



Valued Assets to the  
Business Process



# *100 YEAR ARCHIVE TASK FORCE*

## Overview

September, 2006



# The Problems Of Long Term Digital Archive

- ❖ Applications – aging, compatibility, obsolescence of old formats
- ❖ Computer Systems & infrastructure,
- ❖ Media life & readability
- ❖ On-media Formats
- ❖ File systems, operating systems
- ❖ Search and discovery capability
- ❖ Cost to maintain
- ❖ Migration complexity & cost
- ❖ Metadata & indexes
- ❖ Lack of forward and backward interoperability & compatibility
- ❖ Damage
- ❖ Perceived value of the information
- ❖ Human intervention requirements
- ❖ Lack of standard readers

*Long term preservation of digital content is a big challenge in the Information Society era, digital information in any form is at risk to be lost forever. Technology on which digital content relies becomes obsolete and application versions and files formats change, making data soon inaccessible. Even if content is coded in the simplest format, such as ASCII code, storage media degradation and obsolescence could make it disappear. Even on-line information such as web pages and databases, are vulnerable as much as their web structure become complex thanks to (aging) hyperlinks and cross references.*

Source: Alfredo M. Ronchi - Medici Framework

*The challenges of maintaining digital archives over long periods of time are as much economic, social, and institutional as technological.*

# 100 YEAR ARCHIVE TASK FORCE

## Objectives

- Produce with a multidisciplinary team a “best practices for long term digital archive” paper similar to the Sedona project
- Guide ILM as a core management practice for archive and accommodate the impact of new storage technologies (XAM, Grid)
- Define a standard for on-media formats & long term readability

## THE SEDONA GUIDELINES: *Best Practice Guidelines & Commentary for Managing Information & Records in the Electronic Age*

A Project of The Sedona Conference<sup>SM</sup>  
Working Group on  
Best Practices for Electronic Document  
Retention & Production

September 2004 Public Comment Draft



Reference Model for an  
Open Archival Information  
System (OAIS)

(ISO 14721:2002)

# WHY LAUNCH THIS EFFORT?

- ❖ Existing Archive practices address the process & systems problem and merely say – *keep up with storage technology*
- ❖ Enterprise I.T., RIM, and Security professionals need guidance under the new pressures of compliance and risk management as responsibilities converge
- ❖ New technologies offer new approaches to the long term archive problem
  - ILM as a management practice
  - CAOS systems
  - Grid-based Storage & Data Grids
- ❖ Connect & coordinate with other agencies and organizations working with archive standards

# 100 YEAR ARCHIVE TASK FORCE

## ✓ Operating Plan

- Formation meeting June 22
- Conduct business via weekly telcons and a public SNIA operating portal
  - <http://advisorygroups.snia.org/apps/org/workgroup/100yr/archive/>
- Produce papers & presentations as the group defines
- Coordinate with SNIA TWGs and other outside agencies as needed

## ✓ Initial Member Orgs

- SNIA
- ARMA
- AIIM
- Individuals – US Army, LDS Church, MIT, ...

## ✓ Pending Member Orgs

- NARA, SMPTE
- Many others

# Program Plan

- ✓ Launch – June 22
  - ✓ Deliver program plan – Sept 1
    - Refine objectives
    - Background research
    - Establish relationships
    - Launch sub-committees
  - ✓ Interface with and guide SNIA TWGs - Ongoing
    - SDDF, ILM, CAOS
  - ✓ Present at upcoming conferences -- EIW 2005
- ✓ Deliverables
    - Definition of best practices delivered in presentation and white paper format, jointly published by the associated parties and Associations
    - Definition of requirements and validation of directions for the SDDF, ILM, & CAOS technical work groups and the standards they are working on
    - Speaker's bureau - this team will have access and content to talk about at our mutual events and other industry programs
    - Other deliverables the team identifies

# Important Parallel Activity

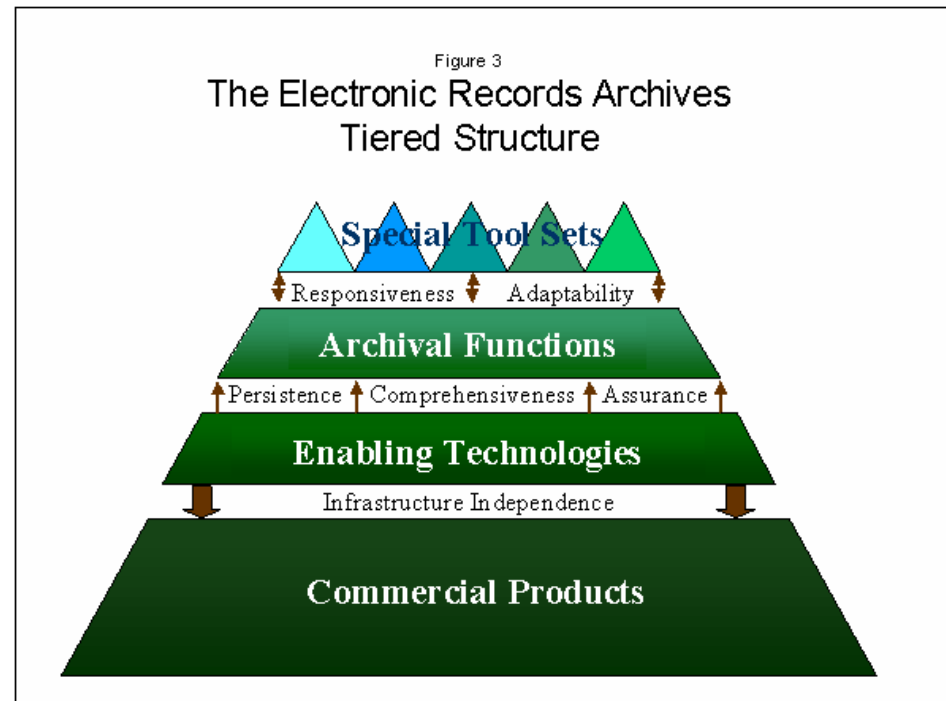
## ✓ **NARA** – Electronic Records Archive

"Mark my words, ERA is going to change the world as we know it."

- John Carlin, Archivist of the United States

## ✓ **ISO** – PDF/A (TC/ 171)

## ✓ **UK Data Archive** – Data Exchange Project



## RECOMMENDATIONS FROM ARCHIVISTS

- ❖ The digital archiving community would benefit from shared services such as a format registry with full documentation of all versions of all open and proprietary formats along with tools for automatic format conversion. A metadata schema registry (or registries) is also needed. - Research Challenges in Digital Archiving and Long-term Preservation, Margaret Hedstrom, University of Michigan
- ❖ The complexity of any single system makes quite difficult to gauge its ability to preserve data over the long term. Traditional data transfer (migration) on removable media is very costly in the aggregate. - Digital Archiving and Long Term Preservation: THIC



# OPERATING PORTAL

- ✓ SNIA AdvisoryGroups area, to signup, go to
  - [http://advisorygroups.snia.org/kmembership\\_info/person\\_signup/](http://advisorygroups.snia.org/kmembership_info/person_signup/)
- ✓ Then with your password, go to this link and 'join the group'
  - <http://advisorygroups.snia.org/apps/org/workgroup/100yrarchive/>
- ✓ Telcons are weekly beginning Thursday 6-30 at 3 PM-ET
  - 323-785-4156 passcode 969-1665#
- ✓ Group Email: [100yrarchive@advisorygroups.snia.org](mailto:100yrarchive@advisorygroups.snia.org)

# OPERATING PORTAL

- ✓ SNIA AdvisoryGroups area, to signup, go to
  - [http://advisorygroups.snia.org/kmembership\\_info/person\\_signup/](http://advisorygroups.snia.org/kmembership_info/person_signup/)
- ✓ Then with your password, go to this link and 'join the group'
  - <http://advisorygroups.snia.org/apps/org/workgroup/100yrarchive/>
- ✓ Telcons are bi-monthly -Thursday 3 ET
  - 2<sup>nd</sup> & 4<sup>th</sup> Thursdays
  - (712)580-0600 code 961226#
- ✓ Group Email: [100yrarchive@advisorygroups.snia.org](mailto:100yrarchive@advisorygroups.snia.org)