Chapter 8

NEW FORMS OF EVIDENCE IN A HIGH-TECH AGE

I. The Essentials of Computer Discovery

Joan E. Feldman*

Introduction

Chances are good that the date you scheduled, the letter you wrote, and the interoffice message you just read have all been recorded on magnetic media. Like any other business record, electronically stored files are discoverable in litigation and can be used as evidence in the courtroom.

What distinguishes computer-based evidence from traditional paper documents in discovery? "Electronic" documents thought to be lost or destroyed can be recovered. Valuable information such as the time, date, and author's name may be embedded in the electronic version of a document. Comparisons of computer backups to existing documents can be used to show that a critical document was altered and when the event occurred. In the case of electronic mail, casual and candid correspondence may be frozen in time like insects in amber.

Glossary

Computer System refers to the entire computing environment. This environment may consist of one large computer serving many users (e.g., a mainframe or mini-computer) or one or more personal computers working individually or linked together through a network. A computer system includes all hardware and peripher-

*Managing Director, Navigant Consulting, Inc., Seattle, Washington. The author has formats for providing notice for nondestruction of computer files, deposing records custodians, and her "7-Step Roadmap to Better Electronic Discovery Management." They are available directly from Ms. Feldman, jfeldman@navigantconsulting.com.

als used (e.g., terminals, printers, modems, data storage devices), as well as the software.

Files are groups of information collectively placed under a name and stored on a computer. Files are organized in various folders sometimes referred to as directories and subdirectories.

Media is the generic term for the various storage devices used to store computer data. For personal computers (PCs), the most common media are the internal hard drive, compact discs (CDs), and floppy disks. Backup tapes, thumb drives, and digital video discs (DVDs) are other forms of storage media.

Networks are the hardware and software combinations that connect computers and allow them to share data. Two common ways PCs are networked are peer-to-peer and client-server. Peer-to-peer networks physically connect each computer in the network to every other computer in the network. Files are stored on the hard drives of the individual PCs with no centralized file storage. Client-server networks connect individual PCs called "clients" to a central "server" computer. In contrast to peer-to-peer networks, files are stored centrally on the server.

Filling in the Details with Computer-Based Evidence

Computer-based evidence exists in many forms and locations within any computer system. The key to finding and using this information is understanding the kinds of information that may exist and where within the system to look for each type of information.

Data Files

The primary function of most computer systems is to process and store information. Information processed and stored electronically can be divided into four basic categories:

Active Data is the information readily available and accessible to users. Active data includes word-processing documents, spreadsheets, databases, e-mail messages, electronic calendars, and contact managers. A listing of active data files can be viewed easily through file manager programs such as Windows Explorer or through "list file" commands in DOS.

File Clones are backup files that may be automatically created and periodically saved when an active file is being worked on by a user. These files are created and saved in order to help users

recover data lost due to a computer malfunction; usually, the file clones are not stored in the same directory as the active file. File clones are useful because they create a copy, or multiple copies, of a document that users may not erase and may not be aware exists. On most networked systems, file clones are saved to the user's hard drive rather than to a centralized network file server. As a result, a document (or some version of it) that was purged from the file server may exist as a file clone on a user's hard drive.

Backup Data is information copied to removable media in order to provide users with access to data in the event of a system failure. Networks are normally backed up on a routine schedule, while individual users tend to back up (or not) on an informal basis. Network backups normally capture only the data saved on the centralized storage media (e.g., the file server) and do not capture all the data stored on individual users' hard drives.

Monthly backups may be kept anywhere from several months to many years. In most businesses, as the backup schedule progresses, the media are "rotated"—recycling older media back into the rotation queue as new backups are created. Backups provide a historical snapshot of the data stored on a system on the particular day the backup was made. Reviewing a series of backup sessions can provide a wealth of information about how a particular matter progressed over several weeks or months. The difficulty with using backup data is that the media (usually tapes) hold a large amount of data that is only loosely organized.[1] Consequently, finding relevant data requires restoring a tape, viewing its directories, and searching within the directories for specific files. If the file is not on the tape, the process must be repeated for each backup tape. Reviewing a large number of backup tapes can be an expensive and time-consuming process.

Residual Data is information that appears to be gone, but is still recoverable from the computer system. It includes "deleted" files still extant on a disk surface and data existing in other system hardware such as buffer memories of printers, copiers, and fax machines. How is deleted data recoverable? In most operating sys-

[1]A single backup tape can store the equivalent of 1 to 5 million written pages of information. In order to fit as much data as possible on a tape, backup programs normally "compress" the data. To access the data on backup tapes often requires decompressing the data and restoring it to a host drive. Because most organizations do not have enough drive space to restore backups without overwriting current data, parties may need to find additional drive space to restore the data.

tems, the term "deleted" does not mean destroyed; rather, when a file is deleted, the computer makes the space occupied by that file available for new data. Reference to the "deleted" file is removed from directory listings and from the file allocation table; but the bits and bytes that make up the file remain on the hard drive until they are "overwritten" by new data or "wiped" through use of utility software. The result is that a file appears to have been deleted, but may still be recovered from the disk surface.

Until data are overwritten or wiped, they can be restored through use of undelete or restore commands contained in many systems' operating software or through specialized programs. As "deleted" files may be overwritten when a new file is saved, new software is loaded, or unused space is wiped through routine system maintenance (e.g., data compression and disk de-fragmentation or optimization routines), the amount and type of residual data that can be recovered will vary. In the case of a partially overwritten file, pieces of the file or "file fragments" may also be recovered.

Residual data can be buried in a number of other places on disks and drives. Forensic specialists have tools that allow them to examine the entirety of a drive for residual data. It is therefore important to note that simple copy commands will not capture residual data. Additionally, most commercial backup programs do not capture deleted files. As will be discussed later in this chapter, gathering this information requires creating an image copy of the drive at issue. Finally, it is important to note electronic mail messages are managed differently than data files, and chances of recovering deleted messages are less likely.

Electronic Mail

E-mail has several characteristics that make it an excellent source of evidence:

- Most people use e-mail informally and candidly.
- Many people believe that e-mail messages are impermanent.
- E-mail is more difficult to get rid of than most users believe. Permanently deleting messages on most e-mail systems is usually a two-step process and many users complete only the first step.
- E-mail is easily copied and forwarded, thus making distribution of a message nearly impossible to control.
- Undeleted e-mail may be captured on system backups.

Although business use of e-mail is skyrocketing, guidelines for its use are lacking. A recent survey conducted by the Cohasset Group[2] revealed that 59 percent of organizations using e-mail did not provide policies concerning either content control or retention periods for saving messages.

Background Information

Although data files and e-mail are often targeted for evidence, they are not the only information that can be gleaned from a computer system. Computer systems can provide a wealth of background information, which may be valuable evidence or can be used to further develop the facts of a case.

Audit Trails and computer logs create an electronic trail regarding network usage. Typically, an audit trail contains information about who, when, where, and how long a user was on the system. Also recorded may be information about who modified a file last and when the modification was made. An audit trail may also indicate when and by whom files were downloaded to a particular location, copied, printed out, or purged.

In addition to using a network's audit trail, an increasing number of companies are also installing software designed to monitor employees' use of company computers. This software records information such as programs used, files accessed, e-mail sent and received, and Internet sites visited.

Access Control Lists limit users' rights to access, view, and edit various files. Access rights often depend on the employee's particular job duties and position in the company. For example, the access rights for a company's billing files may be limited to the accounting department and senior management. Moreover, different personnel may have different kinds of access rights. For example, the accounting department may have read and write access, whereas managers may have read-only access. If litigation centers on a particular file or group of files, identifying who had access rights to the files and the type of access each person was allowed can establish data ownership/authenticity of files. Network security systems allow system administrators to set and maintain varying levels of access to users on the system.

---

[2]Robert F. Williams, "Electronic Records Management Survey: A Call to Action." Cohasset Associates Inc. Co-sponsored by ARMA International and AIIM.

Non-Printing Information carried by most data files is another excellent source of information. The most common example is the date and time stamp the operating system puts on every file. Some word-processing programs store revisions to documents, allowing a viewer to follow the thought process of the author as a document is edited. Some word-processing packages allow users to insert "hidden" or non-printing comments. Many schedule programs track who made changes to a calendar and when the changes were made. This information may never appear in hard copy form, but may be found in the electronic version.

The following checklist summarizes the different types of media that should be collected during discovery.

Electronic Media Collection Checklist

Data Files*
__office desktop computer/workstation
__notebook computer
__home computer
__computer of personal assistant(s)/secretary/staff
__palmtop devices
__network file servers/mainframes/computers

Backups
__systemwide backups (monthly/weekly/incremental)
__disaster recovery backups (stored off site)
__personal or "ad hoc" backups (look for disks and other portable media)

Other Media Sources
__tape archives
__replaced/removed drives
__floppy disks and other portable media (e.g., CDs, thumb drives)

*To ensure that all data, including residual data, are captured, an image copy is recommended when copying data from local computer hard drives.

The following scenario illustrates how computer-based evidence, in all its incarnations, can be scattered throughout a company's computer system.

Case Scenario

To remain in business, VeryTech Corporation needed to launch a new version of their software. The announcement of a June 2003 ship date for the new release gave a welcome boost to shareholders. Stock prices soared, and optimism was high. VeryTech principals and directors did well. The June 2003 date came and went, and VeryTech's software was still in only beta-stage. When stock prices plummeted, shareholders filed suit. Aggressive attorneys for the plaintiffs incorporated requests for computer-based files in their discovery strategy. The following activities had created computer-based documents that provided the jury with a compelling picture of investment security fraud:

Jed Roberts, principal and CEO of VeryTech, wrote a memo to Susan Davis, VeryTech's public relations director, encouraging her to accelerate work on May's media campaign regarding the June ship date. Using the hidden text feature of his word-processor, Roberts wrote this side note to his secretary: "Delivering smoke and mirrors to the press is like carrying coals to Newcastle." He gave his secretary the memo on his thumb drive. The edited copy, generated by his secretary and e-mailed to Davis, did not contain the side note.

In May, Steve in Research and Development was sending his own messages to fellow staff members, making use of VeryTech's e-mail system. "Even if we triple our staff (which you know we won't), we're never going to make it." This e-mail message, sent at 11:00 p.m., was swept into the monthly backup created at midnight.

Jed Roberts, the CEO, automatically received electronic status reports generated using Project Manager software. Project Manager reports included Gantt charts showing critical path, as well as resource, cost, and project status. The Project Manager software and data were stored on R&D's file server and backed up weekly.

Susan Davis in public relations wrote a memo to Jed Roberts expressing her growing alarm that VeryTech's promises to the press were untrue. Before printing the memo, she had second thoughts and deleted it from her hard drive.

The five "smoking gun" documents found were as follows:

(1) hidden text was revealed when the file on the thumb drive was reviewed, (2) Steve's e-mail message was stored on the monthly backup tape, (3) Jed Roberts' hard drive contained saved status reports, (4) Project Manager files were stored on the weekly backup tapes, and (5) the deleted file on Susan Davis' hard drive was recovered.

Gathering Computer-Based Information

There is no question that information stored on computers is discoverable. The Federal Rules of Civil Procedure (and most state rules) include in their definitions of documents "data compilations from which information can be obtained" and permit parties to "copy, test, or sample any tangible things" within the

scope of discovery.[3] Courts facing the issue have uniformly ruled that computer information is discoverable.[4] Courts have further held that deleted files on a party's hard drive are discoverable and that an expert must be allowed to retrieve all recoverable files.[5] In one case, a party failing to produce properly requested data was subject to sanctions, even though the data were not available in a hard copy form.[6] More recent discovery opinions and case law paint a grimmer picture of electronic discovery error, timing disasters, and spoliation issues.[7]

There are three key steps to effectively gather computer-based evidence: (1) preserve existing electronic evidence, (2) get an overview of the systems and users involved, and (3) preserve the chain of custody. Each of these is discussed in the following sections:

Preserve Existing Electronic Evidence

Every time a user enters new data, loads new software, or performs routine system maintenance, some data may be modified permanently. In fact, the simple act of turning a computer off or on will change the information on that computer. To preserve the maximum amount of information, you must put all parties (including your own client) on notice that information contained on computer systems is relevant to the dispute and that all parties must take immediate steps to preserve such information.

The first part of the notice should outline the type of information to be preserved:

---

[3]Fed. R. Civ. P. 34(a); Federal Rule 26 also expressly includes data compilations in the items that must be either produced or particularly described in the parties' initial disclosures. Fed. R. Civ. P. 26(a)(1).

[4]See, e.g., *Anti-Monopoly, Inc. v. Hasbro, Inc.*, 94 Civ. 2120, 1995 U.S. Dist. LEXIS 16355 (S.D.N.Y. 1995) ("today it is black letter law that computerized data is discoverable if relevant"); *Santiago v. Miles*, 121 F.R.D. 636, 640 (W.D.N.Y. 1988) ("A request for raw information in computer banks is proper and the information is obtainable under the discovery rules."); *In re Brand Name Prescription Drug Antitrust Litigation*, 94-C-987, M.D.L. 997 (N.D. Ill. 1995) (e-mail is discoverable); *Seattle Audubon Society v. Lyons*, 871 F. Supp. 1291 (W.D. Wash. 1994) (ordering production of e-mail).

[5]*Easley, McCaleb & Assocs., Inc. v. Perry*, No. E-2663 (Ga. Super. Ct. July 13, 1994). Such access, however, is not unlimited. In two decisions, access to a litigant's computer system was denied because the party seeking discovery could not show a likelihood that relevant information could be retrieved. *Strausser v. Yalamachi*, 669 So. 2d 1142, 1144–45 (Fla. App. 1996); *Fennell v. First Step Design, Ltd.*, 83 F.3d 526 (1st Cir. 1996).

[6]*Crown Life Ins. v. Craig*, 995 F.2d 1376 (7th Cir. 1993).

[7]See www.forensics.com for in-depth, quarterly updates of recent case law.

- electronic mail and information about electronic mail (e.g., message contents, header information, and logs of electronic mail system usage)
- data files created by word-processing, spreadsheet, or other application software
- databases and structural information about the databases; network activity logs and audit trails
- electronic calendars, telephone logs, and contact managers

Explain that this information may exist inactively in places such as network file servers, mainframe computers or minicomputers, standalone PCs, and network workstations. Data may also reside on off-line data storage media including backups and archives, CDs, DVDs, floppy disks, tapes, and other removable electronic media.

The second part of the notice should specify that no potentially discoverable data should be deleted or modified, and procedures that may affect such data should not be performed unless all potentially discoverable data has been copied and preserved. The key is to make clear that the data to be preserved include not just active data, but also archival, backup, and residual data.

With respect to system users who may have discoverable information on their computers, no new software should be loaded and no data compression, disk de-fragmentation, or optimization routines run until there has been an inspection or image copies of the hard drive have been made. Note, however, that most network servers, mainframes, and minicomputers have disk optimization routines that must remain operational. As a consequence, the instruction regarding data compression and disk de-fragmentation is best suited to preserving evidence on the hard drives of desktop and notebook computers.

With respect to backup systems, ask that the rotation and reuse of backup media cease until relevant data can be copied. Requesting parties should ask that existing tapes be held aside and not recycled. Parties should also be instructed not to dispose of any electronic media storage devices that are being replaced due to failure or system upgrade.

Remember that your client will also be expected to follow the same steps that you are instructing your opponent to follow.

Get an Overview of the Systems and Users Involved

Effectively planning and responding to discovery requires you to know how the computer systems are structured. Indeed, your notice letter will be more effective if you can gather some information beforehand on your opponent's system. For both the opponent's and your client's systems, you will need the following information:

- *System configuration.* This includes the types of computers and other hardware used, desktop and network operating systems, and the type of network and communications software and hardware used.
- *Application software and utilities.* Ask for the name and version of all application software and all utilities used on the system; this includes both commercially available applications and custom applications. If you are interested in e-mail, find out what types of e-mail programs are used and ask for current lists of system users.
- *Backup procedure and frequency.* This includes the name and version of the backup software used, the type of media used, the schedules used for incremental and full backups, the length of time backups are kept, and how often backup media are reused. Also ask how the backup media are indexed and stored.
- *Logons and passwords.* Ask about any encryption programs that may be used to "lock" sensitive information. This information will help you when reviewing the data collected and will also assist in the authentication process.

In addition to the discovery directed at the computer system, every witness must be questioned about his or her computer use. Users' computer sophistication varies widely. Knowing how each witness uses his or her computer and organizes and stores data may lead to sources of data not revealed by the discovery directed at general system usage. This discovery should also focus on the secretaries and other people assisting key witnesses.

Perhaps the most overlooked source of electronic evidence is users' home computers. Data can end up on home computers in a number of ways. Data can be transferred to and from the workplace on removable media, via e-mail attachment, or employees

may be able to log on to the company network from outside of the office. With direct access, the home computer acts much like the employee's office workstation. Regardless of how data are transferred, the critical point is to find out whether the witness works from home and if there are data on a home computer.

Witness and client interviews, carefully crafted interrogatories, and requests for production are all excellent ways to gather information regarding the systems. Another extremely useful tool is a section 30(b)(6) deposition of a party's information systems department. The deposition serves two basic purposes: first, it provides the system overview needed to effectively undertake further discovery; second, as with all custodian of records depositions, it helps establish the foundation needed for using the computer records as evidence.

Checklist for System Discovery

__The layout of the computer system, including the number and types of computers, and the types of operating systems and application software packages used.
__The type of electronic mail system, including software used, the number of users, the location of mail files, and password usage.
__The structure of any network, including the configuration of network servers and workstations, and the network operating system.
__Specific software used. This includes software applications for things such as calendars, project management, accounting, word processing, and database management. It also includes industry-specific programs, proprietary programs, encryption software, and utility programs.
__The personnel responsible for the ongoing operation, maintenance, expansion, and upkeep of the network.
__The personnel responsible for administering the e-mail system.
__The personnel responsible for maintenance of computer-generated records and the manner in which such records are organized and accessed.
__Backup procedures used on all computer systems in the organization. This should include descriptions of all devices (e.g., tape drives) and software used to create backups, the personnel responsible for conducting the backups, what information is backed up, backup schedules, and tape rotation schedules.

__The process for archiving and retrieving backup media both on and off site.
__The procedures used by system users to log on to computers and into the network. This includes use of passwords, audit trails, and other security measures used to identify data created, modified, or otherwise accessed by particular users.
__How shared files are structured and named on the system.
__Routines for archiving and purging different types of data.

Preserve the Chain of Custody

A chain of custody verifies that information copied was not altered in the copying process, and has not been altered during analysis. A solid chain of custody is essential to authenticating computer-based evidence copied from your opponent.

- *No information was added or harmed.* Before doing anything else, software and media you intend to use must be virus checked with up-to-date virus checking utilities. It also means that before examining any media or making any copies, the originals need to be "write-protected" so that no data are added or changed during inspection and copying.
- *Make a complete copy.* Accurately copying all data on a drive requires making a sector-by-sector copy of the drive. A sector-by-sector copy (also called an evidentiary image copy) creates a mirror image of the drive being copied, thus capturing all data, including residual data, on the drive surface. Simply making a file-by-file backup captures only active data and may be deemed inadequate for evidentiary purposes.[8]
- *Use a reliable copying process.* In copying data, there are a number of different programs and media that can be used. The following criteria must be met: (1) it must meet industry standards for quality and reliability, (2) it must be capable of independent analysis, and (3) it must create tamper-proof copies. Keep in mind that any copies must be able to withstand cross-examination by your opponent's expert as well as judicial scrutiny.

[8]See *Gates Rubber Co. v. Bando Chemical Indus., Ltd.*, 167 F.R.D. 90, 112 (D. Colo. 1996) (the court criticized a party's expert for not making an image copy, concluding that when collecting evidence for judicial purposes a party has "a duty to utilize the method which would yield the most complete and accurate results").

- *Secure all media.* All copies should be tamper-proof and any original media collected as evidence should be write-protected or otherwise made tamper-proof. All media (copies and originals) should be labeled by time, date, and source, and stored in a secure place. Forensic analysis of the information collected should be done on a working copy created from the secure copy whenever possible.

Summary

Businesses and individuals now use computers to store and communicate information, more than 90 percent of which is never printed on paper. Neglecting discovery aimed at computers and computer-based records thus greatly increases the odds critical evidence will be overlooked. Computer discovery does not require a computer expert. Rather, what it requires is a fundamental understanding of what kinds of information exist, where this information may be stored, and how to ask the questions that will lead you to it.

II. Electronic Discovery: The Current Legal Landscape

Theodore O. Rogers, Jr.*

Introduction

Courts, litigants, and commentators have increasingly been grappling with how to apply discovery rules that were crafted in an age of paper records to massive amounts of electronic data maintained by corporations and individuals alike. Among the difficult topics raised when discovery of electronic data is at issue are the appropriate standards for retention of electronic information, the most efficient means for determining the appropriate scope of discoverable information and, of obvious importance, which party should bear the substantial costs of retention, retrieval, and review of electronic data.

---

*Theodore O. Rogers is a partner with Sullivan & Cromwell, LLP, New York, New York.