Records Management: Life Cycle of Born-Digital Documents / Best Practices

November 19, 2013

William Saffady

(718) 246-4696

wsaffady@aol.com

Life Cycle Issues

- Life cycle defined by retention period: Formalized retention guidance is a precondition for life cycle management

- Importance vs. Paper Records
  - Most important information is born digital
  - Many born-digital records never printed
  - Some born-digital records cannot be printed

- Usability
  - Format obsolescence
  - Media stability
  - Data migration

Different Life Cycle Concerns:
Records Management vs. Archives

- Retention vs. preservation
  - Temporary records
  - Permanent records with no scholarly value

- Emphasis on business significance of records
  - Transaction processing
  - Project management
  - Planning and decision-making
  - Legal and regulatory compliance
  - Discovery

Retention of Born-Digital Information

- Official copies vs. duplicate records

- Criteria for retention decisions
  - Legal considerations
    - Compliance with recordkeeping requirements
    - Retention for possible use in legal proceedings
  - Operational need
  - Scholarly value

- Given the declining cost of digital storage, is retention guidance necessary?
  - Competing demands for storage from data-intensive applications
  - Performance implications
  - Backup issues
  - For compliance and the duty to preserve
  - Discovery costs

Life Cycle Issues for Databases

- A concern for very large, mission-critical databases

- Database performance is the most important motive for retention guidance

- Archiving database records in place
  - In main database
  - Transfer to history/archive file

Life Cycle Issues for Databases

- Purge functionality for older database records may not be supported
  - Deletion of records can have unintended effects
  - Software modification required

- Strategies for permanent preservation
  - Keep database records in original format
  - Convert database records to software-neutral format
  - Keep legacy applications in service
  - Print database records for retention

Unstructured Born-Digital Records

- Common examples
  - Digital documents created by office applications
  - Email
  - Digital photographs
  - Video and audio clips
  - Web pages
  - Collaboration sites
  - CAD files

- Storage locations
  - Network drives
  - Local drives
  - Offline media
  - Cloud services or other remote sites

Life Cycle Issues for Unstructured Records

- Traditional approach to retention
  - Retention schedule lists records by functional categories or custodial department
  - Retention periods based on content not format of records
  - Records must be individually evaluated for retention

- Uniform retention period as an alternative
  - Easily understood and practical
  - Exceptions for records requiring longer retention
  - Some records may be kept longer than necessary

RMA Software

- A managed repository for retention of digital documents as official copies
  - Focus on inactive records: contrast with document/content management products
  - Limited to digital records covered by retention guidance
  - Records imported from originating applications or other sources
  - Role of DoD 5015.2-STD
  - Comparison with offsite storage for physical records

- Records in repository cannot be modified
  - Version control: Revisions treated as new records
  - Deletion when pre-defined retention period elapses
  - No possibility of inadvertent destruction
  - Notice of impending destruction is issued
  - Can suspend destruction for litigation or audits

RMA Software

- ## Retrieval functionality
  - Based on metadata or full-text indexing
  - Boolean searches
  - Common interface regardless of originating application
  - Relevance ranking of search results
  - Records displayed in native application or with multi-format viewer

- ## Security
  - Access tightly controlled at folder and document level
  - Audit trail tracks all retrieval transactions, including failed attempts

Email Archiving Systems

- Single-silo repository for messages and attachments
  - Messages transferred from email servers
  - Messages and attachments archived in original formats
  - Transferred may occur at pre-defined intervals or when mailbox capacity reaches specified level
  - Some systems support real-time capture of messages
  - Organization of archival repository mirrors email system

- Implementation configurations
  - Software vs. archiving appliance
  - On premises vs. cloud-based service

Email Archiving Systems

- ## Access and retrieval
  - Mailbox owners have full access privileges, others may have read-only privileges
  - Access via email client or web browser
  - Searches based on metadata or full-text indexing of messages and attachments
  - Boolean searches
  - Cross-mailbox searches where authorized
  - Retrieved messages can be returned to mailboxes, forwarded, replied to

Email Archiving Systems

- Retention functionality
  - Message deletion based on pre-defined retention rules
  - Retention periods assigned at mailbox, folder, or message level
  - Can override retention periods for litigation or audits

- Other features
  - Single-instance storage / de-duplication
  - Encryption
  - Content screening
  - Automatic categorization assigns messages to topical folders based on subject content
  - Statistical analysis of message activity
  - Audit trail tracks access attempts and other message-related events