Digital Curation – Digital Archiving

Adrian Cunningham
Key Message no. 1

Digital archiving is one form of digital curation
- the two are not synonymous
Key Message no. 2

Digital archives are different from digital libraries
Key Message no. 3

Digital archiving requires active archival intervention across the entire records continuum
Improving Recordkeeping in the Australian Government

• 1990s - NAA repositioned itself as a proactive enabler of good recordkeeping in agencies, providing standards/guidelines

• Response to deteriorating recordkeeping practices and the digital challenge

• Conceptual framework based on AS 4390 (1996)
  • Functional approach to recordkeeping
  • “DIRKS” methodology
Improving Recordkeeping in the Australian Government (2)

• 1998 – NAA/Monash Uni training course
• 2000 – Release of e-permanence suite of standards and guidelines on NAA website
• 2002+ - Audits of government recordkeeping by Auditor-General
• 2007 – “Check-up” – Recordkeeping health check self-diagnostic toolkit
• 2007 – Australian Compliance Standard for ISO 15489
NAA Digital Preservation Strategy

• Project commenced, 2001
• Green paper, 2002
• Uses open source XML technology to wrap and store ‘normalised’ digital records created in proprietary data formats
• Recommended standard archival data formats for use with digital records that are required for long-term retention/access
• Xena version 1.0 tools for normalising digital records, launched September 2004 - now up to Xena 3.0
• NAA digital repository/laboratory, 2005
Migration Strategies – just say no

- Use of File Format Registries (eg PRONOM)
- Has to be repeated every 3-5 years
- Expensive to maintain over the long term
- Every migration is ‘lossy’
- Perpetuates a reliance on proprietary software
Digital Records are Different

• Digital Records are *mediated* by technology

• Preserving the object is meaningless

Need to focus on:
- Interaction between data and technology that interprets the data
- Determining what is archivally important about that interaction and its context (the *essence* of the record)
- Creating and maintaining ability to repeat essence on demand in a sustainable manner
Digital Records are ‘Performances in Context’

Researcher experiences the record via a performance

Source → Process → Performance → Researcher

Preserve the performance and the contextual metadata and you have preserved the record
The NAA and ODF

• NAA joined OASIS Technical Committee in 2003

• A large percentage of government records are created using office applications (DOC, XLS, PPT)

• Xena OpenOffice plug-in uses a Java application with OO2.0 to produce normalised objects in ODF

• Over 30 other file formats can be normalised using other Xena plug-ins (eg MS Outlook, Project, HTML)
NAA Digital Preservation Workflow

• Transfer, 4 Weeks Quarantine, and Ingest
• Normalisation of archival masters
• XML wrapping and retention of original bitstream
• Authenticity mechanisms – checksums
• Secure offline storage of preservation masters
• Access copies stored in separate access repository
Wider Application of Xena

• Anyone with Java programming skills can use or extend/enhance the Xena source code, because it is open source
• The NAA has developed a “Xena lite” for use by Australian Government agencies to preserve non-archival digital records that need to be kept long term
• The NAA wants others to experiment with Xena
• A number of digital preservation initiatives are now basing their approach on open standards.
Total Digital Archiving

• Digital preservation is just one component of end-to-end digital archiving

• MADIRA Project – Managing Digital Records for Access

• Australasian Digital Recordkeeping Initiative – http://www.adri.gov.au
Final Thoughts

• Digital archiving capabilities are in short supply
• Just throwing software at the problem solves nothing
  • You have to know what records you need and how you are going to capture and manage them over time
• It's all about change management
• Records systems are social systems – encourage peer and management pressure by making it easy for people to do the right thing
• Need for industrial scale implementations, not just research projects
HAVE YOU BEEN KEEPING RECORDS?...

ER... SURE...

...THEY'RE ALL ON THE COMPUTER!