Current Digital Preservation Trends and SDB4 Features

Pauline Sinclair, Tessella, PASIG, Madrid, 5 July 2010
Overview

• Introduction to SDB4
• Planets Market Research
• How SDB4 addresses the trends
• Summary
SDB4 - Overview

Based on the OAIS reference model
Evolving technology since 2003
In use at leading institutions

Includes world-leading “Active Preservation” module (safeguard against technology obsolescence)
World Leading Digital Archiving Solutions

The National Archives
Wellcome Trust
Finnish National Archives
Estonian National Archives
Rotterdam City Archive
Dutch National Archives
Malaysian Archives
Swiss Federal Archives
FEDERAL CHANCELLERY - AUSTRIA
Austrian Government

US National Archives and Records Administration

Tessella
Technology & Consulting
SDB4 – Key Features

• Active Preservation
  • Technical Registry - PCR2.1 (based on PRONOM)
  • Characterisation & migration tools wrapped for use
    - DROID, JHove, Stellent, ImageMagick…

• Flexible
  • Drools workflow – custom workflows & steps (API published in SDK)
  • Storage adaptors – network-attached, HSM, in the cloud…
  • Flexible reporting – Jasper Reports
  • Can use any descriptive metadata schema (e.g. Dublin Core, ISAD(G), MARC)
  • Can be configured by 3rd parties – Tieto – Baltic states
Planets Market Research - 2009

• Market Survey
  • On-line questionnaire
  • 200+ participants
    - 67% European, 16% USA & Canada
    - 41% libraries, 30% archives

• Supplier & Vendor Briefings
  • In-depth interviews
  • 18 participants
    - IT companies across digital preservation market (small to multinationals)
    - Headquarters in Europe, North America, Middle East, Australasia
Planets’ Findings: General Digital Preservation

• Confusion over digital preservation - passive vs. active
• Awareness is high, but engagement in digital preservation much lower
• Memory institutions leading the way
• Legislation is the main driver
• Regarded as a luxury
  • Funding is limited
  • Short-term budgets
  • Need for digital preservation cost models
• Policy is a vital first step towards tackling digital preservation
Planets’ Findings: Solution Requirements

• Digital preservation for SMEs
• Scalability
• Demand for digital preservation tools
• Standards
• Emulation
Provision for SMEs

• Related trends
  • Large investment currently required to engage in digital preservation
  • SMEs where archiving is not core activity want to outsource digital preservation

• How SDB can help
  • Multi-tenancy
    - Independent archives
    - Share server configuration & facilities
    - Reduces costs
    - Can be run as a service
Scalability

• Related trends
  • 25-fold increase in material
    - 2009 - < 20 TB
    - 2019 - > 500 TB
  • Scalability is important for:
    - Large volumes of content
    - High ingest rates

• How SDB can help
  • SDB is scalable
    - SDB4 can ingest >1.5 TB / day on a single, entry-level server
    - Ingest is network & I/O bound, not CPU usage
    - SDB4’s storage is limited only by the 3rd party storage system(s) used
Demand for Digital Preservation Tools

• Related trends
  • Demand for
    - Preservation Planning tools
    - Characterisation tools
    - Migration tools

• How SDB can help
  • Active Preservation Framework
  • SDB already makes use of several characterisation & migration tools
  • SDB User Group reporting back on how tools perform in practice & what best practice should be
Standards

• Related trends
  • Compliance with standards important, but too many standards
    - Archiving: OAIS, OAI-PMH, OAI-ORE, TRAC…
    - Metadata: Dublin Core, ISAD(G), PREMIS, METS…
    - ERMS: ISO-15489, MoReq2, DoD5015.3…
    - Domain-specific

• How SDB can help
  • OAIS-compliant
  • Can use any descriptive metadata schema
  • Uses own schema (XIP) internally to allow full Active Preservation
  • Can export to PREMIS (characterisation), & METS (packaging)
Emulation

• Related trends
  • Lack of interest in emulation
  • More complex objects need preservation in future
    - 2009: > 80% organisations need to preserve documents & images
    - 2019: >70% org.s need to preserve video, audio, databases, websites & email

• How Tessella can help
  • KEEP – Keeping Emulation Environments Portable – EU project
  • Plan to incorporate demonstration of emulation in SDB4
Summary

• Trends

• Digital preservation being more widely recognised, but more work needed to make the case for engaging in it & reduce the barriers.

• Trends for digital preservation solutions
  - Provision for SMEs
  - Scalability
  - Digital preservation tool provision
  - Standards
  - Emulation

• Tessella responding to trends and incorporating features in SDB4 to address them
References

• Planets
  • Website: http://www.planets-project.eu/
  • Market Survey:
    - Overview white paper:
    - In-depth report:
  • Supplier/Vendor Briefing – to be published

• Tessella - http://www.tessella.com/
• SDB4 - http://www.digital-preservation.com/
• KEEP - http://www.keep-project.eu/