Point

Public Record Office Victoria provides a wide range of services to enable researchers to access the public records of Victoria. The following pages will help you search for, order and use records from our collection. While many initial research tasks can now be done over the Internet, you will need to visit a Public Record Office Victoria reading room to view most of the records we hold.

WHERE HAVE THE IMMIGRATION INDEXES GONE? Look under our ‘Online Resources’ page to access these and a list of other online material now available.

ACCESS TO THE COLLECTION LINKS

- Resources
  - Archives@Victoria: Search for and order records online
  - PROCustody - Research: Detailed information on searching our collection.
  - Online Resources: A list of datasets and digitised records to help with research.

- How to...
  - Begin your Research
  - Visit Us
  - Online Records
  - Copy Records
  - Reference Research
  - PROCustody - Services

As well as the topics listed here you can also find details of various EVENTS AND PROGRAMS. Public Record Office Victoria runs to help researchers find the information they are after. Our ONLINE EXHIBITIONS show some of the stories to be found in the records.
VERS Digital Preservation Solution
Access – New Entry Point

Public Record Office Victoria provides a wide range of services to enable researchers to access the public records of Victoria. The following pages will help you search, order and use records from our collection. While many initial research tasks can now be done over the internet, you will still need to visit a Public Record Office Victoria reading room to view most of the records we hold.

It is best to plan your research before you start searching for records. If you already know the Agency or Series you are looking for, do a search and consult the Details page for the Agency or Series. If you’re not sure of the Agency or Series, you might find it helpful to start with Research Categories to narrow your search to include Agencies with records relevant to your particular research topic.

Research Categories
- Family History
- Arts & Recreation
- Education & Law
- Science & Environment
- Transport
- Whole of Victoria Government

Perform a search
- Simple
- Find by Number
- Advanced
- Search within a Series

Page last reviewed: 20 Jul 2004
© State of Victoria 2004 | Disclaimer | Privacy Statement
VERS Digital Preservation Solution
Access – New Search Results
Strategic Benefits

- In line with OCIO strategic initiatives
- Interoperability
- Uses mainstream preservation approaches
- Scalable
- Vendor participation supports a sustainable market
Lessons Learnt

- Vision, strategy, program planning and delivery = continuing $’s
- It is part of a long term reform agenda
- Narrow focus on particular record sets or customers provides best outcomes
- Change management is a substantial problem at many levels
- For longevity a sustainable market is required
- Volumes of records, processes, applications and projects will require changes to established processes
Future Goals

- Practical Route Map
- Increased vendor support
- Targeted digital transfers
- Digital Archive in production (May 2005)
- Support of the Australasian Digital Recordkeeping Initiative
- Annual departmental assessments
VERS

Just do it!

Concluding Remarks

- Keep digital records for the long term
- Capture records as close to creation as possible
- Use a single long term format
- To develop scaleable practical solutions
- Build a sustainable market for products and services
- Be standards based
- To develop scaleable practical solutions
Concluding Remarks- Contd..

- Digital records requiring long-term maintenance need to be actively managed through a planned, systematic and documented approach.

- Managing and Preserving relevant digital records before they are carelessly or unthinkingly deleted.

- Positive note: willingness to collaborate across and between governments in Australia and New Zealand.
QUESTIONS??

Have a G'Day
Have a G'Day
"The paradox of preservation is that it is impossible to keep things the same forever. To conserve, preserve, or restore is also to alter. Even if an object survives untouched, it will have changed just by virtue of aging or by a change in its surroundings." ~ Michele Cloonan

"Preservation is the art of managing risk to the intellectual and physical heritage of a community and all members of that community have a stake in it. ...Preservation becomes an ever-changing assessment of value and endangerment." ~ Abby Smith