

Digital Archives Day

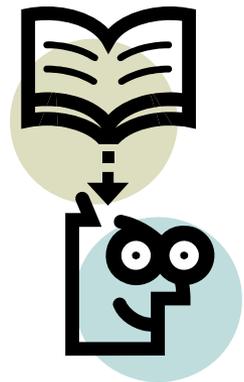
October 24, 2012



Part I

Introduction





Objectives

- Understand the importance of records management to Electronic Records
- Understand Trusted Systems
- Understand Standards
- Understand Records Transfer
- Create a Community for Digital Continuity

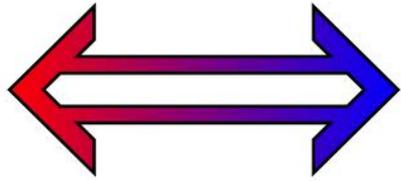


Perspectives

- Records Creators
 - Get the job done
 - Meet legal requirements
 - Maintain digital continuity
- Archives/Library
 - Digital Preservation (forever)
 - After Office life ends
 - Stakeholder in system adoption

We live in the Information Age

- Computers allow us to create many more records
- More than 90% of records are now “born digital”
- Are you including your e-records on your Records Retention Schedule?
- Are you sending E-Records to the Archives? E-Publications to the Library?

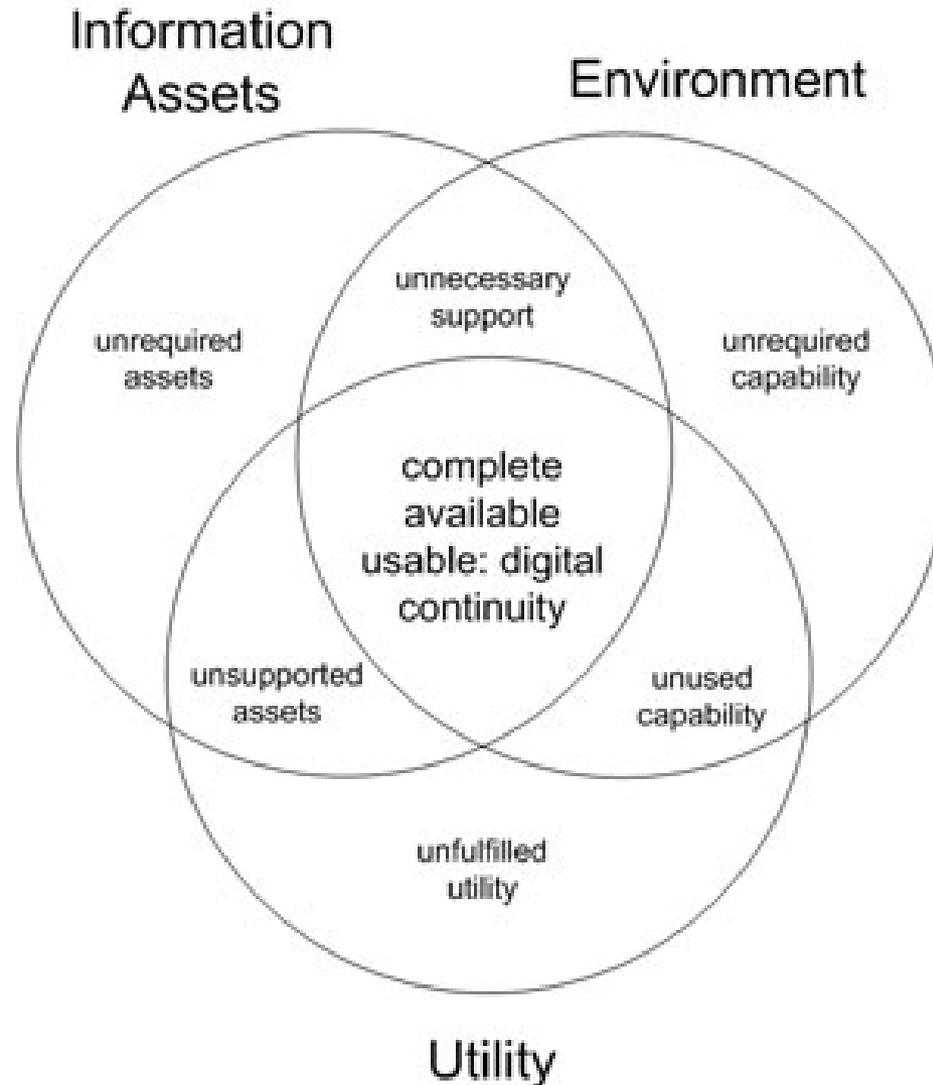


Digital Continuity

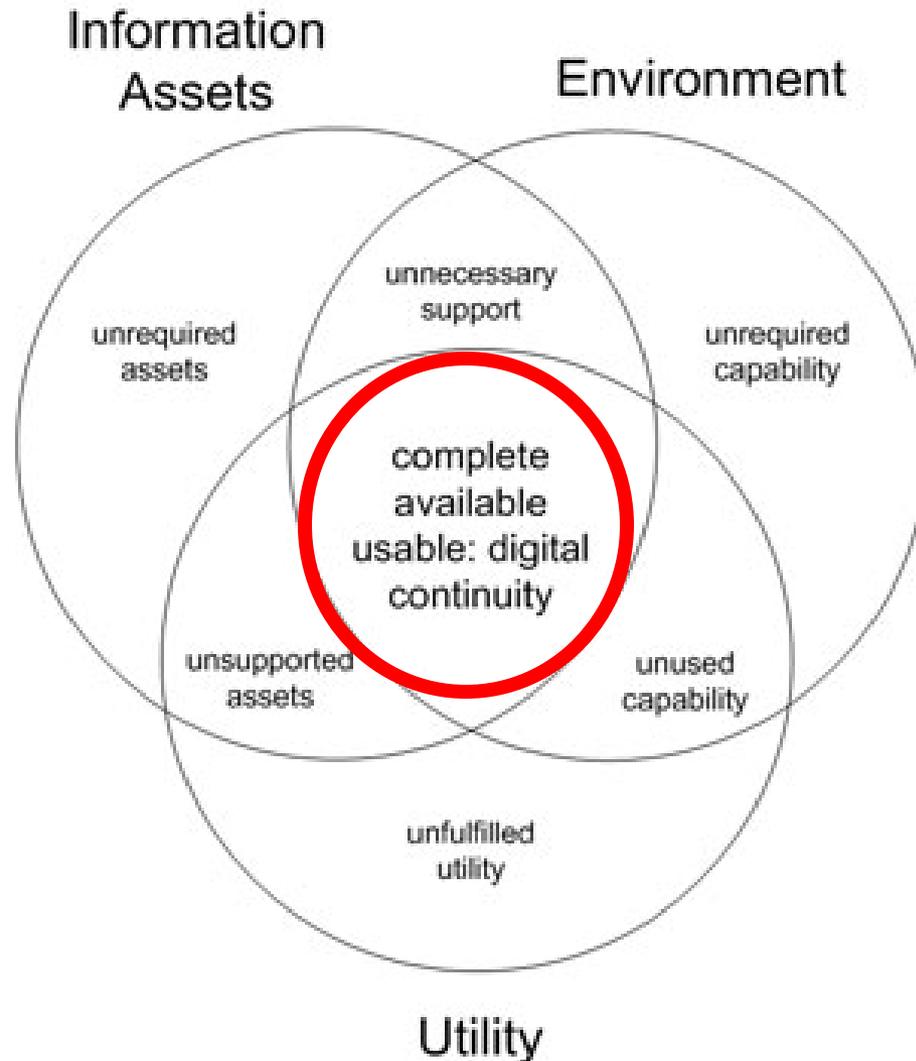
Ensures information is complete, available and useable by those with a need for it.

- You can find it when you need it
- Open it when you need it
- Work with it in the way you need to
- Understand what it is and what it is about, and trust that it is what it says it is

Digital Continuity, Illustrated



Digital Preservation, Illustrated



E-Records

- Archives will “flag” those records on the retention schedule that should be sent to the Archives
- Contact Archives when e-records are ready for transfer



E-Publications

- In the past most would have been printed at State Printing Plant
- To State Library (possibly also State Archives)





Publication

- **Government Code §14902**

“State Publication” or “publication” as herein employed is defined to include any document, compilation, journal, law, report, memorandum, hearing, legislative bill, leaflet, order, regulation, directory, periodical or magazine issued by the state, the Legislature, constitutional officers, or any department, commission or other agency thereof or prepared for the state by private individual or organization...

Part 2

E-Records on Your Records Retention Schedule



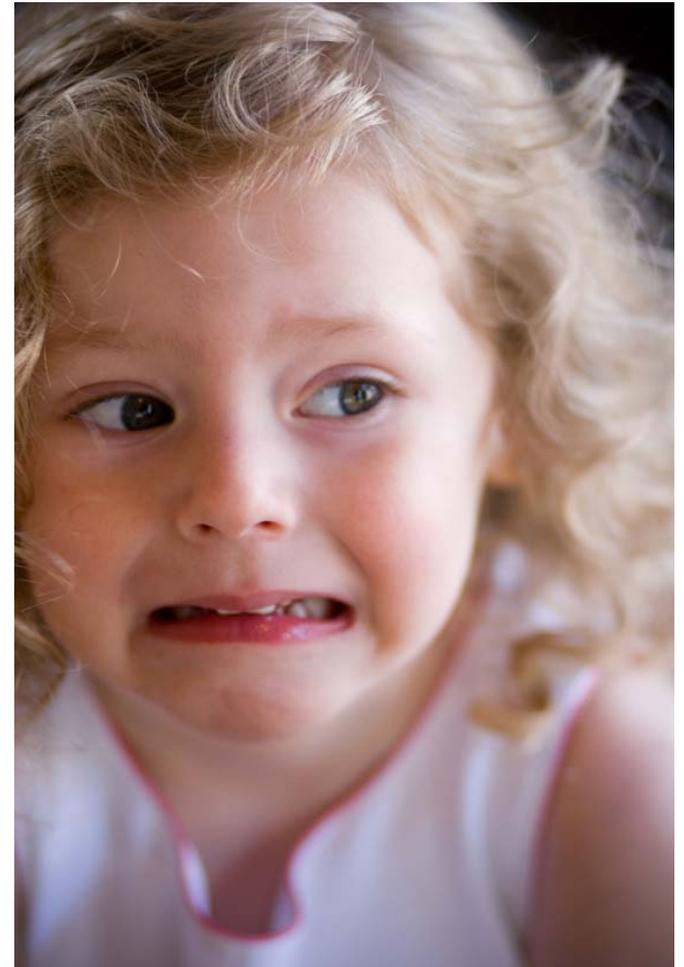
What is a Record?

Definition of a record according to the State Records Management Act:

“Record” or “records” means all papers, maps exhibits, magnetic or paper tapes, photographic films and prints, punch cards, and other documents produced, received owned or used by an agency, ***regardless*** of physical form or characteristics. (Govt. Code section 14741)

The Records Inventory (It's not scary!!)

1. Identify the paper records
2. Find the “E-Go-Withs”
3. Ask Staff about records created electronically & scanned or digitized
4. Follow-up with large shared systems



1. Identify the Paper Records

- If you have a prior schedule for the agency, division, office or unit, confirm that the paper records still exist and add any that were not scheduled in the past.
- If you do not have a prior schedule begin by inventorying the paper records for the division, office or unit.

2. Find the “E-Go-Withs”

Review the paper groups of records
(or record series)

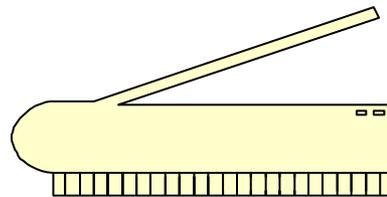
- How does staff access the group of records; by subject, date, assigned number, or name?
- Does an e-index for the records exist?



3. Ask staff about records created electronically & scanned or digitized

Are records regularly scanned, or created electronically, and saved to a shared file?

Examples:



Minutes, Reports, Speeches, Program Files, Images

4. Follow-up with large shared systems

- Does the office maintain a shared tracking system to track projects and project assignments?
- Does the office contribute to a GIS or larger shared system?
- Are records from the office on the agency website/intranet? (Are the website records copies or originals?)



Once You Identify E-Records...

Make a note on your inventory regarding the software/data system used and add it to the remarks column of the schedule.

Examples: *Excel, Access, Word, Powerpoint*, Open Source, or Proprietary System (and what is it?)

Scheduling the E-Record & Gathering the Metadata

Metadata or “Data about data”

- What are the records?
- Who created them?
- Are there restrictions on access?
- What is important about the records?
- How long should they be kept and what makes the records inactive?
- Put this on your retention schedule



Don't Forget



- As with paper records, the ***creator*** of the e-record establishes the retention period not your IT staff

Part 3

Metadata Matters





More About Metadata

- Remember, it's "Data about data"
 - A lot of the metadata should appear on the records retention schedule
 - Some may be automatically created by your software
 - Different types may be needed by different audiences (e.g. public searches may be different than in-office searches)



Metadata Types

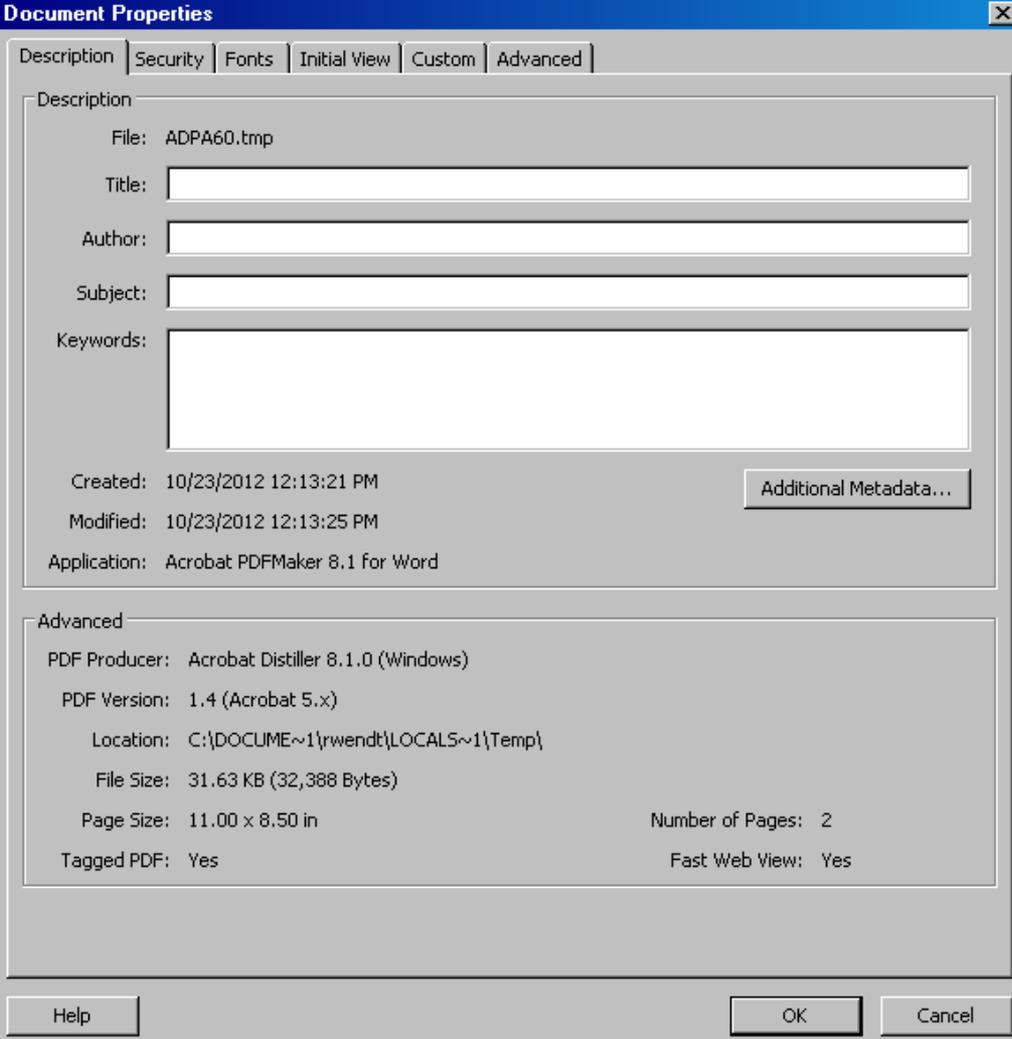
- **Descriptive**
 - What is the record about? (e.g. subject tags)
- **Administrative**
 - Who owns and manages the record?
- **Structural**
 - How should it be displayed and how is it related to other records or parts of records?
- **Technical**
 - The nitty-gritty about file size, resolution, and/or software and hardware
- **Preservation**
 - All of the above +what has been done (e.g. migration) to keep the record over the long term and who has had access

Metadata Standards

- Many options
 - Dublin Core
 - CSDGM (Content Standard for Digital Geospatial Metadata)
 - Others
- Select on basis of needs (legal, business, user)
- Fields to be populated with information about your records
 - ID, Title, Creator, Subject, Date, etc.

Metadata Entry

Acrobat:
(for example)



The screenshot shows the 'Document Properties' dialog box in Adobe Acrobat. The 'Description' tab is active, displaying the following information:

- File: ADPA60.tmp
- Title:
- Author:
- Subject:
- Keywords:
- Created: 10/23/2012 12:13:21 PM
- Modified: 10/23/2012 12:13:25 PM
- Application: Acrobat PDFMaker 8.1 for Word

An 'Additional Metadata...' button is located to the right of the 'Created' and 'Modified' fields.

The 'Advanced' tab is also visible, showing the following information:

- PDF Producer: Acrobat Distiller 8.1.0 (Windows)
- PDF Version: 1.4 (Acrobat 5.x)
- Location: C:\DOCUME~1\rwendt\LOCALS~1\Temp\
- File Size: 31.63 KB (32,388 Bytes)
- Page Size: 11.00 x 8.50 in
- Number of Pages: 2
- Tagged PDF: Yes
- Fast Web View: Yes

At the bottom of the dialog box, there are buttons for 'Help', 'OK', and 'Cancel'.

Part 4

Trusted Systems





Definition

- dreamstime.com Government Code section 12168.7 (c)
 - A trusted system is “a combination of techniques, policies, and procedures for which there is no plausible scenario in which a document retrieved from or reproduced by the system could differ substantially from the document that is originally stored.”

Regulation

- Trustworthy Electronic Document or Record Preservation Regulation
 - 2 CCR Sections 22620.1, 22620.2, 22620.3, 22620.4, 22620.5, 22620.6, 22620.7 and 22620.8
 - Effective for State Agencies Aug. 12, 2012
 - Based on AIIM ARP1-2009

Trusted System Applies

- To e-records created or stored as the official record six months after the effective date of the regulations (Aug. 12, 2012 + 6 months)
- When destroying the original hardcopy and maintaining an electronic version as the official record (defined by statute or business practice)
- To Born-Digital and official digital records as well as digitized records



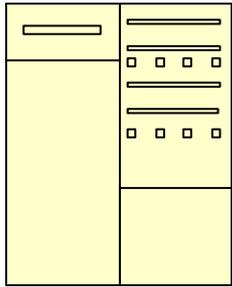
A Trusted System Policy Will Describe...

- How information will be scanned, indexed, and verified
- How the system will be secured from unauthorized access
- How documents will be secured from unauthorized modification or alternation
- How authorized modification of documents will be managed, including audit trail information and the ability to retrieve any previous document version required to be maintained



A Trusted System Policy Will Describe (continued)

- How notes and annotations (if any) will be stored and managed, if they are a part of the record
- How these policies and procedures will be followed
- How the system will adhere to the published records retention schedule



EDMS

- Electronic Document Management System (and related: ERDMS, etc.)
- Manages creation, storage, and control
- Can the system employ retention schedule-compliant policies?
- Discuss with all stakeholders (users, records managers, IT staff, and archivists) before deploying

EDMS Considerations

- Access to public records
- Security of records (including non-public records)
- Transfer (or disposal) of records contained in system
- Legal requirements
- Metadata preservation
- Version control
- Additions/Extensibility

EDMS Selection Should Consider

- Processes
- Procedures
- Technology requirements/definitions
- Business Objectives and Requirements
- Legacy Data/Document Conversion Methodology
- Metadata
- Other

See: AIIM ARP1-2009 “Analysis, Selection, and Implementation of Electronic Document Management Systems”

<http://www.aiim.org/documents/standards/ARP1-2009.pdf>

Trusted Storage

- **ISO 15801** compliant (Document management-- Information stored electronically-- Recommendations for trustworthiness and reliability)
- **ISO 15489** compliant (Information and documentation -- Records management)
- At least 2 separate copies of document
 - one copy of the electronic document to be stored and maintained in a safe and separate location
- No unauthorized additions, modifications or deletions
- Independently auditable



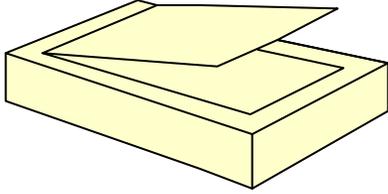
Trusted Formats

- Industry-accepted format in widespread use
- Non-Proprietary/Fully documented
- Non-modifiable through the file format structure (i.e. support “checksums”)
- PDF/A preferred
- Fully-documented TIFF may work

Part 5

Digitization Standards





Digitization

- Can increase accessibility
- High cost
- Should follow standards for increased utility
- Must be trustworthy
- For official records must follow regulations
- Determine your need and audience before starting
- What, legally, can happen to original paper records?
- Quality control a must!
- How will files be indexed?



Digitization Standards

- Good rules of thumb for long-term preservation:
 - Text Master Files
 - 400-600 dpi, 8-bit (grayscale) or 24-bit (color)
 - Photo/Negative Master Files
 - 4000-8000 lpi (because size of original varies it's better to use lines per inch), 8-bit (grayscale) or 24-bit (color)
 - Drawings Master Files
 - 400-600 dpi, 8-bit (grayscale) or 24-bit (color)
- Can create lower-resolution files for Web or other use
- Is redaction necessary?
- Optical Character Recognition (OCR) assists with searching and utility
- Don't forget metadata



Compression

- Lossless or Lossy
 - Lossless is better
- To meet regulations, if compression is used it must not include extraneous information unsupported by relevant industry standards and may not include proprietary alterations of the algorithms

File Formats

- Proprietary
 - Controlled by one provider
 - Accessed by one software system
- Not Proprietary
 - Offered by more than one provider
 - Can be accessed with multiple systems
- Open Source
 - Code is available and can be modified
 - More sustainable
- Open Standard
 - Developed using open source specs
 - Increases compatibility options
 - More sustainable

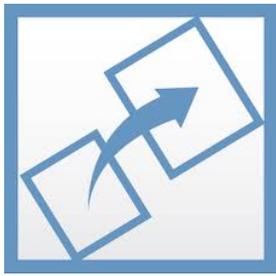


File Types

- Text
 - DOC, TXT, RTF, PDF, PDF/A, etc.
- Graphics
 - Vector
 - EPS, SHP (e.g. GIS files), etc.
 - Raster
 - TIFF, JPEG/JPEG2000, etc.
- Data Files
- Spreadsheets
- Audio/Visual Files
 - WAV, MPEG, etc.
- Markup
 - SGML, HTML, XML, etc.

The File Formats and Types You Choose Will Matter Later

- You can't convert to another version if you don't know what you have
- Migration to another is difficult without indications of format/type
- Both may reach obsolescence quickly; records may become inaccessible



Conversion

- Causes format change
- Maintains accessibility
- Necessary if file format is no longer supported
- Must be done in a trustworthy manner
- Must be done carefully to avoid loss and future accessibility issues

Migration

- Move to other storage medium/place
- Does not necessitate conversion but may be impractical without first converting
- Must be planned carefully to avoid loss



Content/Context/Structure

- **Content** = the information
- **Context** = the way it looks
- **Structure** = information needed for full understanding
- All must be present to maintain the records
- Will your actions change one or more?

Part 6

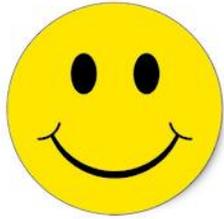
Email & Websites





Email

- Is frequently a public record
- The content of the email message determines whether or not the message is a record
 - Retention should be based on content not on creation method
- Should appear on a records retention schedule with similar records in other formats, not as a separate series
- Record email should include *transactional* information
 - Metadata: the sender and all recipients, date and time of message creation and sending, etc.



Record Email

M																							<p>E-mail records that are classified as official records are subject to the individual department's records retention schedules and must be retained for the same period of time as the records series that most closely matches the subject matter contained within the new e-mail message. If there is no entry that resembles or matches the subject matter of the e-message, the "record" should be added to the appropriate retention schedule as a separate series of records.</p>
M																							<p>Transitory E-mail consists of electronic messages that are created primarily for the communication of informal information as opposed to the perpetuation or formalization of knowledge. Destroy transitory e-mails when they have served their purpose.</p>

Managing Email

- Establish a policy that follows the Records Management Act
- Define roles and responsibilities
- Make management part of users' workflow
- Determine record copy (duplicates will abound!)
- Security is a must

Websites

- Are you posting records on your website?
- Are those records on your Retention Schedule?
- What happens when documents are updated/replaced?
- RRS feed for website can help Archives/Library with records



Part 7

Transfer of E-Records



Transfer of Records

- Please make sure records are on your Records Retention Schedule
- If we “flag” them, they should be transferred
- Contact us and prepare forms
- Transfer may be on portable media or through another method (FTP, possibly)
- We will work with you

What Happens at the Archives?



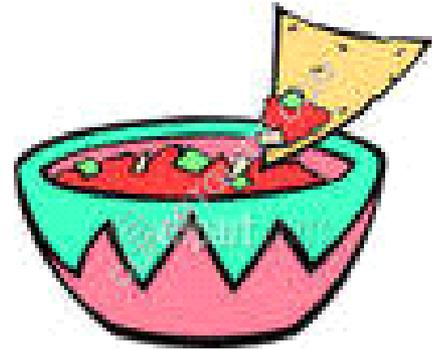
SIP

+



AIP

+

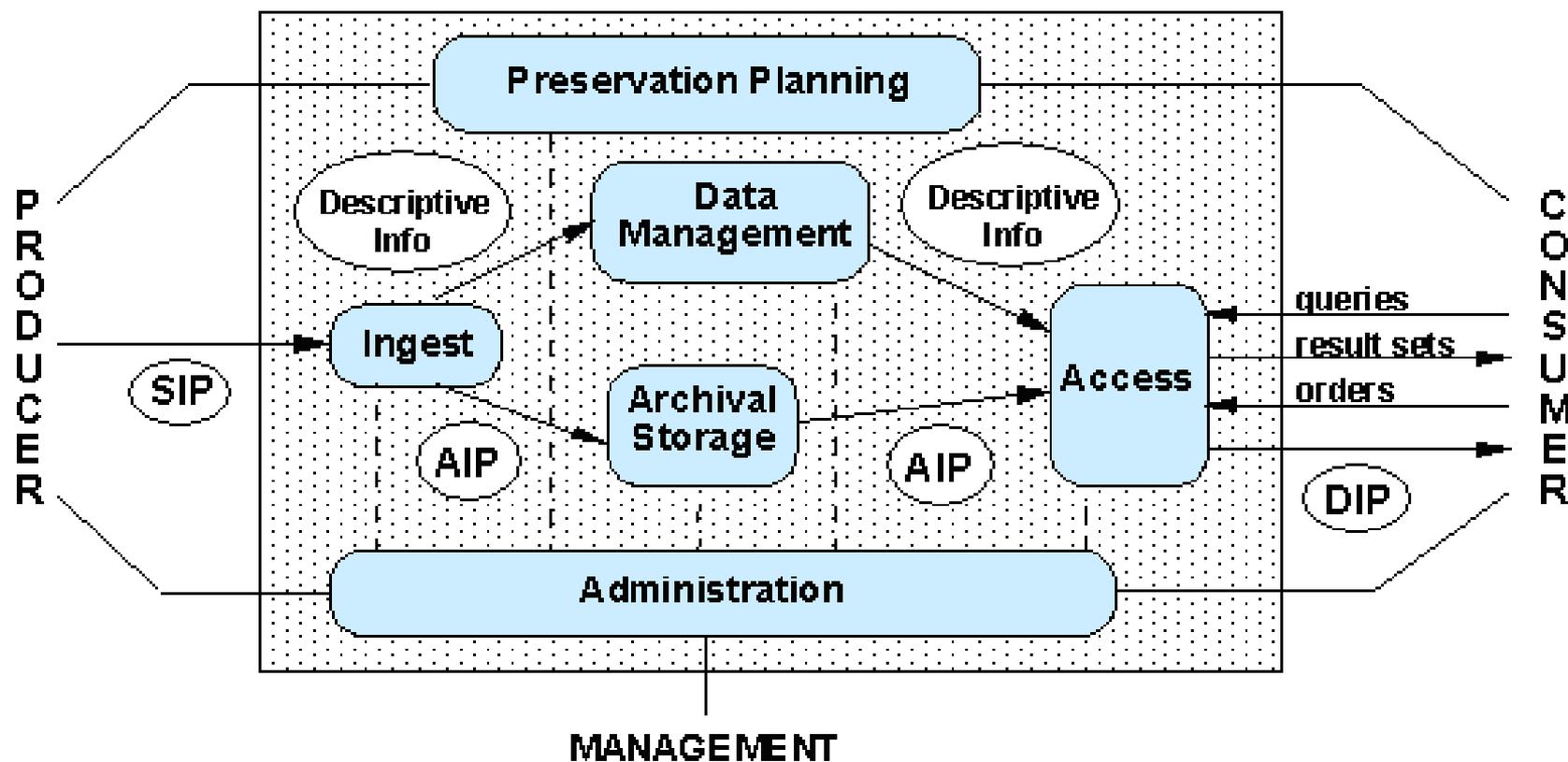


DIP

You're Preparing the SIP

- This is the Submission Information Package
- We'll work with you to transfer your records
- Please contact us

Open Archival Information System



Source: Procedures Manual for the Consultative Committee for Space Data Systems (2001)

Long-Term Digital Preservation

- We're trying to make these records last forever
- We need your help!



Access

- Beyond preservation, there is public access
- We are working toward building a “digital repository” to make records available 24/7
- Until then, we can make copies available at our facility and on demand
- Thanks for helping

Questions/Discussion

