

Data and Record Archiving Best Practices to Assure Continuing Regulatory Compliance and Protection of Intellectual Property

IMACS

Princeton, NJ

May 21, 2008

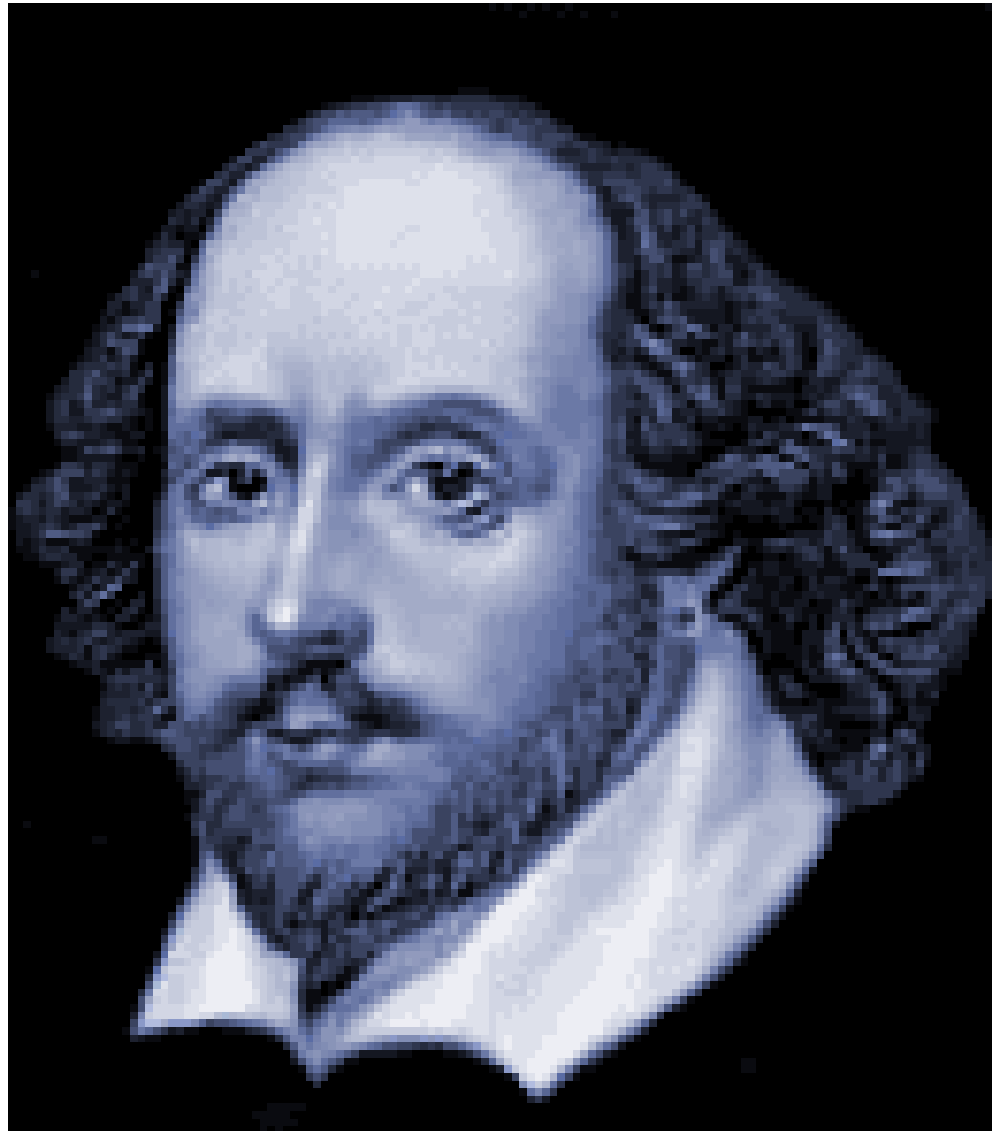
Charlie Sodano, Ph.D.

Senior Consultant

charlie.sodano@eorganizedworld.com



Martinez, California





Review of Records Concepts

Ø What are Records?

Ø What is the Purpose of Records?

Ø Managing the Lifecycle of Your Records



PURPOSE OF RECORDS IN AN ORGANIZATION

- Ø Meet current operational needs**
- Ø Provide legal and regulatory accountability**
- Ø Provide a memory to build on**



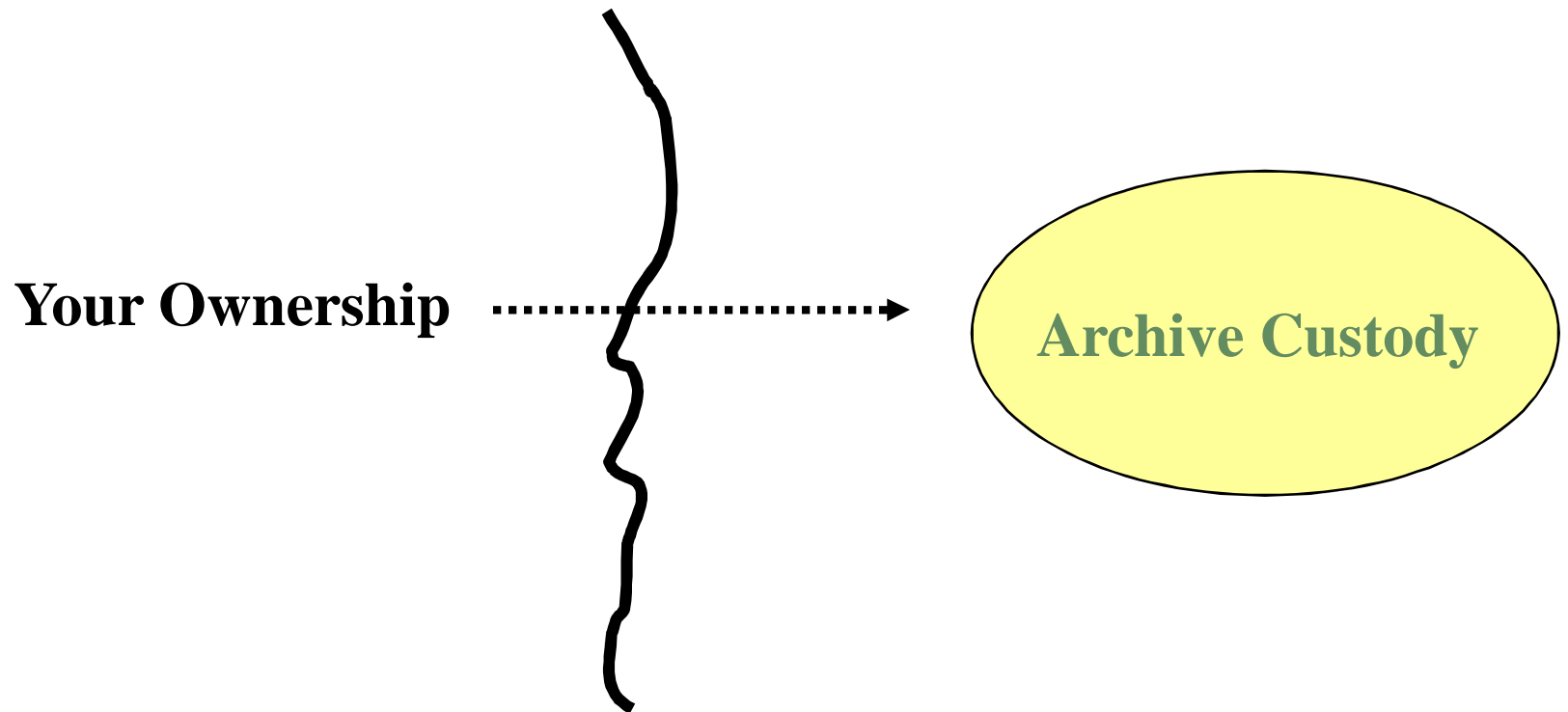
WHAT'S A RECORD?

Documents created, received, and maintained as evidence and information by an agency, organization, or person, in pursuance of legal obligations or in the transaction of business - ISO

Recorded information in any form or medium, created or received and maintained, by an organization or person in the transaction of business or the conduct of affairs - ICA

Review of Records Concepts

WHEN DOES A DOCUMENT BECOME A RECORD?

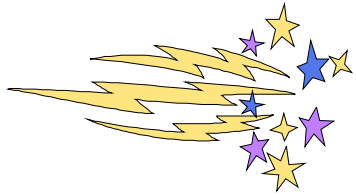


Beginning a Records System

- Ø Identify documents that should be records
- Ø Develop record series (filing plan)
- Ø Develop retention schedules
- Ø Documents that are not records will not be managed

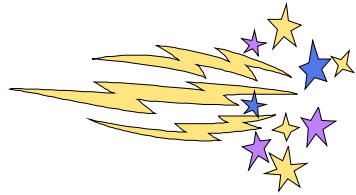
Retention Schedules

- Ø **Federal law requirements**
- Ø **State/local law requirements**
- Ø **Business needs**



RECORD MUST HAVES

- Ø Documented procedures for identification, collection, indexing, access, filing, storage, maintenance and disposition
- Ø Retention schedules
- Ø Disaster recovery



RECORD MUST HAVES

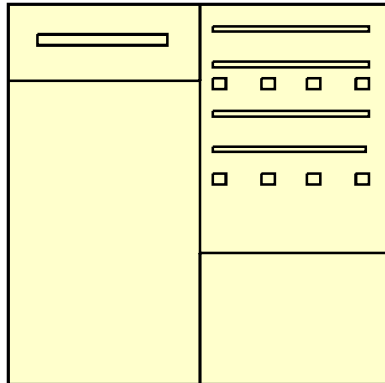
Authentication – Proof that the document is what it purports to be, and that there is a relevant relationship between the document and a particular person, who is frequently a party to the litigation.



E-Records and Databases

E-Record

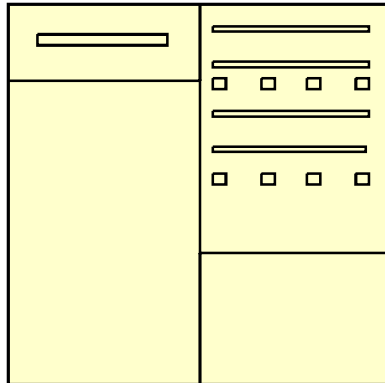
- Ø Like a paper document
- Ø Snapshot in time
- Ø Can be generated from a validated database
- Ø Can cut ties with database after creation
- Ø New audit trail



E-Records and Databases

Databases

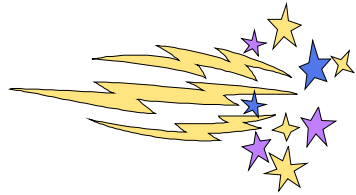
- Ø **Dynamic**
- Ø **Controlled (validated)**
- Ø **Generate reports/views complying with user requirements**
- Ø **Maintain audit trail**
- Ø **Not a set of records, but an authenticated view**



E-Records and Databases

Databases

- Ø Must be kept validated to produce authentic views
- Ø Can be decommissioned or migrated



Record Preservation

- Ø Capture content and structure
- Ø Metadata capture
- Ø Convert record to immutable form
- Ø Lock to prevent modification
- Ø Modify metadata

Record Preservation

- Ø Capture content and structure
- Ø Metadata capture
- Ø Convert record to immutable form
- Ø Lock to prevent modification
- Ø Modify metadata

Review of Records Concepts

HOW DO YOU IMMUTABILIZE?

Ø Pdf /A

Ø TIFF image

Digital Archiving Concepts

The OAIS Model

Open Archival Information System

<http://ssdoo.gsfc.nasa.gov/nost/isoas/overview.html>

Ø Published in 2001 by NASA

Ø Adopted by ISO

Digital Archiving Concepts

The Victorian Electronic Records Strategy

<http://www.prov.vic.gov.au/vers/welcome.htm>

VERS is a framework of standards, guidance and implementation projects which is centered around the goal of reliably and authentically archiving electronic records created or managed by the Victorian government.

The PROV Digital Archive, which is planned to be operational by 2005, will hold electronic records of government departments, local government, statutory authorities, educational and judicial institutions and health organizations.

Digital Archiving Concepts

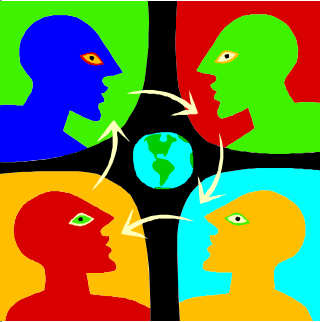
The Victorian Electronic Records Strategy

The VERS Encapsulated Object (VEO) is the mandated format for preservation of the electronic records. VEOs consist of:

- ∅ A record container created in eXtensible Markup Language (XML)
- ∅ One or more documents represented in a VERS-approved long-term document format (at present, PDF, TIFF or ASCII)
- ∅ A digital signature which acts as a "seal" on the record to aid detection of tampering.

eXtensible Access Method (XAM)

- Ø Portability among and between applications and storage devices**
- Ø Storage objects are independent of the platform**
- Ø Unique name (address) for each object**
- Ø Eliminates proprietary API's**
- Ø Archive can be searched via the metadata**
- Ø Storage Networking Industry Association (SNIA)**
- Ø 45 companies contributing to specification**

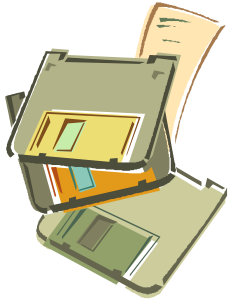


Overview

Long Term Preservation Approaches

- Ø Museum of hardware and software
- Ø Software Emulators
- Ø Migration

Computer Methodologies to Meet Long-Term Archiving Needs



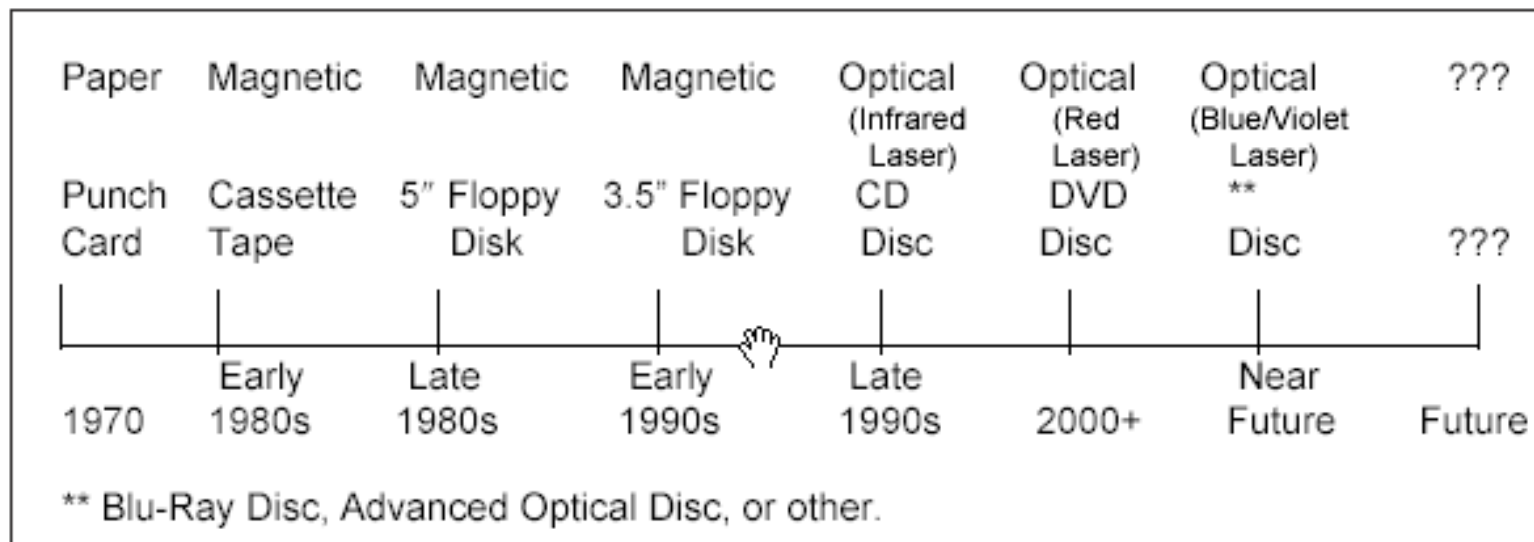
Media

Storage Media Life Expectancy Variables

- ∅ Quality of the media from vendor
- ∅ Quality of the recorder used to write to the media
 - ∅ Number of times data is accessed
 - ∅ Care with which the media is handled
- ∅ Storage conditions (Temp & humidity)
- ∅ Cleanliness of the storage environment

Computer Methodologies to Meet Long-Term Archiving Needs

Media



Computer Methodologies to Meet Long-Term Archiving Needs

Media

Device	25RH 10 C	30RH 15 C	40RH 20 C	50RH 25 C	50RH 28 C
D3 mag tape	50 yr	25 yr	15 yr	3 yr	1 yr
DLT mag tape cartridge	75 yr	40 yr	15 yr	3 yr	1 yr
CD/DVD	75 yr	40 yr	20 yr	10 yr	2 yr
CD-ROM	30 yr	15 yr	3 yr	9 months	3 months

Gold KODAK CD and DVD

∅ “Protects” for 80 to 300 years:

∅ Do not use adhesive labels

∅ Store at room temperature below 50% RH

∅ Protect from light

∅ Store 1 copy offsite



Overview - Hardware

Storage Media Best Practices

- Ø Always follow recognized format standards
- Ø Use only new media
- Ø Purchase high quality stock



Media Migration Plan

Written QC policy

- Ø Verification of all reformatted electronic records
- Ø Complete documentation of all steps
- Ø Backup in different location
- Ø 2 checksums, bit/byte comparison, hash digest
- Ø Convert to microfilm to reduce future costs

Software

Select software

- ∅ Open architecture
- ∅ Follows standards
- ∅ Integrates with other software

Microsoft Office

90% of market

Stable company

Compatible with old versions

Familiar

Computer Methodologies to Meet Long-Term Archiving Needs

Adobe

Post Script

Portable Document Format (pdf)

Digital Signature

Ghost Script

Image Formats

TIFF - The BEST Choice

JPG

GIF

BMP

Computer Methodologies to Meet Long-Term Archiving Needs

Long Term

PDF/A

ASCII

XML

Computer Output to Microfilm (COM)

Optical Character Recognition (OCR)

Computer Methodologies to Meet Long-Term Archiving Needs



Short Term Strategy

2 - 7 Years

Ø What files format(s) would you use?

Ø What storage media would be appropriate?

Ø How much redundancy would you be comfortable with?

Computer Methodologies to Meet Long-Term Archiving Needs



Long Term Strategy

7 - 50 Years

- Ø What files format(s) would you use?
- Ø What storage media would be appropriate?
- Ø How much redundancy would you be comfortable with?

QUESTIONS ??????



*Charlie Sodano, Ph.D.
Senior Consultant*



Charlie.sodano@eorganizedworld.com

Phone: (925) 229 1688

Cell: (925) 231 5909

Martinez, CA 94553