

"Email Mining": Emerging New Business Application

Companies Are Awakening to the Value of Archiving Their Email



...a white paper by Graeme Thickins

 **ComplianceVault06™**

There's a new application coming soon to a corporate email system near you. It's a logical extension, experts say, of the increased emphasis on the archiving and searching of email these days, driven by regulatory compliance and legal discovery. The latter have become cold, hard realities in today's corporate world.

What Employees and Businesses Should Know About Email

According to The Journal of Business Law, there have been several instances where email has played a key role in destroying an individual employee's or a corporation's reputation, and in many cases doing both. "One may ask," an article in this journal states, "what right to privacy does an employee have in the workplace? Although, most electronic communications are protected by the Electronic Communications Privacy Act of 1986, communications in the workplace are commonly held to fall under the ECPA's exceptions....business e-mail is considered exempt from the privacy safeguards of the Act." There is little legal regulation, says the article, on an employer's ability to monitor e-mail, "giving employers much freedom to intrude on workers' privacy."

Many employers now have their own detailed policies regarding employee use of email, and they make sure there is no expectation of privacy regarding email in the workplace. The Journal of Business Law suggests these guidelines for instituting and managing a business email policy:

- Establish a corporate e-mail "Acceptable Usage Policy." including a code of conduct stating what is prohibited (threatening/harassing other employees, obscene/offensive language, sending confidential materials outside of company or to unauthorized personnel)
- Require training sessions to raise awareness of rules and policies concerning electronic communications and remind employees frequently of the usage policy
- Inform employees that their email is regularly monitored
- Educate employees that monitoring of all email through the corporation is for their own legal liability protection as well as for the company
- Designate email and other electronic communications as a corporate asset
- Utilize filters to automatically monitor and block content bared under the "Acceptable Usage policy"
- Attach a legal corporate disclaimer to every email sent from the business email account

This coming new application is called "email mining," because it's all about extracting knowledge from a company's rapidly expanding email stores, using specialized email archiving solutions.

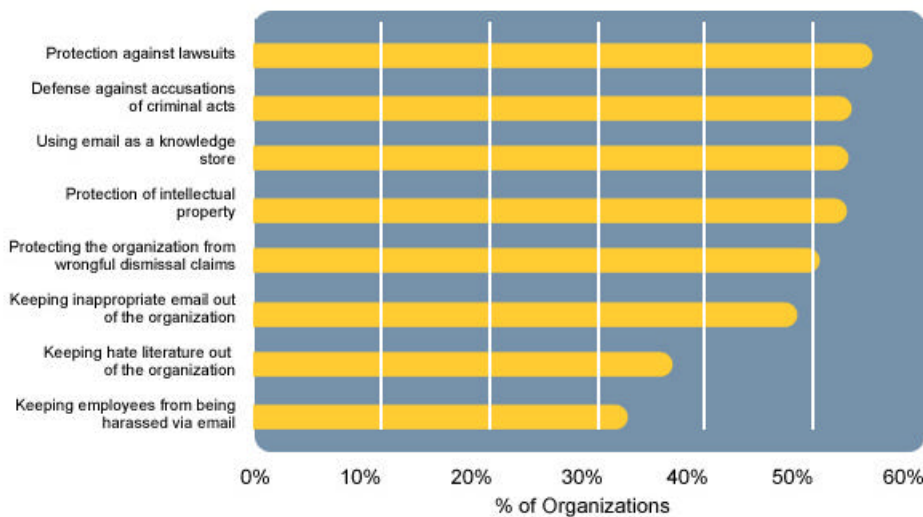
With the ongoing explosion of unstructured data, companies are increasingly finding it necessary to mine various types of knowledge bases, both within and without their corporate walls. Call it what you will: knowledge mining, intelligence mining, or simply a variant of corporate search - it's a good bet for attention and focus in 2006 and beyond, for a number of reasons.

Email mining is one type of intelligence mining that companies are already starting to do inside their own firewall. Email remains the killer app in business communications, and is not expected to relinquish that lead anytime soon. A recent report published by IDC on the future of email talks about "the crucial role for email throughout the collaboration process." And it raises the possibility that email content may "parallel that of other enterprise application data and business processes."

Mark Levitt, the lead IDC analyst behind the report, said this: "Organizations realize that the information contained in email systems is invaluable and in

some cases unique - not found anywhere else. However, most organizations treat their email stores, and archives if they exist, as repositories for personal reference. This is due to the fact that most firms do not have access tools that enable one user to access emails or attachments in another user's mailbox." Levitt went on to say that firms first need to do a better job of educating users that their corporate email mailboxes belong to the firm, not to the users. As far as businesses tapping into their ever-growing email archives, he said: "There is a need for text mining, categorization, and search tools that enable all users to find information that will help them do their jobs."

Reasons to Preserve and Manage Email
 (% responding a 'strong' or 'very strong' reason)



Data: Osterman Research, "Critical Decision Factors for Messaging Management," 2005.

The mining of email actually has a history of sorts, at least in forensics, surveillance, and the government intelligence community. In this respect, it's not the only technological advance to come from the world of government, of course - think the Internet and, in fact, email itself. A com-

pany's email store or archive (if it has one), essentially becomes a large, constantly expanding *unstructured database*. What that means is it's not transactional data - columns and rows of figures - it's basically text, along with a myriad of types of file attachments. And new email archiving solutions (more on that later) don't just search within the email messages themselves, but within the attachments as well.

Just How Big Is Corporate Email?

According to a recent article in *Messaging Pipeline*, 35 billion email messages are generated every business day, up from 10 billion daily just five years ago (data from IDC). The Radicati Group says the average corporate user sends and receives about 84 emails a day, using about 10 Mbytes of storage on a daily basis. By 2008, that same average worker's daily emails will take up 15.8 Mbytes, says Radicati.

The article goes on to say: "Without question, email is the most widely used software application in virtually every corporate enterprise....corporate email messages have achieved the same status as other commonly used business documents."

The Automated Appliance Solution

A simple, in-house solution for preserving email is provided by the ComplianceVault™ Email Archiving & Retrieval Appliance from Intradyn Inc. It is an integrated hardware/software solution that captures all of a company's email on a continuous basis. It can be put to use immediately, requires no integration time or expense, is easy to use, and has extensive, sophisticated search capabilities for fast retrieval of specific emails. In addition, the device has a full audit capability, which tracks all access to the system (required by some regulations). Specific emails can be retrieved in seconds or minutes - in stark contrast to the hours, days, or weeks that many companies have needed to meet the demands of courts or regulatory agencies in the past, which involved trying to search through backup tapes. This lack of responsiveness is the key reason many large fines and penalties have been levied to date - because companies could not retrieve requested email records within the timeframes ordered.

The ComplianceVault™ appliance is available in 1U rackmount models with either 320, 500, and 800 GBytes of disk space, or a model with 1 Terabyte capacity. Each model archives all email onto disk, as well as to non-alterable WORM (write-once, read-many) tape. The latter is accomplished via bundling a Sony® AIT™ automated tape library (also called an autoloader) with the ComplianceVault™ appliance.

Another important feature of the device is that it provides the customer with the choice of encrypting the tapes on which the emails are stored, providing highly secure protection of the data in the event of tape loss or theft. Intradyn's Encryptape™ technology is based on the same type of encryption used by the U.S. federal government, the highest level available, which is AES 256-bit - called "strong" encryption.

The ComplianceVault™ appliance works with Microsoft® Exchange®, Lotus® Notes®, and virtually any IMAP or POP3 email server, which are the standards behind most email systems in use today. When attached to a company's network, all emails are continuously saved and indexed to disk, plus archived and indexed to tape - making them available for retrieval in a manner compliant with the company's corporate policies and industry regulations. Sony's AIT WORM tape media comes in specially marked cartridges, which have a life of up to 30 years.

Additional information about the ComplianceVault™ is available at www.intradyn.com, or via sales@intradyn.com.

And, of course, there is no argument that corporate e-mail represents a litigious gold mine of information for discovery in the event of a lawsuit.

Beyond regulatory compliance and discovery, however, there are very positive reasons for a company to make email management a part of its records management procedures. The content captured in corporate email can be used as part of what some companies call their enterprise content management (ECM) system, for example, to speed decision-making and enhance a company's business processes. A wealth of business intelligence is hidden in corporate email systems, intelligence that can be put to use in business processes like adding prospects to a CRM database, updating customer records, or signing up prospects for online training, for example.

When an organization's email messages are archived, metadata is captured to enable searches, which can be conducted by keyword, sender, or receiver, using such advanced search methods as proximity searches, fuzzy searches, and wildcard searches. Once preserved as a record in an archiving system, access and retrieval of email content is vastly simplified, which can significantly reduce litigation costs. And that means a company could realize a return on investment on their email archiving solution with just one single legal discovery challenge.

An Email Is Forever

Once a user sends an email, he or she has virtually no control over what happens to it. That email can be printed, forwarded to others, edited, and changed dramatically, all without the knowledge or consent of the sender. But by establishing policies that capture outbound email messages as business records, companies can protect themselves against unwarranted claims by providing "digital originals." Because regulators and the courts treat email messages the same as written documents, managing those email messages as records assures that they meet the burden of proof of such regulations as Sarbanes Oxley (SOX), for example.

Regulators are not just concerned with email messages. Instant messaging (IM) is also considered part of the business document trail, and these written communications, with all their electronic shorthand, must also be stored as business records in order to adhere to regulations such as SOX.

Email As "Corporate Memory"

Thus, companies are seeing the value of mining this internal treasure trove of knowledge, for a variety of reasons. Many of them are already doing it for legal or compliance purposes - that is, using the advanced search techniques provided by new email archiving solutions to very quickly retrieve specific emails required for compliance audits or legal discovery. Such solutions range from (1) complex enterprise apps requiring extensive integration, to (2) services available on an outsourced or hosted basis, to (3) simple, internal hardware/software "appliances" that do the job as a plug-and-play solution (see www.intradyn.com). All came about in recent years primarily because companies were facing regulatory pressures and court orders to produce emails very quickly, or risk huge fines and penalties.

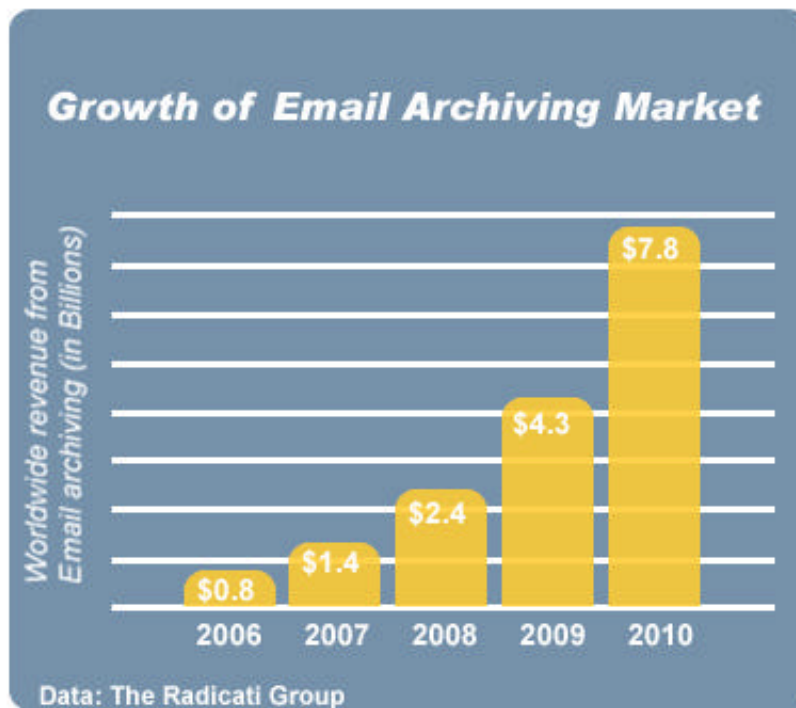
Email volume is growing rapidly - IDC says it more than doubled over the past two years. And, typically, a company's total data is now more than 80% of the "unstructured" variety, according to Enterprise Strategy Group, which of course includes that rapidly growing category of email.

As companies are now actively learning how to mine their email knowledge base for the above "reactive" reasons (legal and regulatory), many experts believe they will naturally expand the practice to more "proactive" uses - to mine intelligence that can improve or speed such functions as product development, marketing, human resources administration, customer support, and other positive, business uses. As

email (including attachments) fast increases as a percentage of the "corporate memory," its value to the corporation becomes ever more critical. Who wouldn't want to be able to mine that knowledge base for all its worth?

Email Archiving As the "Mining" Enabler

"We estimate that, at one time or another, 70% of an organization's intellectual property is transmitted through email," said Brian Babineau, senior analyst at Enterprise Strategy Group. "We expect this to continue as email



remains the collaboration application of choice.

However, the value of the data resides in the intellectual property. Email is simply a communication mechanism. Email does provide some structure to data that allows it to be more useful. More importantly, email archiving solutions add meta-data to messages that enables organizations to intelligently store the emails. This added intelligence, including sender/recipient/keywords, is the key to allowing other applications, like data

warehousing or CRM tools, to leverage the data. Without the meta-data, email would be just another data transport mechanism."

Why will email mining catch on? According to Gary Doan, one of the founders of pioneering email archiving appliance-maker Intradyn, "The concept of email mining is easy for businesses to grasp - moreso than, say, 'email management' or the much broader concept of 'enterprise content management'. No one really wants to manage email. They may want to control it, which is a tough job. But what they very much *do* want is a simple way to search through it." Doan continued: "An email archive is a knowledge base that can be used by management to gather information on

communications processes, negotiations, problems, suggestions from customers, internal idea flow, or even office gossip."

Another expert is Michael Osterman, president of messaging and collaboration research firm Osterman Research. When asked of the benefits of email mining, he said: "There are a number of benefits from a knowledge management focus for archiving, including bringing new employees up to speed on what their predecessors told clients, allowing individual users to see what they or others told customers or prospects, informal reviews for policy compliance even where there is no formal requirement for such reviews, and so forth."

Osterman also noted this: "In a typical company today, I would be surprised if the percentage of unstructured data represented by email wasn't at least 60-70%. I think companies will see the value of email as a source of knowledge increasing, primarily because this will be one of the key secondary benefits that organizations derive after they have deployed an archiving solution. Email is increasingly being used as a source for legal discovery, so email is being taken more seriously as a source of discoverable content."

About the author: *Graeme Thickins has worked in the information technology field for more than 25 years, and has written extensively for more than six years about data storage, compliance, email archiving, and other IT and business trends. He currently serves as vice president of marketing for Intradyn Inc., Eagan, Minnesota.*

References:

IDC: "Worldwide Email Usage Forecast, 2005-2009: Email's Future Depends on Keeping Its Value High and Its Cost Low," December 2005.
<http://www.idc.com/getdoc.jsp?containerId=34504>

The Journal of Business Law: "E-mail: Speed, Convenience, and Liability," March 2005.
http://iblsjournal.typepad.com/illinois_business_law_soc/2005/03/email_speed_con.html

Messaging Pipeline (CMP): "E-mail Management For SOX: More Than Meets The Eye," January 2006.
<http://www.messagingpipeline.com/showArticle.jhtml?articleId=177104024>

InternetNews.com (JupiterMedia): "E-Mail Archiving Stealing ECM Spotlight?" January 2005.
<http://www.internetnews.com/storage/article.php/3465781>

For more information, contact sales@intradyn.com. Additional information is available at www.intradyn.com/compliance.

Copyright © 2006, Intradyn Inc. All rights reserved. ComplianceVault™ is a trademark of Intradyn. Sony® and AIT™ are a trademarks of Sony Corp. All other names used within are trademarks of their respective owners.

This document is for informational purposes only and is subject to change without notice. Although reasonable effort has been made in the preparation of this document to ensure its accuracy, Intradyn Inc. assumes no liability resulting from errors or omissions in the document, nor from the use of the information contained in it. Intradyn Inc. reserves the right to make changes in product specifications without reservation and without notification.

Note: When designing or implementing data retention or data storage solutions for compliance purposes, always seek the advice of an attorney with specialized expertise in electronic records retention and archiving, and the associated laws and regulations that apply for your industry or geographical jurisdiction(s). This document is provided only for informational purposes.