Overview

After December 31, 2022, NARA will no longer accept new transfers of analog records for storage by the Federal Records Centers Program (FRCP) to the fullest extent possible. NARA will continue to store and service all records received by the FRCP by that date until their scheduled disposition date. Once those records reach their disposition date, NARA will accept the permanent records into the National Archives in their original (analog) format.

Beginning January 1, 2023, all other legal transfers of permanent records must be in electronic format, to the fullest extent possible, whether the records were "born electronic" or not. After that date, agencies will be required to digitize permanent records in analog formats before transfer to NARA.

This paper looks at the benefits, risks and challenges of moving from a paper-based process of records preservation to digital format.

Current Process

In addition to the aforementioned NARA policy changes driving the need to accelerate digitization efforts, the shift to an electronic records management program offers many operational benefits to improve agency productivity, efficiency and compliance. Transitioning to an electronic environment enables agencies to work with and leverage information in a way that analog record formats do not always allow. Analog records exist in a static state which limits information from being easily searched, updated or transferred without manual intervention.

Digitization Benefits

Digital records stored with appropriate metadata facilitates a more dynamic interaction such as allowing users to take advantage of text-searchable fields, quickly organize records based on similar characteristics a, or apply policies/rules to an entire records group. Additional benefits of transitioning from a physical to digital records program are:

- Storage
  Real estate space is expensive, and eliminating paper storage can free-up valuable space, reduce rent, reduce off-site document storage fees or potentially open up a new office. Additionally, not all storage space is created equal. Without proper maintenance and environmental controls, physical records can be inadvertently exposed to security threats or degradation. Document imaging offers a safe repository for your data that requires significantly less real estate and can also enable information to be readily shared via a cloud platform or your local document management system.

- Accessibility
  Electronic information is easy to search and access, and more readily accessible via a well-organized document management system. Documents that have been converted can be accessed through the cloud or system using any device that has internet, anywhere or anytime. It also allows increased productivity by enabling employees to find the information they need quickly. Various studies have shown that employees spend roughly

20-30% of the workday (or 1.6 - 2.5 hours per day) searching and gathering information.[1] Transitioning to a digital environment can reduce the time to locate a file and enhance the ability to quickly share, collaborate, exchange and access documents, reducing the turnaround time and further increasing the efficiency for your business.

- Security
  A scanned document with metadata is a trackable document. If needed, only certain users can access the documents and workflows can be set up along with permission groups for an individual, which enhances security and maintains confidentiality of the document. Digitized records provide agencies with tighter chain-of-custody controls because the record itself is secured within a centralized document repository. When accessed by a user or multiple users, the record never physically leaves the repository, which reduces the risk of lost or misplaced files. Agencies can utilize document check-in and check-out features to automatically track access or changes made by authorized users, which also increase accountability.

- Disaster preparedness and recovery
  Whether natural or manmade, fire, flood, earthquakes, etc. may cause a major disaster for your paper documents. During a disaster, access to critical business and operational documents may be limited, which can hinder an agency's ability to maintain continuity of operations and recover quickly. Digitizing records provides a repository that can be shared on cloud or an agencies document management system. This can play an important role in your COOP and disaster recovery planning to minimize downtime and enable remote access to critical documents during emergencies.

- Cost reduction
  Each of the benefits listed above directly or indirectly results in cost savings. Real estate savings gained from eliminating physical storage space can be sizable and realized relatively quickly. Agencies can recoup dollars spent leasing and maintaining real estate, or repurpose the space and resources for higher value activities. Savings gained from improved information access and collaboration has a positive impact on employee productivity. Excessive searching through paper records and maneuvering through manual processes is an inefficient use of employee time. Not only does this result in unnecessary costs and delays, but also contributes to employee fatigue and dissatisfaction. The cost of security breaches and poor disaster planning can be significant depending on the nature of the records and/or the scale of the incident. Taking proactive measures to protect records in the event of a breach or disaster is generally more cost-effective than reacting after an incident has occurred. Depending on the severity of the situation, the cost to remediate exposed, lost or damaged records and/or recover from operational downtime can far exceed initial investments to safeguard records through a modern digitization strategy.

Digitization Risks

---

[1] http://blog.xenit.eu/blog/do-workers-still-waste-time-searching-for-information

Transitioning from a paper-based records management program to a digital environment does not come without risks. The full scope of services, associated costs and required resources of digitization services is still in development. Careful thought and planning is required to develop a successful digital transformation strategy. The following risk factors should be considered before an agency implements a digitization program:

- Specialized/Technical Expertise
  Records management will require additional resources and skill sets to implement this service via external vendors and an internal delivery. Agencies may require additional support, including consulting, interim storage, full-service scanning, IT systems implementation, or software. Additionally, close collaboration between internal RM and IT is critical for integrating records policies and management capabilities into future electronic management systems. Securing the right resources for each stage of the digitization process requires thorough vetting of vendors and ongoing project management. Without proper selection and management of supplementary resources, an agency risks launching digitization projects prematurely without a long-term vision or adequate support to fully complete the transition.

- Funding
  Digitization projects may require a shift in the way agencies fund their records management budget. Migration to the working capital fund may require transfer of IT funding and introduction of additional governance complexities.

- Cybersecurity
  While digitization offers several security benefits including enhanced chain-of-custody, access tracking and authorization, it also introduces cybersecurity as a new/different type of threat. As previously discussed, IT expertise and engagement are required early in digitization process to set up, monitor and protect digital data according to Federal information system standards.

- Mixed Analog and Digital Records Management
  It may not be possible or prudent for your agency to digitize all records. If you digitize only some of your analog records, the result will be a mixed records environment. In these instances, agencies must consider the additional complexities involved in standing up an effective records and information management program that manages both analog and digital records with consistent policies applied regardless of format. Ideally, management of both physical and digital records can be merged under a single, holistic program with centralized oversight. But there is a risk for mixed records environments to diverge into entirely separate records programs with duplicative or uncoordinated policies that are inconsistently managed.

Factors for Initiating a Digitization Project

Agencies must take into account various factors to determine an optimal digitization strategy that balances records accessibility and budget. It cannot be assumed that simply scanning all agency records will automatically produce a more efficient, digital environment. Depending on factors such as activity levels, frequency of access and retention schedules, it may not be practical from

either a financial or operational standpoint to digitize some records. The following factors should be taken into consideration to decide what records should be digitized:

- Access Requirements – Highly active or frequently accessed records are strong candidates for digitization, particularly when these records require access by multiple users at the same time. Digitizing records, then implementing a system to store and access images is an investment. Agencies can realize a greater return on investment with highly active records when the direct cost of physical storage and frequent retrievals combined with the indirect costs of lost productivity and workflow inefficiencies exceed the investment to digitize. In contrast, the cost to prep, digitize, set up and maintain a system for inactive records that are rarely or never accessed will often yield a poor return on investment. For inactive records, it may be more cost effective to continue storing and managing in physical form until disposition, and digitizing only on an as-need basis.

- Retention Schedules – Agencies should consider retention schedules when prioritizing what records to digitize and when. Depending on access and/or final disposition requirements, it may not be worthwhile to digitize records near the end of their retention. The onset of a digitization project is a prime time for agencies to review and refresh their retention scheduled and policies. Assessing records retention is an important step in the overarching digitization process to maximize cost savings. Identifying and dispositioning records past retention can prevent unnecessary scanning, reduce storage costs, and minimize risk of noncompliance.

- Record Workflows – If an agency is implementing a new digital workflow where all new records will be born digital on a go-forward basis, a digitization strategy to convert existing legacy paper records should be a part of the implementation plan. The decision on which legacy records to digitize and when may still vary depending on the other factors previously described. Types of processes that may be better suited for implementing workflow automation include records with standardized fields or inputs that may also require basic approval or rejection reviews (e.g., Accounts Payable, HR, Contract Management).

Other Considerations for Successful Digitization
- Indexing & Metadata Indexing – scanned images are a vital component of any digitization project. Capturing relevant metadata enables scanned records to be indexed for easier search and retrieval, and in some instances can allow some management polices to become automated (e.g., automatically calculate retention dates based on specific date or trigger event criteria). The process for capturing metadata may be completed through various means including manual entry, barcode scanning, or data downloads. In order for metadata to be useful, it must be captured completely and consistently for all relevant records. For this reason, agencies are encouraged to develop indexing schemes with specific guidelines that clearly define how and which metadata must be captured, such as the source, format, character requirements, etc., for each field.

- OCR Technologies – Optical Character Recognition (OCR) is a power capability that can convert paper records to fully searchable text documents. However, agencies should be

aware that not all records can benefit from applying OCR technology depending on the type or quality of the source record. Because applying OCR can be costly, agencies should consider the possible limitations of their source materials and evaluate the usefulness of full text capability before investing in a large-scale OCR project. For some records, partial or zonal OCR may be more beneficial than full-page OCR. In other instances, utilizing indexed fields is a more cost-effective and useful approach to search and locate records.

- Vendor Support – There are many vendors in this space offering a wide range of services spanning all stages of a digitization project including initial strategy/planning, document preparation, imaging, and implementation of document management systems. Selecting a vendor requires agencies to thoroughly vet experience, security, capabilities and pricing. The following are sample questions to assist agencies in the selection process:

  1. Can the vendor support all stages of the digitization process from initial planning through implementation?
  2. Has the vendor adequately detailed the workflows within their solution?
  3. What portions of the vendor's solution are outsourced?
  4. What security and chain-of-custody controls does the vendor have in place regarding transportation of files and external storage devices?
  5. Does the vendor have a national footprint to service records located across the country?
  6. What chain-of-custody controls does the vendor implement to protect records throughout the digitization process, starting from pick-up/handling of physical records through scanning and eventual return to the agency or final disposition?
  7. Does the vendor offer document preparation, compliant storage, secured transportation and/or destruction services to appropriately handle the physical source materials?
  8. How does the vendor track and monitor records as they progress through the digitization process?
  9. What steps does the vendor take to ensure accuracy and quality assurance metrics meet the customer's needs?
  10. What scanning features/capabilities does the vendor provide (e.g., high resolution, OCR, wide-format document, book scanning, fragile/historical records)?
  11. Does the vendor provide digitization services for other record formats (e.g., microfilm, microfiche)?
  12. What IT compliance requirements or certification does the vendor meet?
  13. Can the vendor provide past performance references for successfully completed digitization project of similar size and scale?

Records Storage Risks
There has been progress developing key documentation and formal processes for records storage activities including cost data availability, issue escalation, performance reporting, and risk management, though gaps remain. However, there has not been tracking of the level of effort to map management and administrative cost to storage activities. The following are some risks for storing records:

- Lack of storage space
- Prone to damage
- Document vulnerability
- Inefficient document transportation

Risk Mitigation

Any new process can introduce risk, but this can be limited by developing options/actions to enhance opportunities and reduce threats. The following are some risk mitigation strategies:

- Documented policies, process, and standards in place
- Contracts: defining requirements, establishing expectations, including metrics and quality control
- Training for staff—those involved in digitization project and those doing appraisal of the original source record (e.g., eligible for digitization, should original be destroyed, etc.)

Digitization Challenges

- *Agency policies.* (examples of individual agency limitations)
  - Functional process owners in an agency may demand greater fidelity of the digitized file with the original hardcopy record, necessitating greater scanning quality (e.g., dots per inch, file format selection, etc.) and enhanced quality control measures (e.g., statistically sampling of digitized files to ensure 99.999% quality).
  - Other process owners may demand not only scanning of text but also inclusion of optical character recognition (OCR) with degree of accuracy to be determined.
  - Some process owners may require special tags to be placed on each scanned page (e.g., Universal Product Codes, bar codes, QR codes).
  - Some business procedures may have large quantities of legacy non-electronic records in non-traditional paper sizes (e.g., "E-size" printouts, Computer Aided Design hardcopies, architectural blueprints, microfilm, microfiche, Kodak/Polaroid/Fuji photos and negatives, X-rays, engineering aperture cards).
- *Agency funding.* (examples of individual agency limitations)
  Depending on the scope of the digitization effort, your project may span fiscal years. Work with your financial professionals on the best way to have funds allocated to your scanning project (e.g., Unfunded Requirement (UFR), budget POM, add-on to another funded project, "fallout money") while keeping in mind the "color of money" (use of government funds adhering to specific guidelines).
- *Section 508 requirements.*
  - Section 508 of the Rehabilitation Act of 1973 and Information and Communication Technology (ICT) Standards and Guidelines (Section 508, 36 CFR Part 1194)
    - E205.3 Agency Official Communication. Electronic content that is not public facing shall conform to the accessibility requirements specified in E205.4 when such content constitutes official business and is communicated by an agency through one or more of the following:
      - An emergency notification;

- An initial or final decision adjudicating an administrative claim or proceeding;
- An internal or external program or policy announcement;
- A notice of benefits, program eligibility, employment opportunity, or personnel action;
- A formal acknowledgement of receipt;
- A survey questionnaire;
- A template or form;
- Educational or training materials; or
- Intranet content designed as a Web page.

- *Durability.* A digitized record is subject to the same risks for any unstructured data file– risk of obsolescence of file format, application to read the scanned file, operating system wherein the application resides, storage media (Zip disks, computer tapes, DVD formats), and the devices to read the storage media (Zip drives, tape readers, DVD readers). Plans should be created for the long term accessibility of the scanned records and mitigation risks (e.g., convert scanned records to newer commercially accepted electronic formats 10 years in the future). Cloud computing environments tend to offer more robustness but also to be very dependent on sustained funding and vendor adherence to negotiated contract terms.

- *Classified records.* The same attention to detail and level effort is required for digitizing classified records as for digitizing unclassified records, including dedicated phases associated with the process - material movement, classification review, document preparation, initial digitizing, quality control, re-scans (if necessary), document reconstitution (if required), OCR, and media pre-preparation (if deliverables are required).

  More emphasis, awareness, and attention to detail should be placed on the classification review, especially in relation to identifying portion markings, classified subject lines, and codewords. From a cost standpoint, this increased emphasis/attention to detail may increase processing times and associated costs in labor.

  There is also the potential for increased costs in hardware or software depending on the setup of the system and requirements related to security controls associated with classified networks/workstations.

  Hardcopy storage is also a consideration as there are more physical security requirements and increased costs associated with the storage of classified hard copy material.

  Similar deliberations may need to be taken into consideration in digitizing non-electronic records with Controlled Unclassified Information (CUI).

Strategy

A cost-benefit analysis can provide empirical support for the business decision of when or when not to digitize is a primary driver behind a strategy. Factors that need to be accounted for within the analysis include but are not limited to:

- Labor
- Full Project Life Cycle Quality Assurance
- Hardware (both for digitizing as well as long term storage)
- Section 508 Accessibility (including pictures and graphics)
- Amount of Boxes/Pages to be Digitized
- Types of Analog Records to be Digitized (e.g., letter, legal, tabloid, maps, audio, etc.)

Closing

NARA policy changes are driving the need for digitization efforts, but agencies need to develop a strategy, compare costs and identify risks and opportunities to determine the best path forward to ensure success in digitizing records.