Digital Imaging for Records

Introduction

Digital Imaging Defined…

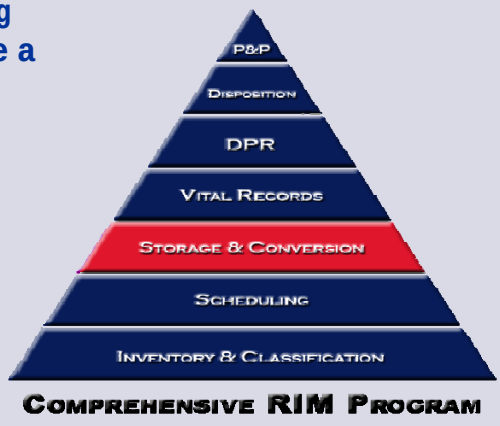
- **The conversion, storage, and distribution of information displayed but not directly modified by a computer…**

GUIDELINES & ADVICE...

- Advisory *not* requirements
- Based on:
  - national technical standards
  - established practices
  - research
- Critical issues to consider...
- Especially important for systems used for mission critical records...
- Public officials are responsible by law for ensuring that their records are protected and accessible *regardless of the media involved.*

PART OF A LARGER RIM PROGRAM...

- A records imaging project should be a component of a broader, comprehensive records and information management program...



Digital Imaging for Records

Project Planning

The "Basics"

- Part of your more, all encompassing RIM program…
- Multi-disciplinary Imaging Team…
- Conduct business analysis & cost justification…
- Set realistic goals & timelines…
- Develop RFI/RFQ/RFP…

Business Process Analysis

- Prior to selecting a digital imaging system, conduct a workflow and records analysis:
  - ...opportunity to reengineer business process for operational efficiency...
  - ...ID & document existing and planned agency information needs...
  - ...determines which records are best suited for imaging...

Cost Justification

- Conduct a cost benefit analysis to justify system purchase / outsourced services...
- Compare the costs of your current operation with the costs of the new system...

Cost Justification

- **Consider**
  - File creation, maintenance, disposition
  - Potential changes/improvements
  - Proposed system architecture
  - Equipment pricing
  - Financial measures (ROI)
  - Analysis tools available at ERC's Web site

Digital Imaging for Records

Technical Specifications & Selection

Open Architecture

- **Require open system or require developers to provide a bridge**
  - **Flexible upgrades without**
    - **Additional costs**
    - **Minimal function loss**
    - **Major risk of losing records**
  - **Bridge to non-proprietary standards**
    - **Allow access from and transfer to other systems**

Non-Proprietary Format

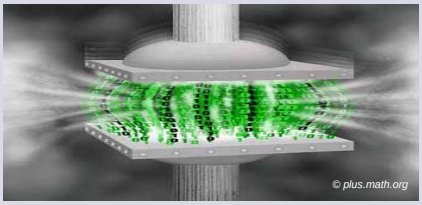
- **Use a non-proprietary digital image file format or provide a bridge**
- **No single industry-wide image format standard**
  - **Tagged Image File Format (TIFF)**
  - **PDF/Archive (PDF/A)**
  - **JPEG2K**
  - **PNG**

COMPRESSION

- **Lossless – a bit for bit recreation of original image…**
- **Lossy – a sampling of the image that "tosses" out actual bits leaving a distorted recreation of original image…**

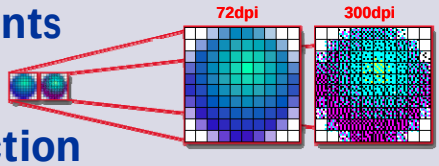
NON-PROPRIETARY LOSSLESS COMPRESSION

- **Black & White**
  - **TIFF w/Group IV (Group III for fax machines)**
  - **JPEG2K**
  - **PDF/A**
  - **PNG**
- **Grayscale & Color**
  - **TIFF w/LZW**
  - **JPEG2K**
  - **PDF/A**
  - **PNG**



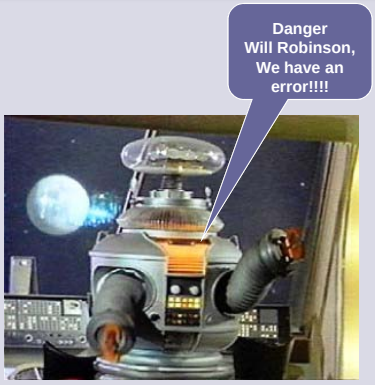
SCANNING RESOLUTION

- Storage requirements
- Throughput Rate
- Accurate reproduction
  - Textual Documents - B&W 300dpi for good quality images
  - Photos, maps, illustrations may need >300dpi and in grayscale or color



ERROR DETECTION

- Equipment conforms to standard error detection and correction methodology…
- System should provide techniques to verify records on digital media…
- System administrator should actively monitor the status…



VERIFICATION

- Verified at the point of writing to storage media…
  - …via "Write and Verify" at the hardware level…
  - …or at the software level
- If not feasible, your quality assurance procedures must be specified (e.g. viewing of documents downstream or spot checks of records as required)

DIGITAL IMAGE METADATA

- …is information about an individual datum or sets of data that is used to facilitate the its understanding, use and management…

Description = Ohio Statehouse

Creator = Unknown

File Name = ohio_statehouse_-_credit_rod_berry_capitol_square_review_and_advisory_board.jpg

Dimensions = 3600pX2400p

INDEXING DATABASE

- **Should provide efficient retrieval, ease of use, and up-to-date information about the images stored**
- **Indexing methods include**
  - **Manual key entry**
  - **Match and merge**
  - **Targeted OCR**
  - **Barcode**
- **Should be selected after an analysis of agency operations and user needs**

AUDITED PROCESSES

- **Based on legal POV…**
  - **"Best available copy"**
  - **Audit trail to ensure that record has not been changed**
  - **Annotations as a separate "layer"**
  - **Records cannot be deleted once entered**
  - **Records cannot be directly accessed by the user--application is an abstraction layer**
  - **Optional use of non-rewritable media**

Digital Imaging for Records

System Implementation

Team Roles...

- **Organization**
  - Project Manager
  - Decision Makers
  - Systems administrator
  - Train the Trainer
  - Users
- **Vendor**
  - Project Manager
  - Implementation Staff
    - Hardware
    - Software
  - Training
  - System Support

Operational Procedures

- **Provide technical & administrative documentation to ensure…**
  - **Future usability of the system…**
  - **Continued access to long-term records…**
  - **Sound foundation for assuring the system's legal integrity…**

Operational Procedures

- **Written record of:**
  - **Procedures**
  - **Operating systems**
  - **Decisions & Changes**
  - **Updated on a regular basis**
- **Consistent with requirements for the admission of records under the rules of evidence laws**

IMAGE INTEGRITY

- **Institute procedures to ensure quality, integrity of scanned images**
  - **Visual verification at start of day**
  - **Inspection of images and system components to insure accessibility**
  - **Regular audits based on sample images**

DIGITAL IMAGING FOR RECORDS

ARCHIVING & LONG-TERM MAINTENANCE

Record Retention Redux

- **Retention and disposal of digital images and corresponding records should be incorporated into retention schedule**
  - **Store images so that they can be identified by that schedule**
  - **More than normal file delete may be required**

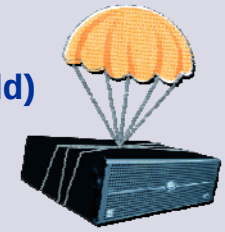
Long-Term Documents

- **Specific plans for creating and sustaining digital images that will be retained more than 10 years**
- **Master image capture**
- **System information and maintenance**
- **Sustainability**

Domesday Book
commissioned by William the Conqueror
completed 1086 and still accessible.
Original Digitization in 1986
inaccessible w/in 15 years

Disaster Recovery Plan

- **Comprehensive records and information disaster recovery plan…**
  - **Identification of vital records**
  - **Off-site facilities (hot, warm, cold)**
  - **Periodic tests**
  - **Redundant recovery site**

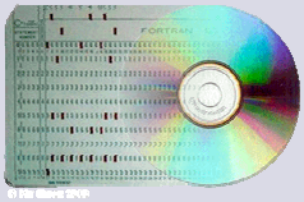


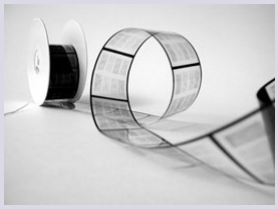
Backups

- **Design procedures to create secure copies of images and related index records**
  - **Regular audits to determine validity and completeness…**
  - **Ongoing testing to verify restore capabilities…**
- **Backups should allow for disaster recovery…**

Impact of Long-term Retention

- **Technology is constantly changing—electronic records may not be stable, reliable, authentic and accessible over the long term, therefore…**
  - *If retention > 10 years, you may need to plan for eye-readable backup (paper or microfilm)*





Refreshing Media

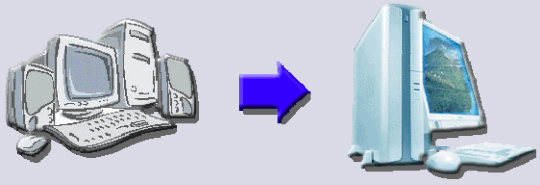
- **Essential to avoid degradation and to facilitate long-term storage**
  - **Copy data onto identical media**
  - **Reformat data from obsolete storage device**
- **Periodically**
  - **w/in time specified by supplier**
  - **as new storage devices are installed**
  - **if audit discloses significant read errors**

MIGRATION

- New software, platforms, file formats
  - Reformatting existing file formats to new ones
  - Migrating one component (e.g. database) to new hardware and/or software platform
  - Migrating the whole system from one hardware and/or software platform to another

MIGRATION

- Document the changes made to the hardware, software, and file formats
- Include changes that could affect data viability

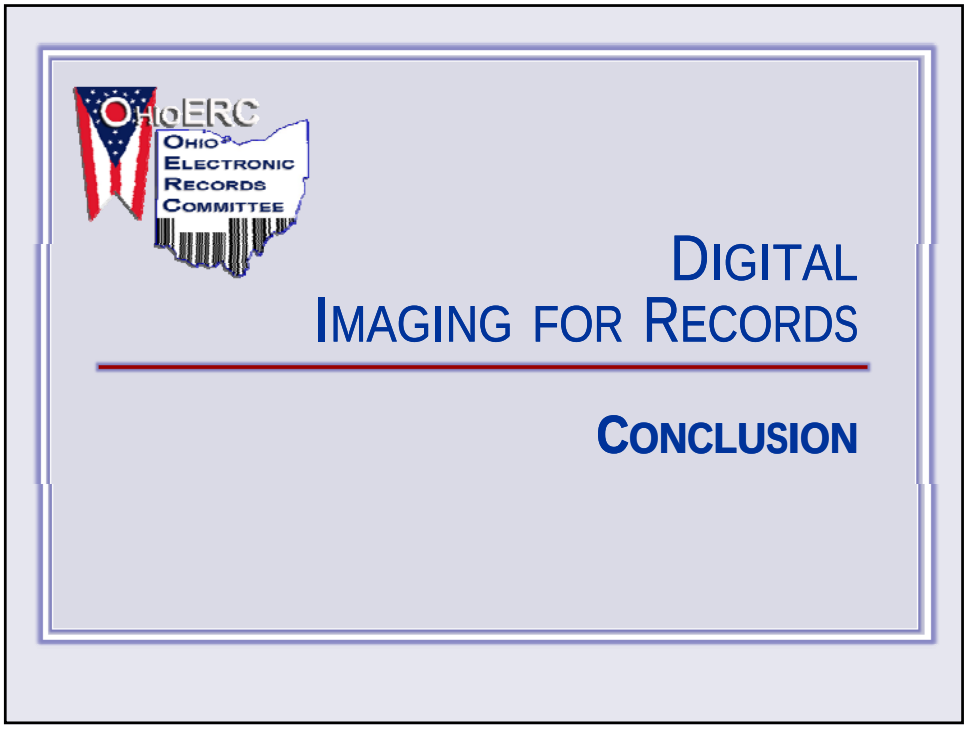


SUSTAINABILITY

- System (hardware & software) could be operational 10+ years
- However, technology often superseded in 2-3 years

ONGOING COSTS

- Budget annually between 20-25% of original system acquisition cost for upgrades, training, maintenance
- If *not* factored into continuing support:
  - system can become obsolete
  - Would require costly outlay to restore effectiveness, if at all possible


OHioERC
Ohio
Electronic
Records
Committee
Digital
Imaging for Records
Conclusion