



THE INFORMATION MANAGEMENT DIGEST

Our Business, Is Keeping Your Business - Your Business!

A Service of WesTex Document, Inc.

October 2010

TO DIGITIZE OR NOT TO DIGITIZE

There has been talk of paperless offices for several decades, only to see paper use continue to grow each year. Recent trending indicates that a growing number of documents are “born digital” and remain in that form throughout their useful life. E-mail is just one of many examples of these types of digital records. What is less clear to some records and information managers are whether documents that are already created as paper documents should be converted to digital format.

MORE TO DIGITIZING THAN MEETS THE EYE

When some people think of digitizing, they tend to oversimplify the process. Because the intent of most conversion activities is to eliminate the original medium, (usually paper or film), there are substantial steps related to quality control that must be followed carefully. Otherwise, digital images may not be legible or may not be captured at all. In addition, care must be taken to ensure that any documents can be retrieved. This involves indexing and capturing metadata. The National Archives and Records Administration is engaged in a major conversion of documents contained in the archive so that they can be placed in the online catalog for easier access by researchers and the public. They provide the same cautions in describing the conversion process.

“For the purposes of this document, “digitizing” should be understood not just as the act of scanning an analog document into digital form, but as a series of activities that results in a digital copy being made available to end users via the Internet or other means for a sustained length of time. The activities include:

- Document identification and selection
- Document preparation (including preservation, access review and screening, locating, pulling, and refile)
- Basic descriptive and technical metadata collection sufficient to allow retrieval and management of the digital copies and to provide basic contextual infor-



mation for the user

- Digital conversion
- Quality control of digital copies and metadata
- Providing public access to the material via online delivery of reliable and authentic copies
- Providing online ordering for reproduction services at quality or quantities beyond the capacity of an end user
- Maintenance of digital copies and metadata”.

It should also be understood that the conversion costs related to digital information do not stop at the moment the capture process is complete. There are recurring costs related to server power, cooling and maintenance. Media and format refreshment and periodic migration of data related to operating system, software, media type and hardware. Again, NARA’s project document illustrates this, in part, by holding out digitizing as an example of conversion in order to preserve information on more fragile or outdated media types.

“We continually reformat at-risk archival materials so that they may continue to be used by the public. A paper document may become so fragile that we need to create a copy for public access; or, a video recording made in an outdated format, such as Betamax, must be transferred to a modern format that can be viewed on current equipment. As supplies of traditional analog

reformatting media diminish due to market forces, digitization is becoming a key activity in NARA's preservation reformatting strategy. We are in the process of adopting a digital workflow for preservation reformatting which will yield tremendous access opportunities as well. NARA commits to leveraging the work done to convert these materials by making them available online to users."

JUSTIFYING THE COSTS OF DIGITIZING

The cost of digitizing paper records can be substantial. According to one study by the City of New York records management department, imaging conversion projects for archival paper records took 16 years to break in the most optimistic scenario and in the most pessimistic scenario would not ever break even. This kind of financial investment must, therefore, be carefully considered prior to launching a project that may have very low or nonexistent return on investment to the organization.

What about preservation as a justification? Many manufacturers cannot produce sufficient test data to establish high confidence in any digital preservation medium for permanent records retention. Media types are simply too unstable at this time, or there has not been sufficient time to do thorough testing. In the example used above, the National Archives and Records administration to facilitate both preservation and expanded access. Because they are a national archive, NARA's preservation timelines are either very long or permanent. This puts them in a different category than most organizations that are preserving records for a more limited period. In cases where records may be damaged (by mold, vinegar syndrome in x-rays, etc.) digitizing to facilitate preservation would be a very good strategy.

Equally good would be selective digitization of vital records as a part of a program of protection through redundancy. In this scenario the original records would be retained but a digital copy would also be created and stored at a different location, such as in a media vault maintained by your information management vendor. This places a copy of the material outside the hands of potential saboteurs. A similar strategy has been employed by some archives using a method known as hybrid capture. In this type of approach information was captured as both a digital file for work use and also as a film image for permanent retention. Some of the most important drivers of digital conversion today are to facilitate frequent retrieval of archival information, improve workflow in the organization, or facilitate distributed sharing of documents across the enterprise.

FREQUENT RETRIEVAL

If there is only one copy of a document in archival storage, the document must be transported to you in some way: physically delivered, sent via fax, or scanned when requested (called scan on demand) and e-

mailed. Some types of archival information may be retrieved frequently. If this is the case then the costs of physical delivery or labor costs to pull and fax documents will increase to the point where digital conversion costs may be completely justified. The types of documents involved will vary widely by organization; working with your records and information management vendor to identify some of these types of documents may help reduce overall retrieval costs while improving operational efficiency.

SEARCH ENHANCEMENT

As Google™ has clearly demonstrated, the ability to search information quickly can help enhance productivity and improve workflow within an organization. According to Network World, "Butler Group, a London-based IT research and analysis organization, this week released a report titled 'Enterprise Search and Retrieval,' which concludes that 'ineffective search and discovery strategies are hampering business competitiveness, impairing service delivery and putting companies at risk.' Specifically, the research firm contends that as much as 10% of a company's salary costs is 'frittered away' as employees scramble to find adequate and accurate information to perform their overall jobs and complete assigned tasks." There are many other studies that have been mentioned online; all of them identify a percentage of employee hours wasted looking for information.

Labor is the most significant cost item for most organizations. If there is an identified problem with time wasted in searches for archival information, overall productivity may improve when information is digitized. A cost comparison of labor savings versus the cost of digitizing could make a digital conversion pay dividends to the organization.

ENTERPRISE USE

One of the most obvious reasons for digitizing information is to make it accessible to multiple persons in different locations at the same time. If there is a need to use archival information across the enterprise then this is clearly a circumstance where the business imperatives of the company far outweigh the costs of conversion.

CONCLUSION

Conversion from physical copies of archival records to digital copies is an expensive process that is not likely to return the investment for many years, if ever. However, there are certain business needs that may be more important than the costs of conversion. Being able to distribute information across the enterprise, protection by redundancy, enhanced search capability, frequency of retrieval, and conversion to a digital medium to preserve fragile or obsolete media types are all good reasons to consider an imaging conversion program.

PDF/A FOR LONG TERM DIGITAL STORAGE AND MICROFILM FOR LONG TERM ARCHIVAL STORAGE REDUNDANCY

At WesTex Document, we are providing services for some of our nuclear industry customers with retention rates for documents of 75 years or more. At the Nuclear Information and Records Management Association (NIRMA) annual conference this year we saw a lot of focus on long-term storage of digital records being discussed. Presentations by NARA and others highlighted and provided education on archiving formats with discussions circling around traditional archiving methods and their ability to provide reproducibility at some point in the future. We got a lot of angry feedback when I suggested that as record managers they may not be doing their jobs if all they are focusing on is electronic storage. Their own admission is that they must keep records a very, very long time.

Anyone that is my age will remember creating documents on Commodore 64's, playing 8-track music cassettes and may even have a few Beta Tapes of their favorite old movies around the house. Remember the ZIP drive? How about those 1950's records in the closet? As a federal agent back in the early 70's I had to store document evidence on microfilm. Case files were typed on an IBM typewriter. I can tell you today, that information from those devices then would be difficult to retrieve today.

TIFF guarantees reproducibility in the long-term and has an established structure. It is also easy to transmit in today's worldwide business environment. It is a raster format and must first be scanned with an OCR engine before it can be searched.

We have all become accustomed to the ease of making a portable data file or pdf. Its smaller file size is very advantageous for electronic file transfer. Meta data like title, author, creation date, keywords and such can easily be embedded for later search and retrieving. PDF files can be automatically classified based upon the metadata, without human intervention. It is today's defacto electronic image and universally accepted as such. It is what you see is what you get (WYSISYG).

The issue however is that PDF in its native form cannot guarantee long-term reproducibility and not even the "WYSISYG" principle. The ISO standard being accepted today for long-term storage of digital records is the PDF/A -1a format. It ensures the preservation of a document's logical structure and content text stream in natural reading order. The text extraction is especially important when the document must be displayed on a mobile device. It is the internationally accepted stan-



dard for long-term electronic archiving. The problem is no one has defined "long-term." Record managers must therefore insure they have a company/organization policy established for retention of information.

Microfilm, with its reported 500 year storage claim, remains the preferred long-term (more than 10 years) preservation reformatting strategy even as digital imaging has assumed a prominent role in enhancing access to such materials. What I suggest, is that a dual use of microfilm for preservation and PDF/A1a digital imaging for enhanced access for long-term digital retrieval is the proper solution for archiving information beyond 10 years retention. Although digital imaging can be used to enhance access, preservation goals will not be met until a microfilm or computer output microfilm (COM) recording of digital image files has been achieved as film is the best way to extend the life of information. You will always be able to put film on a light table with magnification and read it. Old film can be converted to digital; once converted you can put the digital back to film blipped for ease of retrieval now as COM and I have found that COM created from 600 dpi 1-bit can meet or exceed national microfilm standards for permanence and image quality. The quality of COM will principally be determined by the quality of the initial scanning. In some cases, a 600dpi grayscale will be needed to enhance quality, but 600 dpi bitonal scanning will produce digital files that faithfully render all textual information.

What is imperative, is that requirements be established based on your records and your organization's future plans for those records. Each organization must make its own decisions and implement them considering the importance of the information to be retained, its future use and the cost to get there. Utilize PDF for its ability to capture a lot of information and for its enhanced WYSISYG abilities. Use PDF/A1a for longer term digital archiving and use "scan to microfilm" for long-term archiving. If you are being faced with a conversion solution, then taking PDF or PDF/A to TIFF will need to be accomplished as TIFF still is the only image for COM.

LEGAL TECHNOLOGY CONFERENCE 2010

Nov. 6-7, 2010
Moscone West Convention Center, San Francisco

Wondering how to navigate information governance in the 21st Century? Look no further, LIT-Con 2010 is where legal and information technologies merge.

Mitigating risk and managing information to ultimately maximize revenue have never been more important. Two leading associations, ARMA International and the International Legal Technology Association (ILTA), come together to bring you the premier event on legal technology and information policy and governance – LIT-Con 2010. They bring together years of unparalleled resources and expertise that has not been offered ... until now! You won't want to miss this conference, which will address some of the most pressing concerns facing the legal and RIM professions. LIT-Con is also pre-approved for 10.0 ICRM maintenance credits.

Anyone whose role includes mitigating risk in a firm or legal department needs to take advantage of this exceptional event, including:

- Legal technologists
- Records managers
- Conflicts professionals
- Litigation support
- Legal administrators

LIT-Con is being held prior to ARMA International's 55th Annual Conference & Expo in San Francisco. LIT-Con attendees will have access to an Expo Hall packed with more than 200 exhibiting companies, education, activities, and networking opportunities with nearly 4,000 industry professionals!

<http://www.arma.org/LitCon/2010/SanFrancisco.aspx>

NATIONAL ARCHIVES AND RECORDS ADMINISTRATION RELEASES FEDERAL WEB 2.0 USE STUDY

Washington, DC...Have government agencies embraced new social media? The National Archives and Records Administration (NARA) shares the answer in its new study: A Report on Federal Web 2.0 Use and Record Value.

The report is online [<http://www.archives.gov/records-mgmt/resources/web2.0-use.pdf>]

The report explores how Federal agencies are using web 2.0 tools to create and share information. Tools examined include internal and external blogs, wikis, social networking, and other collaborative web-based technologies.

Archivist of the United States David S. Ferriero, a blogger and Tweeter himself, stressed the importance of such virtual outreach: "Social media and other web 2.0 tools are key aspects in furthering transparency and open government and through this study NARA provides a foundation for understanding and addressing the records management challenges these tools present."

Six Federal agencies were assessed on their use of and policies regarding web 2.0 tools. Representatives from an additional 19 Federal agencies participated in a focus group session.

The study concludes that based upon function and use, records created should continue to be appraised based upon business, evidential, informational, and contextual values.

Study recommendations include clarifying the statutory definition of a Federal record, addressing transfer requirements for permanent web 2.0 records into NARA's holdings, mitigating public expectations of content longevity, and integrating records management into agency social media policy.



WesTex Document, Inc.

815 S. Gilbert Drive
Reese Science & Technology Center
Lubbock, TX 79416

CALL (806) 885-2906 or visit our web <http://www.westexdocument.com>

ARCHIVE - IMAGING - DESTRUCTION SERVICES

West Texas's most trusted source for meeting your record management needs. We offer professional storage, digital imaging and secure destruction of any type of records you may have. We can shred material to three different particle sizes - no strip cutting used; we can send two mobile trucks that will destroy up to 7 TON an hour between them; we can image and copy (printing) documents up to 24 x 48"; professional record archive services and consulting for all of your professional RIM needs. Call or come visit us to learn more and tell us how we might help you.

Lubbock Chamber Member Since 1997

ISO 17799:2005 (Information Security) Compliant

OUR BUSINESS, IS KEEPING YOUR BUSINESS - YOUR BUSINESS!

