



Helping You Work
Better, Faster, Smarter

IMERGE Consulting InSights™

April 2005

Welcome,

If you haven't visited the "Articles" page of www.imergeconsult.com in awhile, you're in for a treat! Several articles are new this quarter.

We feature dozens of industry-related articles, white papers, industry reports, books and Webcasts all written, edited and presented by IMERGE analysts.

Topics include electronic records, document management, workflow, content management, portals, customer relationship management, electronic signatures and insurance claims payment and processing.

The IMERGE Consulting website has long been a favorite in the industry for information and advice. Visit it today, and see how we can help you do business **better, faster, smarter**.

Editor's Note:

Regardless of your industry, your company, your own individual role and where you are in your career, IMERGE Consulting has many opportunities that can help you do business **better, faster, smarter!** Here are just a few:

- (1.) **The AIIM On Demand Conference & Exposition May 17-19, 2005 in Philadelphia.** <http://www.aiimondemand.com>. The AIIM On Demand Conference & Exposition is the largest enterprise content and document management event, and has been an event for over 50 years. You'll want to attend if you're a corporate executive, IT professional, consultant, or system integrator seeking technologies and solutions to develop or improve your enterprise IT, content, and document management business processes.
- (2.) On May 16, 2005 IMERGE is conducting a **pre-AIIM workshop for professionals pursuing the Information Capture Professional (ICP) certification**. The registration fee is just \$495. If you're planning to go to AIIM, this workshop couldn't be more convenient! For more information, [CLICK HERE](#) or call 888-824-3004. For more information on ICP certification, visit <http://www.tawpi.org/icp.html>
- (3.) **IMERGE Partners, Principals, and Associated Consultants are well represented at the upcoming AIIM conference.** Partner **Don Post** is Chair of the 2005 AIIM Conference, and several persons from IMERGE are featured speakers! *Again, IMERGE has more speakers at this event than any other single organization!* These accomplishments are additional testaments to the high-caliber professional that IMERGE is proud to have on its own roster.

The IMERGE speakers and their sessions are:

- **Partner Arthur Gingrande's** topic is, "To Conform Is the Norm: The Latest Advances in Forms Processing" (Wednesday, May 18).
- **Partner Jim Minihan** presents, "ECM 101: Workflow" (Tuesday, May 17), which is part of AIIM's ECM Practitioner Level 1 series.
- **Principal Gary Gershon** focuses on a case study: "Putting WebDAV, Web Services & Rich Clients to Work" (Wednesday, May 18).
- **Principal Jesse Wilkins** is speaking at the Messaging On Demand Conference on May 18.

His topic in the Technology Basics Track is, "Email 101: The Basics of Email Management."

- **Principal Vigi Gurushanta** speaks on, "Don't You Trust Me? Demonstrating the Integrity, Reliability & Trustworthiness of Your System" (Wednesday, May 18).
- **Ronald Kral**, an IMERGE Associated Consultant, speaks on, "An Exercise in Self-Control: IT & Auditing Mitigate Risk & Create Business Value" (Thursday, May 19).
- IMERGE Associated Consultant **Robert Schleyer** speaks on, "Demystifying Change Management" (Wednesday, May 18).

- (4.) **IMERGE is featured in AIIM's Advice Center.** Professional Members can access industry experts for free one-to-one advice on one of a number of topics (e.g., workflow/process management, records management, email management, outsourced services, and many more). To learn more, [CLICK HERE](#)
- (5.) IMERGE Partner Jim Just is a featured speaker at the Electronic Document and Records Management conference (April 26-28), sponsored by International Quality & Productivity Center (IQPC). He is leading the interactive evening workshop on, "Common Pitfalls To Avoid when Implementing an Electronic Document and Records Management System." For more information, [CLICK HERE](#).

Partner Arthur Gingrande Moderates TAWPI Webinar

IMERGE Partner Arthur Gringrade recently moderated a TAWPI Webinar, "Auto-Classifying the Toughest Documents: Invoices, Mortgage Applications, and Mail. The webinar's purpose was to teach end users what they need to know about processing invoices, mortgage applications and mail. Industry experts presented real-life case studies that spell out the benefits and illustrate the R.O.I. that can be achieved by fully optimizing these auto-classification software applications. TAWPI is The Association for Work Process Improvement (www.tawpi.org).

IMERGE Welcomes Pierre Heymann as Principal

Pierre Heymann recently joined IMERGE as Principal. Pierre has 20 years of experience in document production with print vendors and output management companies. His deep understanding of processes and technologies for the creation, production, distribution, archive and disposal of documents. His expertise covers software, hardware equipment, papers and supplies, policies, procedures, compliance and best practices. Pierre has played a key role in the architecture and design of many of the largest Automated Document Factories in the US. Pierre is very active at the regional and global XPLOR and at the Print On Demand conferences where he speaks regularly for the past 10. For more about Pierre, [CLICK HERE](#).

Industry Quick Notes

TAWPI is working on an update to the Information Capture Professional (ICP) exam. The ICP Blue Ribbon Commission has been working on updating the ICP Handbook, the Glossary of Terms, and the ICP exam questions for several months. The updated materials should be available by the end of April. The next ICP exam cycle is scheduled for July 23-30, 2005 and will include new questions based on the updated Handbook. For more information visit <http://www.tawpi.org/icp.html>.

Xerox has announced it will produce storage hardware that uses the UDO specification. The company will deliver four new products: the XUDO-100 single media drive, the XUDO-244 storage library, with two drives and 24 media slots, providing up to 720 GB of storage, and Xerox Write Once and Rewritable media. The products should be available in April 2005. For more information [CLICK HERE](#).

Getting Involved in the Industry

Professional associations are a great way to network with others who have had similar problems, to get a fresh perspective, or simply to keep abreast of what is going on in the industry. Associations also offer a way to get more involved in the industry – whether as a chapter leader or speaker or simply through the educational opportunities they offer.

There are thousands of associations in the U.S. Here are some that are directly relevant to the document and content management industry.

AIIM – The Enterprise Content Management Association. AIIM covers the technologies used to capture, manage, store, preserve, and deliver information. Chapters meet monthly September – June and are in most major cities in the U.S. plus Toronto. The AIIM 2005 Conference and Exposition will be May 17-19 in Philadelphia. <http://www.aiim.org>

ARMA – The Association of Records Managers and Administrators. ARMA members focus on managing records, regardless of media. ARMA chapters are throughout the U.S. and Canada, and meet monthly September – June. The 50th Annual ARMA Conference and Exposition will be September 18-21 in Chicago. <http://www.arma.org>

TAWPI – The Association for Work Process Improvement. TAWPI focuses on mail, remittance, document, and forms processing strategies and technologies. There are eight chapters in the U.S., and chapters meet quarterly. The TAWPI 2005 Forum and Exposition will be July 24-27 in Baltimore, MD. <http://www.tawpi.org>

Xplor – The Electronic Document Association. Xplor is focused on the users and suppliers of products and services used to create, modify, and deliver information through a variety of document technologies and strategies. Xplor has chapters throughout the U.S. and world-wide. The Xplor Global Conference and Exhibit will be next February 1-5 in Miami Beach, FL. <http://www.xplor.org>

"Migrating to Better Media"

Jesse Wilkins

jesse.wilkins@imergeconsult.com

As Appeared in AIIM *e-DOC Magazine*, Jan/Feb 2005. Reprinted with permission

Most of the information we generate is stored in digital format for two good reasons: first, because it starts out that way, and, second, so we can find it when we need it. According to a study conducted by the University of California at Berkeley, the world produced 5 exabytes of information in 2002, or more than 800 megabytes for every man, woman, and child on the planet. Up to 94% of that information is created digitally and stored on magnetic hard disks. If this information were printed and stored, the resulting stack of paper would stretch around the earth more than 1,400 times – and the rate of growth is continuing to increase.

However, digital information suffers from a key limitation: longevity. Information printed on paper is expected to last decades, at a minimum; archives suggest a life expectancy for paper documents of more than 500 years using special paper and inks, and up to 500 years for documents stored on archival-quality microfilm.

In contrast, the information stored digitally may be unreadable in as little as a couple of years. The first problem is that of storage media stability. It is not uncommon for CDs and DVDs to become scratched or the readable layer corroded and thus be rendered unreadable. Tapes stretch; hard disks get dropped or infected with viruses.

The second problem is finding hardware that can read older media. Many organizations still have information stored on media for which they have no hardware, either because of technological advances or because they selected proprietary media which have since been discontinued. Archives and records management offices still have information stored on 5.25" floppy disks, ZIP drives, and even 8" floppy disks that can't be read because there are no working drives available.

Even if the media survives, and hardware can be found to read them, the organization may not have software available to access the information on the media. Proprietary file formats come and go, and if you choose the wrong one, your information may not be accessible in the future.

So how do we manage this fragile information so that we can get at it in five, or fifty, or five hundred years? There are several approaches available. For example, we can try to replicate all the different computing environments, including hardware, software, drivers, etc. and preserve them indefinitely. But the sheer number of systems required to accommodate the different combinations, and the resulting centralization of access, renders this solution impractical.

Emulation is another option – write software to mimic the operating system, hardware, software, and drivers. But emulators that run on Windows 2000 to emulate DOS or VIC-20 are unlikely to run on Windows 2050, resulting in a requirement either to rewrite the emulators for every generation of operating system, or write emulators for the emulators ad nauseum. Neither of these choices is particularly appealing.

The solution that has the best track record to date is migration, wherein the information to be retained is periodically copied to new media and/or new formats. Migration has a number of advantages that make it both a good short-term prospect and offers benefits over the longer term. This article will take a more in-depth look at what migration offers to organizations concerned about long-term access to electronic information.

Choosing What to Migrate

The first step in migration is to determine what to migrate. The obvious choice is to migrate documents in their native format. Insuring that information remains trustworthy and reliable drives migration. By taking the information in its native format and periodically recopying it, the underlying information structures of the individual documents remain the same.

It is also important to consider how the migration process will affect the provenance of the information. For example, during a migration changes are often introduced to the bitstream, or the pattern of the stored bits, on the storage media. This can lead to questions about the validity of the migration process, particularly as organizations verify the results of the process.

This issue becomes more difficult with the complex, dynamic documents that are increasingly found in organizations. Complex documents with hyperlinks or embedded information such as multimedia are more difficult to migrate and particular care must be taken to ensure that those link relationships are preserved. If this is not possible, the conversion process must document what information is not being preserved and why. This has particularly been an issue with email messages containing attachments and web pages. While there are applications that manage attachments and compound documents, it can still be problematic to migrate the information from its original storage media to new or different media.

The migration process must also take into account metadata. Ideally the migration will not result in changes to the metadata of the individual documents, such as file size, file date, etc. If this is unavoidable, it must be documented for each migration to mitigate the effect on the provenance of the documents.

Approaches to Migration

There are several different approaches to migration that organizations should consider. The first, media refreshing, involves periodically recopying information on tape or optical disk to a new tape or disk. The individual documents are not changed, and often the entire bitstream stays the same as the entire contents of the volume are copied bit by bit. This is the least invasive approach but can be difficult to justify due to the rapid rate of technological innovation.

The next approach is to copy information from older storage technologies to newer ones – for example, from CD-R to DVD-R. Like refreshing, this approach leaves the original documents unchanged but allows the organization to take advantage of improvements to storage technologies. The bitstream often changes during a migration operation, for a couple of reasons. DVDs hold about seven times more information than a CD, so it makes sense to consolidate the information from the CDs onto fewer DVD volumes, with the result that the bitstream is different. Different storage media may actually store things slightly differently as well, so information on tape would have a different stored bitstream than the same information stored on DVD.

Both of these approaches assume that the documents themselves remain unchanged. But as noted above, software changes over time as well, and organizations may need to convert the documents from older, proprietary technologies to newer ones. For example, an organization may have documents it is required to keep for twenty years stored as WordStar or GeoWrite, two formats that are increasingly difficult to support. The documents may be stored in various different versions of Microsoft Word, each of which provides different capabilities. As these documents are converted from older file formats to newer, the underlying bitstream changes.

One of the considerations for the organization is to determine whether to continue to use proprietary formats or to move to a more standard one for long-term preservation and ease of migration in the future. This changes the bitstream, of course, but more importantly can lead to a significant loss of functionality. For example, an organization that is storing its financial documents as Microsoft Excel spreadsheets may consider translating the documents to a more standard format, such as ASCII text. In this case, the resulting ASCII file would not store the formulas used to calculate the values, but only the numbers themselves. More recently, XML is becoming a promising option for translation because it has the flexibility to include not only the numbers but the formulas and even to encapsulate the metadata directly into the document. The use of formal or market standards such as TIFF, PDF, and XML can significantly ease the burden of migration from a software perspective.

Once the migration process is complete, the information should be validated to ensure that the documents were not changed (refresh or migration) or, if they were (translation or conversion), that the content contained in the original documents survived the process. For the former, the organization can do a bit-level comparison and verify that the bitstream was unchanged. For the latter, the process is more involved and will require that the organization make some decisions about tradeoffs and then document those decisions.

Longer-term, the organization periodically needs to review the storage media and the information stored on the media to ensure that it remains accessible. The US National Archives recommends that organizations review their electronic information at five years. With the current rate of technological innovation, an organization that waits ten years to review its digital documents may have waited too long and incur significant expense to regain access to the information.

Migration is not a cheap solution. Organizations that pursue migration as a long-term strategy for digital preservation must understand that it will have to be done repeatedly over the lifetime of the documents and storage systems. Money has to be budgeted for it as a recurring expense – but if it is done regularly it is less difficult to ensure access to the organization's information assets. Migration can be an effective strategy for ensuring access to digital information over time.

Jesse Wilkins, CDIA+, LIT, EDP, ICP is a principal with IMERGE Consulting (<http://www.imergeconsult.com>), a leading unbiased technology and management consulting firm. A frequent industry speaker, Jesse is involved in both AIIM and ARMA chapter activities and is a member of the 2005 AIIM Conference planning team.

Coming Events, April 2005

Date: 4/19/2005

Event: AIIM Content Management Solutions Seminar, Denver
Program details: Free educational forum on ECM technologies.
More information: [CLICK HERE](#)

Date: 4/20/2005

Event: ARMA Baton Rouge/Lafayette Spring Seminar
Program details: Jesse Wilkins will speak on "Managing Electronic Records"
More information: [CLICK HERE](#)

Date: 4/21/2005

Event: AIIM Content Management Solutions Seminar, Dallas
Program details: Free educational forum on ECM technologies.
More information: [CLICK HERE](#)

Date: 4/25-27/2005

Event: The InterDoc Symposium
Program details: The InterDoc Symposium will help you to better understand how to establish information architecture in support of your Enterprise Content Management implementation. Jesse Wilkins speaking on active

email management. More information: [CLICK HERE](#)

Date: 4/26-28/2005

Event: ARMA Houston Spring Seminar

Program details: Three days of records and information management tools & techniques. Jesse Wilkins will speak on "Accessing Your Electronic Records in 5, 50, 500 Years".

More information: [CLICK HERE](#)

Date: 4/27-28/2005

Event: IPQC Conference on Electronic Documents and Records Management

Program details: Two-day conference on improving business processes using document technologies. Jim Just presenting workshop on "Common Pitfalls to Avoid When Implementing EDRM Systems".

More information: [CLICK HERE](#)

Date: 4/26-29/2005

Event: Society for Imaging Science & Technology (IS&T) 2005

Program details: Conference on electronic archival and preservation.

More information: [CLICK HERE](#)

New publications, Jan-March 2005

"Automated Forms Processing", Arthur Gingrande, AIIM e-DOC Magazine, Jan/Feb 2005, [CLICK HERE](#)

"Fundamentals of Collaboration Systems", Bernard Chester, AIIM e-DOC Magazine, Jan/Feb 2005 [CLICK HERE](#)

"Non-Records Management: What's Your Excuse?" Jesse Wilkins, AIIM e-DOC Magazine Online, [CLICK HERE](#)

"Migrating to Better Media", Jesse Wilkins, AIIM e-DOC Magazine, Jan/Feb 2005, [CLICK HERE](#)

"The First Step: Inventory Your Electronic Records", Jesse Wilkins, AIIM e-DOC Magazine Online, [CLICK HERE](#)

To **unsubscribe** email lynnea.brown@imergeconsult.com with "Unsubscribe" in the subject line. If a friend or colleague is forwarding "IMERGE InSights™" to you, subscribe to receive your own free copy: email lynnea.brown@imergeconsult.com with "Newsletter" in the subject line.