Fending Off the Digital Dark Age – Assuring Future Digital Access and Preservation

Don Post
IMERGE Consulting, SNIA LTR
SavingTheDigitalWorld.org
PASIG Austin January 11, 2012
SavingTheDigitalWorld.org Goals

www.LTDPRM.org

- To a strong advocate in North America for promoting long-term access and preservation of digital information
- To facilitate interchange of standards and best practices across participating organizations in different industries
- To encourage industry wide development of standards, best practices and knowledge necessary and sufficient to attain acceptable levels of persistent usability
- Scope: North America focus, but applicable worldwide
  - Europe is further ahead, e.g. funding and EU projects
  - US and Canada is a natural community of interest.
Conclusions About Urgency LTDP

- At the current rate, it is possible that in 20 years only 50% of critical info will be protected in TDRs
  - Much of the remainder will be lost forever
  - Many times worst than the 1973 disastrous fire at the NPRC destroyed 16-18 million files (OMPF)

- Now, it is virtually impossible to access knowledge from fragmented and specialized organizations dealing with LTDP. Today, learning from others is limited.

- There is a strong need for a comprehensive collaboration and links between fragmented organizations, knowledge and achievements
% of “Permanent” eRecords in 2010 that Are Still Accessible in Future Years:

Digital information is in jeopardy: it is likely that in 35 years that 50% of critical info will be unavailable.

Reasons:
• Inadequate EDRMS
• Policy & process issues
• Technology Issues
ITEM = Data, Document, or Record

Records “Declaration”
→ Retention, Disposition Rules

EDRMS PROCESSES
AIIM, ARMA, NAGARA

DIGITAL CURATION PROCESSES
SAA, ACA, NAGARA

STORAGE
SNIA, IEEE

Governments
Corporations

Conventional Storage Processes
<10 yrs

Libraries
Archives
Organizations with no EDRMS

TDR Functionality
to support 100 year access, preservation & usability

Long Term Retention TWG
AUTHORITATIVE KNOWLEDGE CENTER

- AUTHORITATIVE
  - Power (n.) Influence Clout
  - Confidence (n.) Conviction Certainty Experience
  - Citation (n.) Source Evidence

- KNOWLEDGE
  - Information (n.) Information Facts
  - Comprehension Realization
  - Expertise Skill Know-How
  - Wisdom (n.) Education Intelligence

- CENTER
  - Consensus (n.)
  - Focus (n.) Concentrate
  - Place (v.) Arrange, Align
Welcome to the Long-Term Digital Retention and Preservation Reference Model

Help Define the Long-Term Digital Retention and Preservation Reference Model

While OAIS exists as the ISO standard of a reference model for a digital archive, it doesn't map adequately to the needs of large-scale distributed organizations operating on the internet or in the cloud and faced with the added cost burdens of legal and security risk, litigation, compliance and business requirements.

New capabilities and broader architectures are needed. Many organizations are extending OAIS, and it is time to begin gathering and defining best practices and common architectures. A simple example is all the noise around “Cloud Archiving.” We are leading the discussion, planning, using case definition, and best practices definition on the utilization of the Cloud for digital Archives. (Meaning preservation stores, not tiering.) 32% of the respondents in SNIA’s new survey declared that they plan on using the cloud for long-term retention periods more than 10 years. Sorry folks, traditional IT practices don’t support this. Significant portions of that data will be lost in 10 years. Is that acceptable?

Join LTDP and help define these important areas.

The LTDP reference model community is being organized to respond to these many needs as digital preservation becomes an IT datacenter practice. Come join in and contribute. The time for development of cost effective, interoperable digital archives in the business and organization datacenter has arrived.

Go to LTDP to Comment and Contribute... Register Here.
Best Source for Reference Material

www.LTDPRM.org

Google Site Wiki

Register Free

Contribute & Comment
CIRCUMSTANCES LIKELY TO ALLOW A DIGITAL DARK AGE

- Inadequate control and classification of items requiring preservation
- Reduced funding in government and other sectors
- Expectations of digital usability, compared to paper or microfilm is higher
- Early stages of TDR requirements and specs evolution.
- Virtually non-existent vendor and integration firm supply chain
- Complexity of considerations in architecting and sustaining solutions.
APPROACHES AND STEPS TO HELP AVOID A DIGITAL DARK AGE

- Gain broad consensus of approaches and steps
- Strongly advocate for actions which largely are not being adopted today
- Take steps to radically improve availability and low costs of both EDRMS and TDR solutions.
- Gain rigorous credentialing of software, systems, processes and people
Records Management as an Analogy to LTDP-TDR Growth

- eRecords Management solutions history is a useful parallel to LTDP-TDR evolution, as described below:

- 1989 Provenance Systems (Ottawa Canada) launched ForeMost Software.

- In 1997 ForeMost world’s first certification against the US DoD 5015.2 standard.

- In 1999, Tarian Software was founded first e-Records software engine for business software.

- In 1999, there were essentially three ERM (i.e. electronic record keeping systems) commercial software vendors – Provenance, PSSoftware and TRIM.

TDR ANALOGY: Today there arguably are only three commercial software/solution vendors of TDRs.
Records Management as an Analogy to LTDP-TDR Growth

- In 1999, eRecords Management solutions was standalone software having no interoperability with well-established ECM document versioning systems like FileNet, Open Text, Documentum. ERM was of interest only to records managers.

  TDR ANALOGY: In 2012, OAIS-TDRs are architected as standalone solutions of interest only to archivists and digital curators.

- In late 1999, Open Text acquired PS Software (5015.2 certified)

- At AIIM 2000 spring of 2000, Don Post did a session at AIIM on 5015.2.

  TDR ANALOGY: In 2011, there are emerging sessions on OAIS-TDRs at SAA, SNIA, PASIG, but not AIIM, ARMA.

- In October 2001 the Enron scandal eventually led to the bankruptcy of the Enron and the dissolution of Arthur Andersen. At this time records management became a Board/CEO level issue.

  TDR ANALOGY: NONE. Unclear what will be the tipping point for general consciousness about long-term risk, preservation and access issues.
Records Management as an Analogy to LTDP-TDR Growth

- By June 2002, DoD 5015.2-STD “Design Criteria Standard for Electronic Records Management Software Applications” was implemented. This standard defines requirements for the management of records in DoD.
  - Standard defines requirements for the management of records in DoD.
  - Became in the past decade (2002-2010) become the accepted standard for many governments and corporations worldwide.
  - All serious providers of EDRMS (electronic document and records management systems) obtain DoD certification. While other ISO and other standards/practices have evolved, DoD 5015 is the only testable certification program of wide recognition.

  **TDR ANALOGY:** A DoD software certification of TDRs would accelerate and validate “requirements” across users and software/solution suppliers.

- In early 21st Century, eRecords Management became essential EDRMS.
  - In November 2002 Tarian acquired by IBM.
  - In 2002 TruArc (was Provenance ForeMost) acquired by EMC/Documentum.
  - About the same time, TRIM was acquired by HP.
  - Other records management products were introduced as part of suites.

  **TDR ANALOGY:** LTDP-TDR will become an essential extension of and interoperable with EDRMS Suites, products and processes.
Records Management as an Analogy to LTDP-TDR Growth

- I has taken eRecords Management over twenty years (1989) to reach the current level of maturity after the first well designed software was introduced.

- There were unique changes which have not occurred yet for LTDP-TDR
  - And 13 years after being certified by DoD
  - Ten Years after bankruptcies, threat of executive jailing
  - Several years after eDiscovery became widespread.

- TDR ANALOGY: LTDP-TDR possibly take twice as long – 40 years – to reach a similar level of Maturity in absence of “tipping point” of political, legal and governance impact.
An Ultimate Metric of the Success of the ECM & Records Program:

Set goals and compare to estimated % of electronic content and paper records that are:

- Formally versioned, retained and disposed of through EDRMS implemented policies.
- Total content under control by the end of 2014 might still be insignificant? By 2021?
- Limitations to attaining a high percentage under control need to be addressed in the Strategy and Roadmap from this project.

The next slide shows what future metrics might look like.
ECM & Records Program Content % Controlled
Versioning, Retention, Disposition?

??

Graph showing the percentage of content controlled from 2010 to 2016.