ZyLAB®
The Paper Filing Company

E-mail Archiving: A Vaccination Against Regulatory and Legal Distress

White Paper

Priscilla Emery
November 2003

President
e-Nterprise Advisors

E-mail: The New "Smoking Gun"

A recent study by the ***META Group*** points out that e-mail has become the preferred communication tool for business executives. This pervasive business tool has created new challenges for IT network managers, records managers and legal departments in many organizations. In addition to the impact that the increase in e-mail volume has had on storage requirements for e-mail administrators (you know it's growing – just look at your own in-box), the ways in which e-mail have been used by the business community are causing organizations to view "managing" e-mail as more than just an exercise in conserving storage resources.

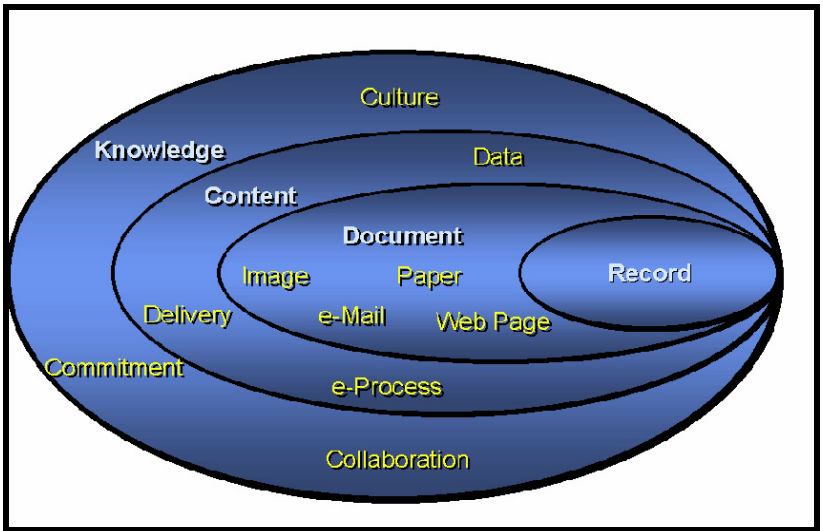


Figure 1: e-Nterprise Repository Source: e-Nterprise Advisors

One first has to look at the position that e-mail is taking in the overall enterprise repository of information to better understand its impact (see Figure 1).

If you look at where e-mail sits in this diagram it fits squarely in the Document space and it has become a very crucial document, not just in terms of transferring information from one person to another, but sometimes serving as a historical record of a business transaction. You can see that not all documents or data elements are records but records can take the form of any number of different delivery and format mechanisms, e-mail being one of them.

This mix of potential delivery mechanisms coupled with new regulatory issues associated with retaining corporate information is causing a widespread interest in how to effectively manage not just e-documents but e-mail in particular.

Headline grabbing events such as the shredding of evidence at Enron; the e-mail trail of conflict of interest at Merrill Lynch, Citicorp's Smith Barney and other financial investment firms; and even the U.S. Department of Justice vs. Microsoft case have proven just how crucial e-mails can be in the document trail of evidence. As a result of those events government regulators have become significantly tougher about making sure all communications that occur during the course of providing financial data to shareholders and other public entities are reproducible in an audit or investigation. E-mail has become the "smoking gun" in many of these investigations and has also become a "lightning rod" of attention for auditors and lawyers alike.

Although the Sarbanes-Oxley Act of 2002 outlines a variety of different activities that now must be supported by public companies, such as having truly independent audit committees, the bottom line of this act is that all of these communications need to be archived for a specified length of time and immediately recoverable and reproducible on request by shareholders, audi-

tors or regulatory agencies. The underlying assumption is that if the document or e-mail cannot be produced when requested there must be an intent to defraud or to circumvent the system, which consequently involves some serious financial penalties and potential jail time.

Whoever thought that deleting e-mails could be a criminal offense? Well, now it can happen. Government mandates such as HIPAA (Health Insurance Portability and Accountability Act) also involve the potential auditability of e-mails for all organizations (not just healthcare companies) since health information could potentially be exchanged between employees, managers and Human Resource departments.

Still, there have always been negative consequences for using e-mail to participate in what may be viewed as unethical or illegal conduct — that hasn't changed. Several financial services companies have been slapped with significant fines over the last several years for conflict of interest activities that have come to light through the investigation and eventual recovery of e-mail-based documents. E-mail has been heavily used as evidence in the recent Credit Suisse First Boston investigations, where Mr. Frank Quattrone has been charged with allegedly telling his employees to delete e-mails and other potential evidence just prior to a government investigation.

E-mail is a prime target of evidence in many litigation activities these days and that activity continues to increase. Sexual harassment claims and suits subject to potential private litigation rely even more on e-mail-based evidence as well. And with the Sarbanes-Oxley act providing an outlet for shareholder grievances the rate of shareholder lawsuits to recover lost investments will probably not reduce this trend.

The cost of attempting to deal with these types of lawsuits can be astronomical (whether or not the company is at fault). For example, American Home Product's Wyeth-Ayerst Pharmaceutical Division became the subject of a lawsuit in Massachusetts related to its "Fen-Phen" diet drug. It was claimed that Wyeth-Ayerst had known that Fen-Phen produced some serious contra-indications for some patients but had failed to disclose this knowledge to physicians and the public.

The deposition process required that the Company recover the e-mail of about 15 employees from more than 800 back-up tapes. The defendants estimated the cost of restoring the tapes for electronic discovery would cost anywhere from $1.1 million to $1.7 million. Wyeth-Ayerst chose to settle out of court so it may not have had to incur those particular costs but guilty or not, it remains an expensive proposition to defend any company when e-mail-based documents are required as part of the deposition process.

Why Is E-mail So Troublesome?

It doesn't help that e-mail can be a "troublesome" document and/or record to actually manage. E-mail systems allow messages to be changed before forwarding so that the "original" is not what it seems to be to the receiver making it imperative that e-mail records be "locked" before being re-transmitted. Annotated items and comments can also be deleted that may be crucial to the record of a transaction. E-mail systems also allow for indirect addressing through distribu-

tion lists, and blind copies, making it difficult to track whether or not someone who is NOT really supposed to have access to the information is being included. E-mail messages can contain embedded links that can take a reader to an information item that exists (or at least used to exist) elsewhere.

And the most troubling problem — most users keep hundreds of old messages on their desktops and on their personal folders on e-mail servers. As the number of daily messages grows, and those messages increase in size, organizations may see a growth in storage overhead of 100% to 150%. Storage of these messages is having an impact on overall e-mail server requirements. E-mail administrators are constantly asking e-mail users to delete unwanted or unnecessary messages so that e-mail servers can operate efficiently. Unfortunately, this request can be counter-intuitive to the notion that e-mail should be saved as a record. The challenge for many organizations is to keep e-mail servers optimized for peak performance while at the same time making sure that the right e-mails are being archived (and/or deleted) at the appropriate time.

Mitigating Risk By Managing E-mail Assertively

Given all these potential threats, what can IT managers to avoid the consequences of non-compliance or minimize the costs associated with potential litigation. Well, the one thing they can't do is nothing. Of course, sometimes doing "something" isn't really effective enough, such as doing daily back-ups of e-mail servers at the end of the day. A back-up file only provides a snapshot of what is still left in the e-mail server at the end of the day. A lot of e-mail that should have been archived for regulatory reasons could have been deleted during the course of the day. That e-mail will not show up on a back-up tape. And, even if it did end up on a back-up file, finding these un-indexed e-mails several months or years later would be very difficult. As a consequence, organizations should evaluate e-mail storage and archival alternatives to address this issue.

Before evaluating e-mail archiving alternatives it is very important to do some serious internal planning and answer some key questions that will impact your implementation approach.

Understanding what your organization is trying to accomplish from a compliance standpoint goes a long way to understanding what types of internal procedures need to be developed and what tools need to be evaluated. For example, compliance with Sarbanes-Oxley is really focused on the use of archiving tools while compliance with HIPAA may be focused on privacy tools along with archiving tools.

In all cases it is mandatory that appropriate policies and procedures be in place first. Most archiving tools only help to enforce or manage approaches already in place. For example, a file plan with record categories defined should already be in place before an organization can effectively use an e-mail archive product for e-mail records management. No automated system (even ones that automatically index e-mails and records) can produce effective results if categories have not been defined prior to implementation.

That said, a secondary step to creating a policy is enforcing it and enforcing it consistently. Organizations that do not consistently enforce record keeping policies are subject to the same legal

liability as those that don't have any policies at all. In the case of enforcing record keeping standards, this involves a combination of training at appropriate levels of the organization and timely quality assurance checks on record keeping practices. In the case of the misuse of e-mail (such as internal sexual harassment), perpetrators have to be actively admonished or expelled, as outlined in any internal policy, in a consistent way (i.e., the organization cannot admonish one person but ignore someone else) or be subject to a potential lawsuit. Again this type of enforcement also involves providing training and the appropriate compliance checks.

Many other issues still need to be sorted out when evaluating e-mail management tools and services.

- What part of the e-mail is being scanned for content, viruses, etc.? The Header? The body of the Message? The Attachments? All of the preceding? When it comes to archiving it may be sufficient to scan the Header if you use only standard headers for certain types of messages that have to be stored. This scenario is highly unlikely but every organization needs to identify what types of e-mails need to be archived and figure out how to identify them as easily as possible.

- How intrusive is the product you plan to use? Not just to users but to the e-mail administrator as well. It should go without saying that adding new fields for users to classify and file e-mail, as either a record or other repository-based document, is additional work that most users will not welcome. Finding a system that aids in pre-population of file plans, and providing a familiar user interface (i.e. Microsoft Outlook, Lotus Notes, Internet Browser) will go a long way to making sure the product gets used appropriately. E-mail administrators already feel overworked and sufficiently challenged on a daily basis so that any new "management" tool should not add a significant amount of "overhead" to the e-mail server's storage needs and provide value-added reporting and tracking capabilities for the administrator.

- Is auto-categorization or automatic indexing an option you want to consider? For knowledge gathering applications, auto-categorization can be a useful tool to automatically file and categorize e-mails based on the content. These applications can be a little more fluid than more rigorous archival applications and although consistency is important, it is not as mandatory as in a record keeping system. Record keeping systems can also use auto-categorization but it is recommended that a significant quality assurance testing and implementation effort be completed before rolling out the capability *en masse*. Auto-categorization can enhance compliance efforts by making it easier and less time consuming for users to file all kinds of records including e-mails.

- How much customization is required to implement the product in your environment? As we all know, customization can run into a significant sum of money, especially if the product is completely incompatible with the normal procedures used in an organization. In addition, integration with legacy systems must be taken into account.

- Is it easy to access and retrieve archived e-mails when required? Ease of access is a two-way street. On the one hand, people who have authorization to access information should

not have to wade through confusing interfaces and several layers of software to get to the e-mails needed. On the other hand, there should be sufficient security on the system to limit access to only those who should have it. Another question related to this one is, what software do users need on their desktop for retrieval of archived e-mails? This could be asked of either software or service providers. If plug-ins are required on the desktop how will they be proliferated and how much time does it normally take for them to "boot-up" when needed?

- Can the tool handle all your e-mail servers? Many large companies have multiple e-mail products supported in-house and multiple e-mail server nodes as well. Many e-mail archiving products support a variety of different e-mail systems but it is important to note whether or not they can support different systems at the same time. And, even if only one type of e-

(Paper continues on page 7)

ZyLAB's Approach to Archiving E-mails

With the ZyImage XML Plug-ins for Microsoft Outlook, a user can move their e-mail messages (including attachments) from the proprietary MS Outlook or Exchange data repository into an open source XML format.

ZyLAB converts an e-mail message to one XML file (holding e-mail structure, address information and plain body text), one RTF file with the formatted body text (if present) and one XML wrapper plus the original file format for every attachment. (See Figure 2). The screenshot in Figure 3 illustrates how this process looks to the user.

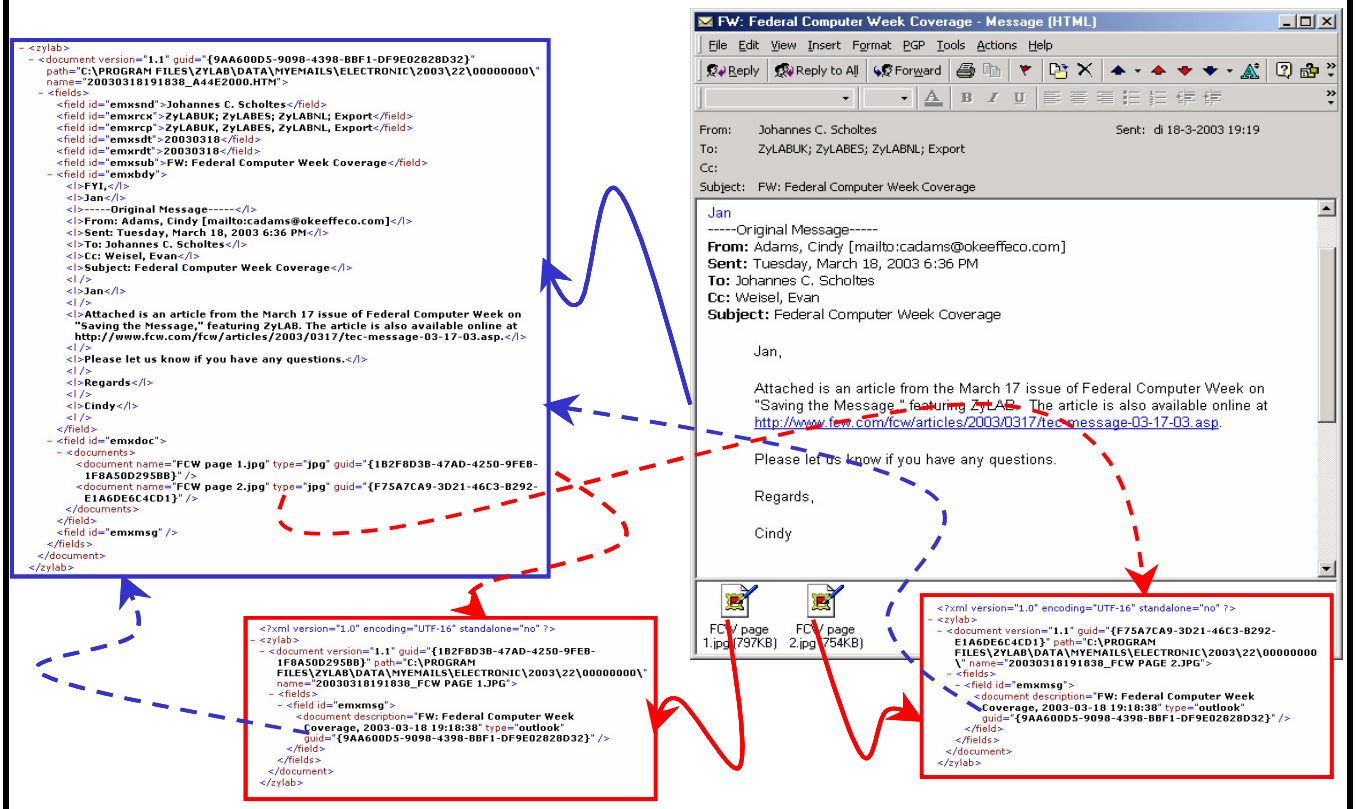


Figure 2: Converting MS Outlook e-mail to XML format

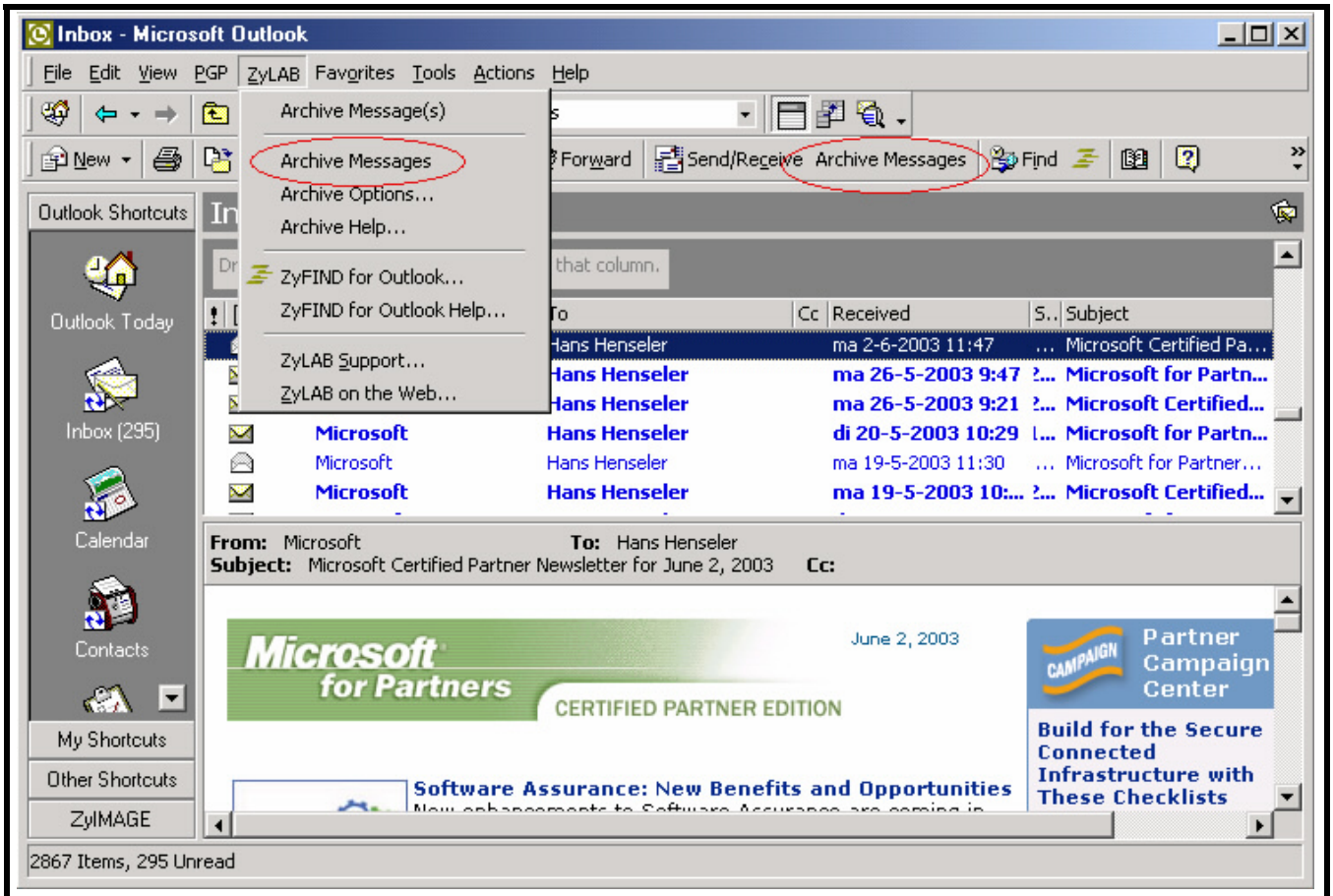


Figure 3: Archiving Plug-in for Microsoft Outlook

After they are archived the e-mails and the attachments can be searched with either ZyFIND or ZyIMAGE Web-server.

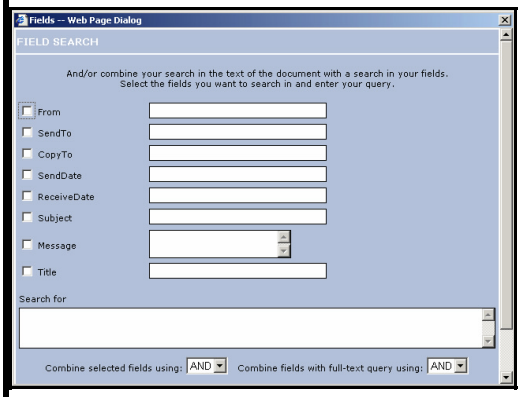


Figure 4

A user can open a message to access an attachment or open an attachment to access the associated message (see Figure 6). Archived e-mails can be marked by adding some text to the SUBJECT line such as “archived on <date and time> in

The screenshots to the right and below illustrate how e-mails can be accessed through ZyLAB’s web-based full text search capability. Using key field search (Figure 4) e-mails and their associated attachments can be easily retrieved (Figure 5).

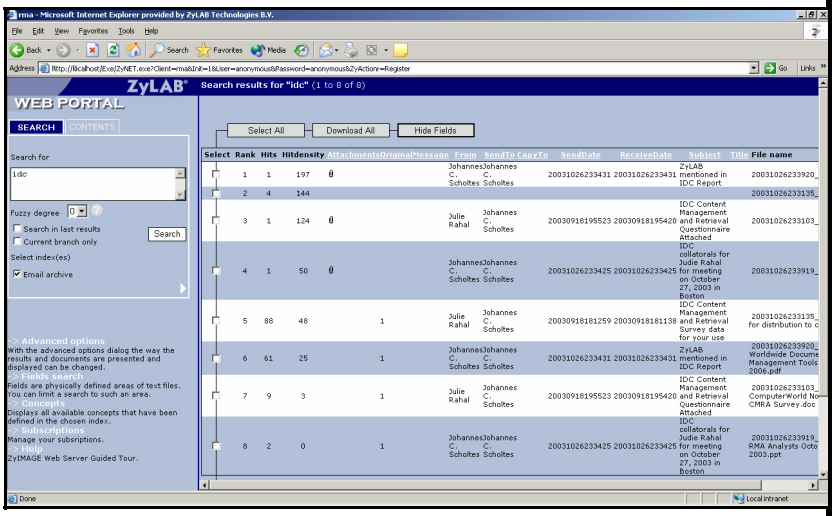
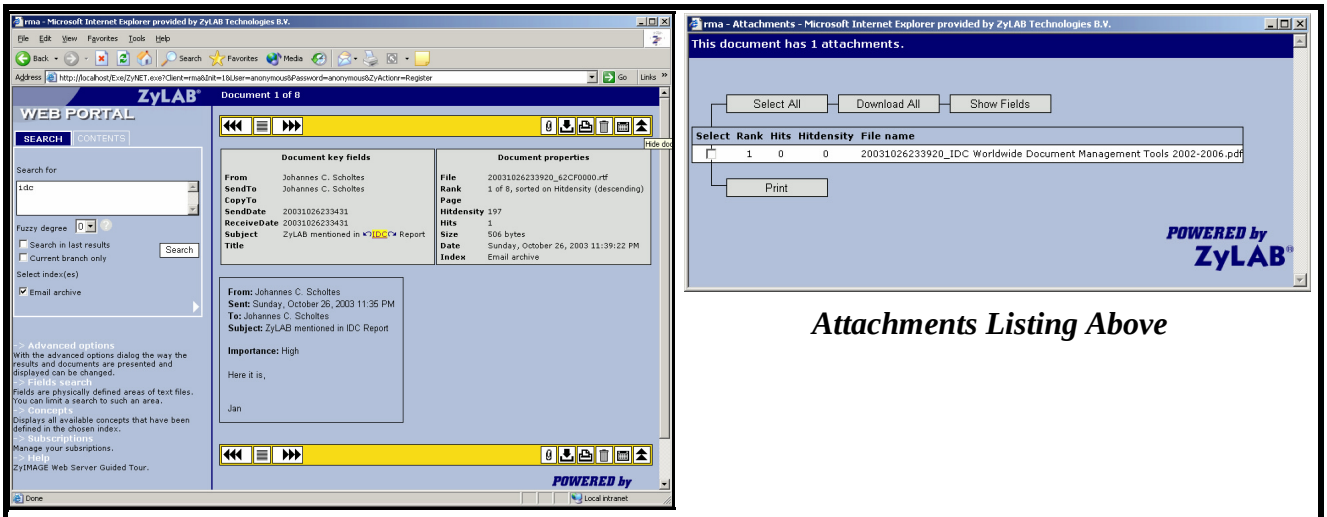


Figure 5: Search Results



Attachments Listing Above

Figure 6: Linking between Message and Attachment Can be Done in Either Direction

archive <x>." Simplicity is the key element here. In addition to being fully text searchable, records management functionality can be incorporated into the archive allowing for the storage and deletion of e-mails based on defined file plans. Policies can be enforced by using the ZyIMAGE DoD 5015.2 module. ZyLAB's e-mail archiving capabilities are already being used by several companies, located worldwide, for compliance with Sarbanes-Oxley regulations.

(Continued from page 5)

mail application is used can the product or service handle traffic from multiple e-mail servers?

- What type of auditing and logging facilities are provided? This question should be addressed by all applications. How much overhead do they add to the system or to the service?

- Then there is the long-term plan of how to view these e-mails (and other stored records and attachments) after they've been archived over a long period of time. Will you still be able to read these e-mails and attachments after seven years (a virtual eternity in technology years)? Are the e-mails and attachments saved in a proprietary format? Is there an option or do you have a plan to copy e-mails to a non-revisable media such as optical disc for long-term preservation? Having a defined plan for either migration or long-term viewing is essential.

- If you are archiving for record keeping purposes is the product DoD 5015.2 certified? Certification doesn't guarantee that the product has all the records management functionality that you may require but it does help to know that the baseline requirements for handling metadata are supported correctly. In addition, you should also be constantly aware of changes to the "standard" that may impact implementation efforts in the future.

If managing e-mail as a record is your primary objective some other functions and features must also be taken into consideration. The authenticity of the e-mail record has to be maintained and the e-mail has to be "unalterable" from creation to its final disposition.

In order for an e-mail record to be deemed as "usable evidence" it needs be prepared to be subject to legal scrutiny to overcome any legal objections during any potential court or discovery process. The following activities need to be supported:

- The capture of incoming and outgoing e-mail messages at time of creation or receipt
- Retention rules should be applied systematically
- Consistent application of a file plan with policies and retention schedules.

Record integrity also depends on three attributes: content, context, and structure. Therefore, moving a complete e-mail record and its attachments, optimally in native document format, out of mail servers may change content, context or structure or loss of e-mail metadata. Retention information should be integrated into message stores and/or repositories.

Managing E-mail: More Than Just a Regulatory Issue

The knowledge sharing aspects of managing e-mail should not be overlooked. An archive can also serve the purpose of making e-mails a part of overall corporate memory, and providing a way of accessing messages and attachments so that they can be found when needed.

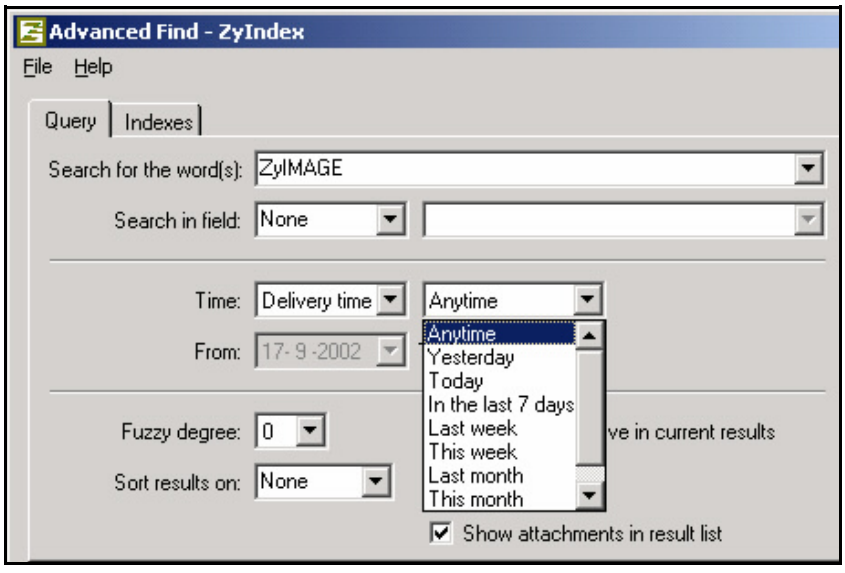
According to a KPMG survey conducted in 2000, 60% of employees spend more than one hour per day duplicating the work of other employees. If the average professional employee costs an organization at least $100,000 a year, employees spend roughly $12,500 duplicating work each year. That's a lot of productivity and money going down the drain.

ZyLAB addresses this issue within its product suite. It can index Outlook .pst files (including attachments) using ZyFIND for Outlook. Although a search option is already present as a feature of Outlook itself, searching with this option is very slow. Dr. Johannes Scholtes, president of ZyLAB points out, "ZyFIND's raw search power saves me lots of time every day. No structure—just save all e-mails in a .pst or .ost file and search the repository. I have seven gigabytes of e-mail on my laptop—all fully searchable." The tool is also useful for forensic investigations, corporate e-mail analysis and other knowledge-based applications.

A user can also search for information based on the meta information of the e-mail messages, or fields. These different field types include: Filename, Subject, Sender, Mailto and CC. The e-mail repository can be also searched based on delivery time.

Other advanced options that ZyFIND supports include:

- Fuzzy and progressive searching
- Sorting methods: sort on, ascending/ descending
- Show attachments in the result list.



Waiting Till You Are Investigated Is Not the Time To Implement

The bottom line for e-mail archiving is that regulators, auditors and lawyers will not be sitting idly by waiting for your organization to "get ready" to archive e-mails. They are moving forward with investigations and audit activities whether your systems are ready or not.

You Are Being Watched

E-mail archiving should be viewed in much the same way as you may few getting your annual flu shot. You can take the chance that you won't get the flu this year and not get a shot but if you do you may get lucky enough to catch it in time to treat it quickly or you could wind up being out sick for a week or two. The other alternative is get the shot and be prepared so you don't have to worry about getting sick and be productive throughout the season. If you don't archive your e-mail systems quickly you are playing "Russian Roulette" with authorities and with your companies exposure to litigation. If you do implement e-mail archiving tools and practices you will at least be prepared for the attack and save time, money and resources in the long run.

About ZyLAB Technologies

Founded in 1983, ZyLAB is the leading provider of document imaging and paper filing software that helps Global 2000 companies and governments digitally file and manage millions of pages of paper and electronic documents.

ZyLAB's comprehensive investigative capabilities, with its high quality search and retrieval features supporting over 200 languages, giving users the ability to organize and easily share all information online, makes ZyLAB software the preferred solution for intelligence agencies, law-enforcement organizations, prosecutors, law firms, courts, in-house legal departments.

With over 7,000 installations worldwide and over 300,000 users including ***Amtrak, the FBI, the CIA, the INS (BICE Dept. of Homeland Security), the New York Stock Exchange, Pepsico, Riggs Bank, the State of New York,*** and ***Walt Disney***, ZyLAB has a wide breadth of experience and knowledge across a variety of different industries and business applications. The company has offices located in McLean, Virginia in the U.S., the U.K., Germany, Spain, France, the Netherlands, Singapore and Australia to provide global service to its client base.

For more information you can access www.zylab.com

About e-Nterprise Advisors

e-Nterprise Advisors provides market research, strategic planning and advisory services in the dynamic area of Enterprise Content Management (ECM) to both vendor and user organizations. Technology and market areas covered include Records Management, Electronic Document Management, Content Management, Knowledge Management, Electronic Imaging, Business Process Management and Search and Retrieval Technologies.

Priscilla Emery is President and founder of e-Nterprise Advisors. e-Nterprise Advisors provides market research, strategic planning and advisory services in Enterprise Content Management to both vendor and user organizations. Prior to establishing e-Nterprise Advisors, she was Senior VP of Information Products and Services for AIIM International where she was responsible for the development and delivery of publications and other information-oriented products and services to AIIM members and associates. Prior to her position at AIIM, Ms. Emery was VP and Director of Gartner's Electronic Workplace Technologies research center and New Science's Intelligent Document Management service. She has provided many Fortune 500 user and vendor organizations with strategic planning advice in the areas of document management and the assimilation of new and emerging technologies. Ms. Emery has also worked at Blue Cross & Blue Shield of Connecticut (now Anthem), Combustion Engineering (now ABB), Primerica Corp. (now Citicorp) and Bell Telephone Laboratories.

Ms. Emery has over 25 years experience in the information systems industry, has been a featured speaker at international industry events, has been quoted in business and industry publications such as *The Wall Street Journal, The Washington Post, Computerworld, InformationWeek, Software Magazine* and *PCWeek,* and has written numerous articles for publications, such as *Imaging and Document Solutions, e-doc, KMWorld,* and *DB2 Magazine*. She has also written *Knowledge Management: The Essentials*, an AIIM International publication, and *E-Mail Management Tools: Sorting Through the Options*. She has a B.A. in Mathematics from Lehman College (part of the City University of New York) and is listed in the 16th edition *of Who's Who of American Women* and the third edition of *Who's Who of Emerging Leaders*. Ms. Emery is an Advisory Board Member of the Electronic Document Systems Foundation and was also listed as one of the top 50 influencers of the document management industry in *KMWorld* magazine. She has recently been named conference chair for the 2004 AIIM Conference Program committee and one of the 20 Leaders to See in 2003 by CMSWatch. Ms Emery also holds the Masters of Information Technology and Laureate of Information Technology for Electronic Imaging designations from AIIM International and is a member of both ARMA International and Xplor International.