THE JOURNAL OF DIGITAL FORENSICS, SECURITY AND LAW

Journal of Digital Forensics, Security and Law

Volume 1 | Number 1 Article 1

2006

Electronic Data Discovery: Integrating Due Process into Cyber Forensic Practice

John W. Bagby
The Pennsylvania State University

John C. Ruhnka
University of Colorado at Denver and Health Sciences Center

Electronic Data Discovery: Integrating Due Process into Cyber Forensic Practice

John W. Bagby
Professor of Information Sciences and Technology
College of Information Sciences and Technology
Co-director Institute for Information Policy
The Pennsylvania State University
301C IST Bldg
University Park, PA 16802 USA
jbagby@ist.psu.edu

John C. Ruhnka
Professor of Law and Ethics
Academic Director of the Bard Center for Entrepreneurship
Graduate School of Business Administration
University of Colorado at Denver and Health Sciences Center
1250 14th St., Suite 242
Denver, CO 80217-3364 USA
John.Ruhnka@cudenver.edu

ABSTRACT

Most organizations and government agencies regularly become engaged in litigation with suppliers, customers, clients, employees, competitors, shareholders, prosecutors or regulatory agencies that nearly assures the need to organize, retain, find and produce business records and correspondence, e-mails, accounting records or other data relevant to disputed issues. This article discusses some high visibility cases that constrain how metadata and content is routinely made available to opposing parties in civil litigation, to prosecutors in criminal prosecutions and to agency staff in regulatory enforcement litigation. Public policy, as implemented in the rules of evidence and pretrial discovery, restrict electronic data discovery (EDD) as it becomes a predominant and potentially costly pre-trial activity pivotal to modern litigation. This article discusses these constraints while identifying opportunities for the interdisciplinary activities among litigators, forensic experts and information technology professionals.

Keywords: electronic data discovery, cyber forensics, pre-trial discovery, litigation hold, spoliation, obstruction of justice, electronic records

management, metadata

"As a litigator, I will tell you documents are just the bane of our existence. Never write when you can speak. Never speak when you can wink."[1]

1. INTRODUCTION

The quote above might be adapted to the emerging arena of electronic data discovery (EDD) and cyber forensics as follows: "Never email when you can write, never write when you can phone, and never phone when you can meet face to face." While such defensive practices might seem to avert potential "smoking gun" disclosures, they also can promote a corporate culture of cover-up and risky behavior inconsistent with good public and corporate policy. This article employs traditional doctrinal legal research to review the emerging field of EDD under criminal, civil and regulatory process and provide analysis of the implications for cyber forensics and sound electronic records management practices.

Eventually every company, not-for-profit organization, non-governmental organization, self-regulatory organization or government agency will end up in litigation or involved in a governmental investigation or regulatory action, some as a plaintiff to enforce rights or claims, but most as a defendant. Claims can be filed by customers, shareholders, former employees, competitors, or by watchdog regulatory agencies such as the Security and Exchange Commission, Federal Trade Commission, Internal Revenue Service or the U.S. Justice Department. In the event of litigation or governmental investigations, it is a virtual certainty that records will be subpoenaed, including business records, correspondence, and phone logs and emails related to the disputed issue.

Many people who use email for its speed and efficiency, as compared with phone calls or face-to-face meetings, run a significant risk - phone calls and personal conservations are evanescent unless wiretapped or recorded. By contrast, email is persistent since it is recorded and preserved by the sender, the recipient and various network servers and ISPs during transmission. Recent high visibility cases in which emails and electronic evidence played a critical role, now clearly signal that nearly any electronic record is subject to compulsory "discovery," the mandatory pre-trial process of obtaining information from before trial through demands on opposing and third parties. The discovery of electronic information includes the recovery of both content and meta data of such records during the pre-trial discovery phase of criminal and civil litigation as well as in investigations by regulatory agencies. Anyone

[1] Statement of panelist Jordan Eth, *Sarbanes-Oxley: The Good, The Bad, The Ugly*, Nov.10, 2005, panel hosted by the National Law Journal and Stanford Law School's Center on Ethics, *reprinted in* National Law Journal, 12 December 2005.

with a legally-enforceable interest involved in the litigation, has the right to demand traditional paper records, as well as electronic communications, to use to prove or to contest facts at issue in legal proceedings. The term, electronic data discovery (EDD) is generally becoming applied to a wide range of electronic document acquisition from opposing parties, potentially adverse parties or third parties, in preparation for civil, criminal and regulatory proceedings.

Many people still remember the famous enforcement action against Merrill Lynch when it's star security analyst Henry Blodget referred in internal emails to a "dot.com" company that Merrill had underwritten and which he had recommended to Merrill Lynch's investor clients as "a piece of junk" and "a powder keg." These email disclosures revealed blatant conflicts of interest between Merrill Lynch's underwriting activities and its investment advisory divisions. Public disclosure of these conflicts of interest by New York Attorney General Elliot Spitzer caused significant damage to Merrill Lynch's reputation and was a pivotal factor in the $100 million civil fine in the settlement with New York State and the U.S. Securities and Exchange Commission.[2]

Litigants on both sides can be required to help identify and to disclose emails and archived internal and external communications (including email files, chat files, network server logs, and network back-up tapes) which may be relevant to disputed issues involved in litigation or investigations once such records are requested by an opposing party. Many persons may still incorrectly believe that such internal electronic communications and records are "private" or can be relatively easily destroyed. The reality is that EDD often has the ability to provide a roadmap of "who knew what and when," which is so pivotal in litigation. The costs of electronic records management (ERM) and of responding to mandatory EDD discovery requests are very significant, including the search of email and document files residing on employees' desktop and laptop hard drives and restoring network backup tapes, and even "recovering" deleted files from hard drives of persons potentially relevant to the litigation or investigation. Such costs of responding to EDD discovery requests can easily exceed a quarter of a million dollars in business litigation. The threat of discovery costs, even before EDD became a predominant focus in litigation, could sometimes be used as a "club" to force opposing parties into pre-trial settlements in order to avoid much higher discovery costs. Judicial penalties for an insufficient response to a discovery request or for concealment or destruction of requested or potentially relevant records can be severe, including legal sanctions for spoliation in civil and regulatory investigations and obstruction of justice charges in criminal cases. Another famous 2002 case is illustrative, when Martha Stewart was prosecuted for obstruction of justice for trying to cover-up behavior that could indicate that her sudden sale of 3,928

[2] SEC v. Merrill Lynch, Pierce, Fenner & Smith Inc., 03 CV 2941 (S.D.N.Y. April 28, 2003).

shares of ImClone stock was based on inside information.[3] Her criminal conviction for obstruction of justice in 2004 largely rested on electronic evidence indicating tampering with potential evidence.

2. DEFINING ELECTRONIC DATA DISCOVERY

After litigation has commenced with the filing of a complaint and well before the trial takes place, there is a (usually lengthy) period of trial preparation during which facts are gathered and information is requested from all parties, and parties must make requested but unprivileged information available to the opposition. The rules of pretrial discovery require production of requested information, so long as the information is potentially relevant to the issues in dispute, even if this disclosure would weaken or compromise the producing party's interests. All litigants, both plaintiffs and defendants, may demand production of documents and "data compilations" relevant to the prosecution, complaint or defense of an issue from almost any party in possession, custody or control of that information. This document production duty is required in U.S. federal courts in civil cases under the Federal Civil Rules of Procedure (Fed.R.Civ.P.)[4] and in criminal cases under the Federal Rules of Criminal Procedure (Fed.R.Crim.P.) as well as under similar civil and criminal discovery rules in all the U.S. states. Fed.R.Civ.P.Rule 34(a) broadly defines the potential sources of discoverable data as "writings, drawings, graphs, charts, photographs, phonorecords,[5] and other data compilations from which information can be obtained, translated, if necessary, by the respondent through detection devices into reasonably usable form." Therefore, data compilations exposed to discovery include reports, correspondence, memos, text files, spreadsheets, PowerPoint presentations, digital photos, graphics and art, email including all attachments, instant messages and any other data created on or stored on a computer, computer network or any other electronic storage media.

Often one party in litigation (or an investigation) may not be in possession of sufficient facts or evidence to prove the other party's fault. Under U.S. pretrial discovery rules, parties may not hide, destroy nor deny an opposing party access to most forms of potentially-incriminating evidence by claiming proprietary control. The formulation of "justice" in the U.S. gives all litigants the right to request and examine records, files, or other evidence from the opposing party if the latter holds exclusive control over such information. With great frequency, pre-trial discovery often unearths evidence critical to the requester's success. Consider the high visibility $253 million jury verdict in

[3] *U.S. v. Stewart*, 305 F. Supp. 2d 368 (S.D.N.Y 2004).

[4] FED. R. CIV. P. 26(a) (1).

[5] Under U.S. copyright law the term "phonorecord" is broadly defined to include analog vinyl and streaming tape recordings as well as digital recordings "fixed" in magnetic and optical media.

2005 against Merck, maker of the prescription painkiller Vioxx - a Cox-2 inhibitor.[6] Merck defended this claim that Vioxx had caused the plaintiff to suffer a fatal heart attack by arguing that arrhythmia or irregular heartbeat had caused the death and not a blood clot produced by Vioxx as the plaintiff alleged. However, the jury was likely greatly influenced by a 1997 internal email authored by Merck's own research scientist Alise Reicin, that was disclosed in response to a discovery request from the plaintiff. In the email, sent two years before the FDA approved Vioxx for sale to the public, the Merck research scientist worried that "the possibility of increased CV (cardio vascular) events is of great concern." Vioxx was subsequently withdrawn from the market by Merck in September, 2004.

Until the 1990's, discovery requests for the production of documents mostly involved paper records. Much of the content of personal communications including phone calls and face-to-face conversations and meetings was effectively unavailable unless there were contemporaneous written notes, tape recordings or the conversation was recalled in a witness' testimony under oath or in a deposition. The probative value of witness recall depends on witness' credibility as tested by cross-examination and on the strength of rebutting evidence. Today, however, most business records are in electronic form and a massive amount of both external and internal communication within and between organizations is transmitted through networks in digital form. Because of this, EDD has become perhaps the most important source of evidence in criminal and civil litigation as well as in government and regulatory investigations.

According to one study published in 2003 on the proliferation of electronic data, over 90% of business documents are created and stored electronically.[7] While the rapid increase in the overall percentage of business records and communications in electronic form is slowing, the absolute percentage of such records in electronic form will only continue to grow. Most intra-corporate and intra-governmental communication is conducted via email with a growing use of even newer electronic communications technologies including instant messaging and blogs. Each email contains an electronic record that includes not only the content of the email and any file attachments, but also the metadata embedded in the message which reveals the sender, the recipients, and the sequence and timing of linked messages and attachments. Such metadata helps to provide additional "who, what, when" detail to the content of communications. Electronic telephone logs and electronic voice mail repositories are also discoverable and will likely increase in importance as usage of voice-over Internet protocol (VoIP) proliferates.

---

[6] *Ernst, et. al. v. Merck & Co. Inc*., No. 19961- BH02, (D.C. Tex. Brazoria Co. 2005).

[7] Lyman, Peter and Hal R. Varian, *How Much Information*, 2003. Retrieved from http://www.sims.berkeley.edu/how-much-info-2003 on [12.20.05].

In addition to email, other persistent electronic records may be useful in proving facts, impeaching witnesses or casting doubt on an opponent's depiction. These electronic records can include documents created, altered or stored on computers, logs of network user access, records of web site visits, records of web site content downloaded, identity and dates of voice mail callers, employee security access code usage (access cards to physical locations, floors and offices, and parking facility access), some of which may be further corroborated by video surveillance tapes. E-Zaps and many other automated toll collection authorities frequently must respond to subpoenas for toll use records identifying customers and registered vehicle, time, date and location for various evidentiary purposes in legal and regulatory matters. Such advances over the past decade in both the amount and variety of information created and retained in electronic form has transformed litigation in the U.S. Innovations in technology continue to expand the types of electronic data discoverable. Consider how voicemail is shifting from individual desktop recorders to networked digital storage of voicemail messages. Unless effectively purged, voicemail may become as permanent and accessible as is email today resulting in much more voicemail discovery. This architecture holds the promise for unpleasant surprises because currently voicemails are most often preserved only by the recipient, not the sender. Indeed, it is unlikely senders can produce voicemails if sought in discovery making employers of most senders ignorant of the precise voicemail content and ignorant of the general content of renegade employees. It is costly to identify all potential voicemail recipients and request recipient recordings. Voicemail search technologies are still developing and as yet are neither as sophisticated nor effective as are search methods for screening email records for keywords in key fields such as sender, recipient(s), time, date, routing, attachments and subject matter.

3. ELECTRONIC RECORDS SUBJECT TO DISCOVERY

Targets of electronic data discovery requests or subpoenas can include anyone possessing information potentially relevant to issues, subject matter and participants in an investigation or litigation. Many potentially-discoverable email messages are archived on network server backup tapes, by third party service providers, at ISPs (e.g., Hotmail, Yahoo, AOL), on personal digital assistants (PDA) and Blackberry devices, on cell phones, on personal laptop and home computers and in various personal storage media (e.g., floppies, portable hard drives, thumb/jump/USB flash drives). External archives and caches of Internet content can be used to retrieve the past content of web sites and web pages previously posted but later "taken down" and such Internet web archives are often admissible evidence in the proof of infringement, internet domain name misuse, defamation, trade libel, harassment, misrepresentation,

and fraud.[8] Litigators now often refer to doing a "Wayback", a search of proprietary archives of website materials for a potential adversary's prior Internet postings, to assist litigation strategies.

It is important to note that both federal and state discovery rules permit requests for preservation of potentially relevant records to be served on persons in control of such records even if a lawsuit has not yet been filed or such persons are not yet parties to a legal action or governmental investigation. Compliance with discovery requests is mandatory for opposing parties and their employees and agents. With some very limited exceptions for "privileged" communications, such as the attorney-client communications concerning pending litigation and investigations and the attorney work product privilege, almost all information is discoverable from almost any opposing party or third party source. Depending on the situation, there are other important, but typically narrowly construed privileges, such as the spousal privilege, the doctor-patient privilege, the priest-penitent privilege, and in much more limited situations the accountant-client privilege and the self-evaluation privilege. Even when unprivileged information is obtained in discovery, its use and public disclosure may be limited by various types of confidentiality, such as court records held under seal by a judge's order for trade secrets or national security matters. Increasingly, confidentiality requirements or secrecy orders may suppress disclosure of divorce records and information about juveniles and minors. Unless specific public policy exceptions like the very limited privileges discussed above are applicable, there is no general right to privacy in the U.S. that prevents discovery of potentially incriminating information or other evidence damaging to a party's interest.[9]

Metadata is another important aspect of electronic records that many business executives and government leaders are still unaware of. Metadata is literally "data about data" that accompanies data files and electronic communications. A "metadata tag" will indicate various aspects of every file created on a computer, including the file creation date, the identity of the computer on which it was created, the IP address from which it was accessed, dates of editing or simply last viewed/opened and by whom, and all subsequent alterations and deletions. Email metadata may also show the sender's address book information, dispatch and receipt date of messages, information about forwarding or replies, and the existence or content of attachments. Metadata can be useful in showing the sequence and authorship of critical

[8] The Internet Archive claims that its Wayback Machine contains "approximately 1 petabyte of data and is currently growing at a rate of 20 terabytes per month." *See* http://www.archive.org/.

[9] In the high visibility antitrust case against Microsoft, an email by Bill Gates sent from his home computer, on a Sunday evening, to someone outside Microsoft, concerning the competitive objective of bundling Internet Explorer with Windows was held not to be protected by Mr. Gates' personal right to privacy and was admitted as proof relevant to Mr. Gates' state of mind.

communications. For example, in the 2004 criminal trial of Martha Stewart for obstruction of justice in the Imclone affair, the U.S. Department of Justice used metadata from an electronic log of office phone calls that indicated that a log of one phone call had been erased and later restored on a different computer. The Department of Justice used this evidence of tampering with the evidence, when correlated with an electronic diary entry showing a meeting with her Merrill Lynch broker, to argue that the defendants had met to develop a "cover story" to explain the sudden sale of her Imclone stock, and later attempted to cover-up the meeting. Despite evidence that Ms. Stewart had second thoughts almost immediately and allegedly reentered the erased phone log entry on her own office computer, this metadata record supported the government's contention that she had attempted to tamper with evidence of the alleged meeting and cover-up. Additional forms of metadata are evolving as the architecture of communications methods evolves. The capture and association of user IP addresses with particular web site visits by ISP web servers and the potential for court orders to require disclosure of identity of web users will likely significantly limit the anonymity of web communications and transactions. Such IP user access records are used by law enforcement agencies to identify purveyors and users of child pornography on the web or to identify both those serving and those downloading content using peer-to-peer networks.

3.1 "Litigation Hold" Compliance

All persons, businesses, institutions and government agencies have a legal duty to preserve records relevant to "reasonably expected" investigations and litigation. Thus, when litigation or a governmental investigation becomes "reasonably expected" the target of such litigation or investigation must immediately issue a "litigation hold"[10] throughout the organization that preserves all potentially-relevant evidence. Delays of only a few days in implementing a litigation hold might result in the automatic overwriting and permanent loss of network backup tapes containing emails potentially relevant to the pending litigation or investigation, which in turn can lead to charges of spoliation. Because of this, all organizations need to have electronic records management (ERM) programs in place that specify document retention and destruction policies depending on various factors such as time, date, subject matter of file, internal needs, etc., and which may be suspended in the even of a litigation hold. Consider the landmark (and extremely lengthy) EDD case of *Zubulake v. UBS-Warburg*[11] in which UBS was ultimately ordered to pay $29.3

[10] A provisional definition of litigation hold is that it is a legal duty to suspend routine document destruction when a lawsuit is filed or is reasonably anticipated. This definition is amplified in the ensuing text.

[11] Eight related Zubulake decisions were issued between 2003 and 2005: *Zubulake v. UBS Warburg*, 217 F.R.D. 309 (S.D.N.Y. 2003) (Zubulake I: allocating discovery costs for email production from backup tapes); *Zubulake v. UBS Warburg*, No. 02 Civ. 1243, 2003 WL 21087136 (S.D.N.Y. May 13, 2003) (Zubulake II: Zubulake's reporting obligations); *Zubulake v.*

million for gender discrimination and sexual harassment claims filed by a former female stock broker. The female stock broker plaintiff made discovery requests for all UBS internal emails sent by her alleged harasser that were expected to support the claim of harassment.

The judge in *Zubulake* described the litigation hold process for electronic evidence and email as follows:

> Once a party reasonably anticipates litigation, it must suspend its routine document retention/destruction policy and put in place a "litigation hold" to ensure the preservation of relevant documents. As a general rule, that litigation hold does not apply to inaccessible backup tapes (e.g., those typically maintained solely for the purpose of disaster recovery), which may continue to be recycled on the schedule set forth in the company's policy. On the other hand, if backup tapes are accessible (i.e., actively used for information retrieval), then such tapes would likely be subject to the litigation hold.
>
> However, it does make sense to create one exception to this general rule. If a company can identify where particular employee documents are stored on backup tapes, then the tapes storing the documents of "key players" to the existing or threatened litigation should be preserved if the information contained on those tapes is not otherwise available. This exception applies to all backup tapes.[12]

When about to initiate litigation, an opposing party's counsel may send a "preservation demand" to parties to the lawsuit demanding that they preserve email files and other evidence potentially relevant to a threatened lawsuit, which in turn will trigger a responsibility on the part of the targets of such preservations demands not to destroy backup tapes where such email records are likely to be located. Once a lawsuit is filed and pre-trial discovery begins, interrogatories are a useful tool that can be used to identify the existence and location of potentially useful electronic records and witnesses. These are written questions requiring the opposing party's answers under oath.

Once a litigation hold arises, there are numerous and significant legal

---

UBS Warburg, 216 F.R.D. 280 (S.D.N.Y. 2003) (Zubulake III: allocating costs between parties for restoration of email backup tapes), *Zubulake v. UBS Warburg*, 220 F.R.D. 212 (S.D.N.Y. 2003) (Zubulake IV: duty to preserve emails; defendant bears plaintiff's re-deposition costs) *Zubulake v. UBS Warburg*, 2004 WL 1620866 (S.D.N.Y. July 20, 2004) (Zubulake V: sanctions granted; UBS ordered to pay costs; defense counsel ordered to monitor compliance and preserve with a litigation hold); *Zubulake v. UBS Warburg*, 231 F.R.D. 159 (S.D.N.Y. Feb.2, 2005) (Zubulake Va); *Zubulake v. UBS Warburg*, 382 F.Supp.2d 536 (S.D.N.Y. March 20, 2005) (Zubulake VI: preventing admission of various evidence); and *Zubulake v. UBS Warburg*, 02-CV-1243 (April 6, 2005) (Zubulake jury verdict: $29.3 million in damages - $9.1 million compensatory, nearly $20.2 million punitive).

[12] *Zubulake v. UBS Warburg*, 220 F.R.D. 212, 218 (S.D.N.Y. 2003).

consequences. Legal counsel must immediately notify the information technology personnel to immediately discontinue destruction of records and backup tapes containing potentially relevant files. It is advisable to request the IT department to explicitly acknowledge the implementation of a litigation hold in order to avoid misunderstandings or slip-ups. Document destruction must cease immediately and preservation efforts must be undertaken for relevant emails, electronic documents and records, and backups. A written preservation plan with linked procedures should be developed that designates specific electronic evidence and emails for preservation and then the plan should be distributed to all employees potentially in control or possession of relevant evidence. Outside or independent directors of corporations should also be instructed to preserve relevant documents and records.[13] Litigation hold notices should provide reasons for identifying and preserving specific records, such as potential future litigation or regulatory enforcement actions, filed lawsuits or complaints, preservation letter, or a court order requiring the preservations of records. The notice should also warn of serious sanctions for failure to comply.

Large organizations that are regularly engaged in litigation need to have EDD teams already organized to quickly and effectively implement a litigation hold including experienced members from IT, legal, human resources and other in-house and third party records management professionals. Litigation hold compliance must be actively monitored to guard against inadvertent destruction of relevant records or the appearance that such records have been tampered with. All files containing potentially relevant evidence should be preserved in an electronic mirror image "snap shot" as soon as the litigation hold attaches, in order to authenticate the pre-discovery state of such records and to establish the "chain of custody" before any internal review of these records for potentially relevant files is undertaken. An internal review of files from desktop, laptop and handheld computers of the key employees potentially involved in litigation is usually necessary determine whose emails, files or Internet activity may contain potentially relevant emails or documents, but this should be done only *after* an archival "snapshot" of such files is made in unopened form, since opening or forwarding such files for in-house review will alter metadata in the files which may give the appearance that such files were subsequently tampered with. Electronic forensics experts, often third party service providers are frequently needed to recover or preserve evidence, particularly where relevant evidence may exist on backup tapes or erased or deleted files on employee computers. Such experts employ procedures and software that can authenticate the chain of custody and thereby refute charges of alteration or destruction post-litigation hold date. Backup media must be preserved to avoid charges of post-litigation hold date tampering.

[13] See In re Triton Energy, Ltd. Securities Litigation, 2002 WL 32114464 (E.D. Tex.

4. ALLOCATING COSTS OF ELECTRONIC DATA DISCOVERY

Some critics of U.S. litigation practices have argued that pre-trial discovery is so costly that it can be intentionally deployed to coerce defendants into arguably unjust, but perhaps cheaper, settlements. Indeed, several proposed reforms of product liability, tort liability and securities litigation involve a litigation "stay" that serves to halt litigation in order to enable defendants to challenge unsubstantiated litigation claims before significant discovery costs have to be incurred.[14] Judges have long had the power to consider the oppressive costs of discovery. The basic discovery rule in federal civil litigation, Federal Rule of Civil Procedure 26(c)[15] grants targets of discovery requests the right to seek protective orders limiting the scope of discovery if the request is overly broad, seeks privileged information or would result in "annoyance, embarrassment, oppression, or undue burden or expense."

In the past, when paper records were predominant, discovery targets were typically required to bear the costs of locating, screening for relevance, assembling and copying requested records. Today, the much larger volume of potentially-relevant electronic records, exacerbated by very significant cost to "convert" (decompress and reformat) network backup tapes into reviewable form, has prompted some trial judges to shift some of these electronic record discovery costs to the requesting party. The *Zubulake* decisions differentiated five categories of increasing recovery difficulty and expense that are currently relevant to determining EDD cost sharing:

1) **Active** online data on hard drives or active network servers;
2) **Near-line** data on removable media (e.g., CD-ROMs, floppy disks, magnetic tapes);
3) **Off-line** storage or archived data of organized files on off-line removable media;
4) **Backup** tapes organized not for convenient retrieval but for disaster recovery, produced in time sequence (network throughput backups made daily, weekly, monthly) and therefore intended primarily for massive re-image restoration:
5) **Deleted**, fragmented data files recoverable only by forensic experts.

Active data files, generally found in categories one through three above are accessible with the least costly restoration so discovery targets would usually be expected to bear the costs of production. Archived, compressed and hidden data recovery is more expensive and in some cases may require the specialized

[14] *See e.g.,* Private Securities Litigation Reform Act (PSLRA), Pub.L.No. 104-67, 109 Stat. 737 (1995).

[15] FED. R. CIV. P. 26(c).

services of outside forensic experts to decompress, reformat, organize and search such files. In such cases, judges are becoming more sympathetic to cost sharing between the requester and the discovery target. Sometimes a court may agree that requested network backup tapes be sampled first, in order to predict the effort and costs of the data recovery project before making decisions as to requested cost sharing. Such cost balancing is authorized by Federal Rule of Civil Procedure 26(b)[16] permits judges to weigh the *potential relevance* of requested documents against the *burden* on the target. The likely benefit of discovery is evaluated against the litigation stakes, the parties' resources and the likely relevance of the proposed discovery to resolving the issues.

The *Zubulake* case provides an example of a court attempting to allocate electronic discovery costs between the parties. SEC rules require broker/dealers like UBS Warburg to preserve all external emails sent to customers for three years, which UBS stored on optical disks making them relatively easy to search. UBS was ordered to assume 100% of those search and production costs. By contrast, proof of a gender discrimination claim would likely require production of internal emails, and these were preserved by UBS only on network backup tapes that existed in various files formats. UBS argued that Plaintiff Zubulake should pay part of the cost of searching these backup tapes because there were potentially 94 different network backup tapes needed to comply with the discovery request. Instead, the *Zubulake* court ordered that 5 of the 94 backup tapes selected by Ms. Zubulake as most likely to contain relevant emails be sampled to see how likely the remaining backup tapes were to contain relevant information. UBS's outside forensic expert billed nearly $12,000 for producing 600 potentially relevant emails from these five backup tapes, and the judge determined that 68 of the 600 recovered emails had potential relevance. Generalizing from the sample, the judge ordered UBS to pay 75% of the total $165,000 estimate for internal email restoration and 100% of the attorney's fees for reviewing and transmitting the relevant documents, estimated at over $100,000.[17] The scale of EDD restoration and recovery costs illustrated in *Zubulake* are not atypical. Indeed, the $300,000 restoration and search costs for emails in *Zubulake* is not unusual for a larger organization, raising legitimate concerns that the cost of threatened discovery can be used unfairly to force settlements. Some cases are limiting backup restoration for this reason. Consider the "fishing expedition" concerns of Federal Magistrate John Facciola when he declined to order the Department of Justice to do a full search of potential backup tapes in the case of *McPeek v. Ashcroft*, a retaliatory discharge claim by a former DOJ whistleblower.

> If the likelihood of finding *something* was the only criterion (for ordering a full search of backup tapes), there is a risk someone will have to spend

[16] FED. R. CIV. P. 26(b)(2)(iii).

[17] *Zubulake v. UBS Warburg*, 216 F.R.D. 280 (S.D.N.Y. 2003).

> hundreds of thousands of dollars to produce a single email. That is an awfully expensive needle to justify searching a haystack. . . . . ordering the producing party to restore backup tapes upon a (automatic assumption) likelihood that they will contain relevant information in every case gives the plaintiff a gigantic club with which to beat his opponent into settlement. No corporate president in her right mind would fail to settle a lawsuit for $100,000 if the restoration of backup tapes would cost $300,000.[18]

The privately developed Sedona Principles advocate more equitable cost sharing and hold some promise to influence judges' allocation of these burdens.[19]

5. AVOIDING CONSEQUENCES OF EVIDENCE DESTRUCTION

Parties always have a disincentive to produce documents that are incriminating or which disclose proprietary information or strategy. However, litigation process rules have long penalized the intentional destruction of evidence. *Spoliation* is the improper destruction of documents, including email, of potential relevance to a pending or forthcoming legal proceeding. The potential sanctions for spoliation vary depending on the spoliator's degree of fault and either punish the spoliator and/or give justice to the requesting party. This degree of fault varies with the spoliator's "culpable state of mind" - that there was a knowing destruction of evidence relevant to the opposing party's claim or defense. In spoliation cases the courts balance the degree of fault against the degree of prejudice caused to the requesting party. Sanctions may include:

(1) Discovery sanctions (monetary fines);

(2) Attorneys fees and costs payments to the requesting party for spoliation hearings;

(3) Additional discovery with the costs charged to the culpable party;

(4) Adverse inference instructions to the jury – permits the jury to infer that the spoliation destroyed evidence favorable to the opposition's case and harmful to the spoliator's case;

(5) Default judgment on a defendant (imposing liability) or dismissal of a plaintiff's case (an extreme remedy);

(6) Tort (civil) liability for injuries due to the spoliation.

5.1 Role of Document Retention and Destruction Programs

Electronic records management programs often include document management

[18] *McPeek v. Ashcroft*, 202 F.R.D. 31, 34 (2001).

[19] The Sedona Principles, The Sedona Conference, (Sept. 2005) http://www.thesedonaconference.org/dltForm?did=TSG9_05.pdf.

and document retention protocols to periodically eliminate unneeded documents and records. There are justifications beyond the obvious self-serving effort to eliminate potential "smoking guns." For example, the costs of storage, management and discovery response are well-known. When information is no longer useful or necessary for business purposes and no federal or state law requires its retention, there is strong incentive to accept the business reasons for destruction. Courts accept the balancing of retention/management costs against the possible future use of such records in litigation, if the document elimination policies are applied consistently and without a sole purpose to thwart opposing parties in litigation. The key standard is whether the potential discovery target could reasonably anticipate such documents could be relevant to a likely or pending government investigation or litigation.

Furthermore there must be clear proof that the document destruction program or its execution is implemented with a specific intent to destroy relevant evidence. Consider the recent U.S. Supreme Court reversal of the conviction of Arthur Andersen for Enron-related document destruction.[20] The 2002 criminal conviction of Andersen was overturned because the trial judge's instructions to the jury were flawed, leaving this intent element unclear. An in-house Arthur Andersen lawyer's email reminded staff to faithfully execute "document-retention" policies on their Enron audit papers and this allegedly triggered a extraordinary shredding of Enron audit documents just when, DOJ argued, Andersen could "reasonably anticipate" the forthcoming SEC investigation. This Arthur Andersen criminal conviction was arguably the direct cause of Arthur Andersen's collapse as clients defected. Thus, a simple reminder of document retention program procedures is not, in itself, wrongful without further proof of a conscious wrongdoing. Such wrongdoing would require contemplation of a particular official proceeding in which the destroyed documents might be material.

5.2 Record Retention Constraints

Legal requirements for document retention generally depend on three factors: (1) the line of business, (2) the type of records in question and (3) the significance of archived information to the business model. Government regulations from various sources, federal, state, local even international regulations, require the creation and maintenance of various business records. Some retention requirements apply to most businesses, such as tax records under IRS rules, wage and hour records under the Fair Labor Standards Act, cradle to grave toxic chemical handling under the Resource Conservation and Recovery Act (RCRA) or employee exposure records under regulations of the Occupational Safety and Health Administration (OSHA). Other requirements

[20] *Arthur Andersen v. United States*, 125 S.Ct. 2129 (2005).

are specific to regulated-industries such as for railroads, airlines, public utilities, nuclear power plants, federally-chartered or insured banks, securities broker/dealers, commodities future commission merchants or registered investment advisers. The Code of Federal Regulations (CFR) contains over 2,800 sections requiring record creation and/or maintenance. Retention periods range from 30 days up to 50 years. Few federal agencies have email retention rules yet, the SEC is at the forefront with rule SEC Exchange Act Rule 17a-4(f) requiring broker-dealers to retain all correspondence with customers for three years, and this specifically includes emails.[21] Retained copies must be in a "non-rewriteable and non-erasable" format that the SEC describes as follows:

> Under the rule, the electronic storage media also must verify automatically the quality and accuracy of the storage media recording process; serialize the original and, if applicable, duplicate units of storage media, and time-date for the required period of retention the information placed on such electronic storage media; and have the capacity to readily download indexes and records preserved on the electronic storage media to any medium acceptable under paragraph (f) as required by the Commission or the self-regulatory organizations of which the member, broker, or dealer is a member.[22]

Despite the SEC's detailed email retention requirements, such regimes at other federal and state agencies remain generally less detailed and frequently are unclear. Most retention laws are adapted to electronic communications rather than stated in language that is clearly and unequivocally applicable to electronic communications. Indeed,

Donald S. Skupsky, a leading author in document retention practices argues that existing federal records retention requirements may not clearly apply to emails.[23]

5.3 Perils of Aggressive Document Destruction

It has been reported that at least some organizations employ aggressive email destruction policies – erasure of emails after short time frames of 30 to 60 days. The tacit reasoning is that if internal emails can be routinely eliminated, this in turn will reduce the probability of the subsequent discovery of damaging emails. Some document retention experts argue that most email is private informal conversation deserving of the same treatment as face-to-face or phone conversations. Thus they would argue we do not need to archive all emails for

---

[21] 17 CFR 240.17a-4(f).

[22] Electronic Storage of Broker-Dealer Records, SEC Exchange Act Rel.No. 47806 (May 7, 2003).

[23] *See e.g.,* Skupsky Donald S., (1996) Discovery and Destruction of Email, Chapter 5 in: THE INTERNET AND BUSINESS: A LAWYER'S GUIDE TO THE EMERGING LEGAL ISSUES, (Computer Law Association).

the same reason that we typically do not memorialize phone calls or personal conversations. Such a position might require the identification and separation of more official, "permanent" emails from "informal" emails. However, giving senders or recipients an option to either archive or to eliminate prior communications that may involve personal or corporate malfeasance or criminal intent will likely seem unacceptable to many policy makers.

Institutional policies of aggressive email destruction can also increase the risk of spoliation or obstruction penalties. Consider the SEC's discipline and civil fines against Deutsche Bank, Morgan Stanley, Goldman Sachs, U.S. Bancorp and Salomon Smith Barney in 2002 for their alleged systematic destruction of network email backups and departed employees' hard drives.[24] Second, the Sarbanes-Oxley Act,[25] the SEC rules issued there under and the attendant internal control accounting standards issued by the Public Company Accounting Oversight Board[26] can arguably be interpreted to prohibit aggressive email destruction, at least for publicly-traded companies. Aggressive destruction arguably violates the Sarbanes-Oxley Act duties to maintain "adequate internal controls." Further, employee customs and institutional practices regarding email often preserve copies in multiple locations, on employee hard drives and in electronic and paper printouts. Therefore, aggressive email destruction policies may effectively eliminate only the organization's primary copy of potentially incriminating messages, leaving untouched secondary copies preserved by employees, service providers, suppliers, customers, online and Internet service providers. All too frequently such copies are held in less-friendly hands. A very invasive organizational effort would be needed to effectively limit and destroy all internal emails, and it must be expected that evasion and circumvention will frequently occur as copies are surreptitiously retained by both senders and recipients for individual defense and whistle blowing or retention is inspired by retaliatory motives in support of opposing parties or perceived victims.

Electronic record management practices increasingly include deletion of electronic file metadata using "wiping" programs that erase automatically-generated metadata on files, communications and Internet browsing history. Consider how commercially available applications such as *Metadata Assistant* by Payne Consulting[27] enables removing metadata such as a document's author(s), various dates of file creation, printing or access and edits from files to preserve confidential information before such files are released externally.

---

[24] *In Re Deutsche Bank Securities, Inc., Goldman, Sachs & Co., Morgan Stanley & Co. Incorporated, Salomon Smith Barney Inc., And U.S. Bancorp Piper Jaffray Inc*., SEC Exchange Act Re. No. 46937 (Dec. 3, 2002) Admin Proc. File No. 3-10957.

[25] Sarbanes-Oxley Act of 2002, H.R.3763, 107 P.L. 204, 116 Stat. 745 (July 30, 2002).

[26] http://www.pcaobus.org/.

[27] http://www.payneconsulting.com/.

Such tools are also useful to examine files submitted by opposing parties. Metadata wiping is prohibited following a litigation hold. Care must be taken because metadata can be falsified or surreptitiously supplied after wiping.

6. DISCOVERY NEGOTIATIONS

While courts are sensitive to limiting overly-broad and unjustifiably costly discovery requests, legal negotiations are typically undertaken to narrow the scope of a discovery request or to force a target to respond to costly pre-trial discovery requests. For example, the *Zubulake* gender discrimination case involved extensive motion practice, frequent court appearances and eight separate court decisions, most on EDD discovery issues, resulting in very significant fees for lawyers and forensic experts, and, ultimately, penalties for delay tactics as well as adverse publicity for UBS. Non-responsive production is all too frequently met with subsequent demands or court sanctions for failure to respond. Hence it is often advisable, at the outset of litigation or regulatory investigations, to negotiate mutually-agreed parameters for electronic discovery – including the subject matter, time window, electronic formats and production timetables. Many seasoned litigators can be willing to negotiate discovery limits with experienced and reputable counsel and forthright clients. Another area for negotiation can be the search criteria, keywords, search strings, concept searching and other screening criteria used to identify and search potential electronic files that might be relevant to the litigation.

7. SOME CONCLUDING OBSERVATIONS ON EDD

Despite incremental judicial reforms, there are no clear signs on the legal front that aggressive discovery of electronic information will slow anytime soon. Indeed, it can be expected that a steady stream of publicity will reveal alleged wrongdoing exposed via discovery of email and other electronic evidence. The preservation and strengthening of an organizations' "ethical culture" is a fundamental concern in how an organization responds to EDD requests. Organizational policies concerning preservation, monitoring and disclosure of electronic records can significantly impact both employee morale and the organization's long-term survival and prosperity. Organizations that attempt to deliberately limit electronic evidence accessibility and disclosure using frequent email and electronic record destruction, risk establishing a perception among their employees that this is done to hide unethical practices or malfeasance, and therefore, by implication, risk establishing a corporate perception that such activities will be shielded from pubic scrutiny or, by implication, are acceptable. By contrast, an organization that is able to make both internal and external electronic communications accessible for internal review as well as electronic discovery requests, where applicable, projects a culture of intolerance towards illegal or unethical activity. The ability to monitor employee and corporate activities and to take corrective action as necessary is an essential element of a strong corporate ethical culture. This

approach might be implemented with routine preservation of all emails and general backup in readily-accessible, well-organized and well-indexed central archives. This in turn, should help to reduce the financial burden to responding to EDD discovery requests and of screening potentially relevant files.

Paper record systems imposed very significant document storage costs as well as discovery search and production costs. By contrast, the costs of electronic storage continue to drop precipitously. New archiving, recovery and search technologies from various third party EDD experts can simplify and reduce EDD costs in the event of document requests. The SEC penalties against Deutsche Bank, Morgan Stanley, Goldman Sachs, and others in 2002 for the failure to preserve emails, discussed earlier, as well as the 2005 award of $1.58 billion by a Florida jury against Morgan Stanley and in favor of Ronald Perelman, which was assisted by an "adverse inference" instruction of the trial judge to the jury that they could infer that Morgan Stanley's failure to provide

requested email records was an indication of guilt, indicates that reliance of slipshod retentions of emails only in network backup tapes can add significant expense to EDD targets while also increasing risks of spoliation and obstruction.

In summary, electronic data discovery has the unique capability of uncovering evidence of internal and external actions and communications of many types that were formerly inaccessible, and which can often be determinative in civil, criminal, and regulatory litigation and investigations. Legal discovery practice is evolving to catch up with rapid technological developments in communications and data preservation and as a result, EDD can be expected to continue to be a critical driver of modern litigation practice and therefore of the burgeoning professional fields of cyber forensics. Landmark cases such as *Zubulake* and emerging requirements of regulatory agencies for the preservation of electronic communications in closely regulated business sectors such as financial services are setting potential standards for retention and management of electronic records that can be expected to spread to other industries as well.

AUTHOR BIOGRAPHIES

John W. Bagby, BA, JD, MBA, is Professor of Information Sciences and Technology and Co-Director of the Institute for Information Policy at the College of Information Sciences and Technology at the Pennsylvania State University. His interdisciplinary research in information assurance, litigation risk management, regulatory process, tort and product liability reform, intellectual property and technology transfer is published in various journals in law, economics, business and engineering as well as in numerous college textbooks.

John C. Ruhnka, BA, JD, MBA, LLM is professor of law and ethics at the

Graduate School of Business Administration and is Academic Director of the Bard Center for Entrepreneurship at the University of Colorado at Denver. He has published on corporate compliance, reporting and liability litigation issues in various journals including Corporate Information and Privacy Reporter, Securities Regulation Law Journal, Journal of Business Venturing, Journal of Health Politics, Policy and Law, Journal of Business Ethics, CPA Journal and Harvard Business Review.