

The Essentials of Computer Discovery

Larry Johnson

Director of Electronic Discovery Services

Fios

and Joan Feldman

Computer Forensics Inc.

When it comes to computers, most lawyers are like everybody else: they may know how to use them but don't know how they work – or what surprising secrets they may hold. So when documents and other information stored on computers become the focus of discovery efforts in litigation, you'll most likely need to hire a computer forensics expert to help you through the electronic jungles. Before you do that, though, it's a good idea to get a general overview of what those jungles look like; what the opportunities and pitfalls are; and what discovery tactics work best in this unfamiliar terrain.

Let's first look at the nature of the beast itself, computers, and the wealth of information they hold, some of which is not obvious or readily visible.

A discovery strategy will start with the most obvious sources of information on computers. These are the "data files" created by software applications: word-processed documents; reports generated by databases; spreadsheets, and e-mail.

Often you can obtain the same documents in printed form through traditional discovery methods, but not always. It has been estimated that up to 30% of the typical documents generated by businesses remain in electronic form only, especially e-mail.

Discovery tactics that do not include the opposition's electronic data may overlook some very important evidence, again primarily e-mail, which tends to be casual, candid – and careless.

With the readily available data files, software can search through gigabytes of data for key words or phrases much faster and more efficiently to find key documents than one could with hours of brain-numbing sifting done by an army of paralegals. Because electronic data offer this ease of search, courts have invoked Rule 1 of the federal Civil Rules of Procedure ("[the Rules] shall be construed and administered to secure the just, speedy, and inexpensive determination of every action") to let litigants surf electrons swiftly rather than forcing them to rummage through boxes of paper.

After the data files there are the so-called "replicant data" or "file clones," normally just so much garbage created in the course of normal computer use. But, ah, what useful garbage this can be!

Many software manufacturers build in automatic backup features that create and periodically save copies of a file being worked on by a user. These help users recover

data lost due to a computer malfunction (e.g. system crash or power loss). 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 would not normally erase and are usually not aware of. 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 still exist as a file clone on a user's hard drive. So an ideal discovery plan should not consider redundant whatever is on a user's hard drive, simply because "final version" files are on the network server.

Besides these useful clones or relics, there are also backup data to consider. Backups are maintained and stored, again 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.

A typical network backup schedule would be as follows: A full weekly backup (usually done on Fridays) with incremental backups to capture all new or changed data made on all other days. At the end of the month, the last full weekly backup will be pulled from its rotation queue and saved as the monthly backup.

Reviewing a series of backup tapes 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. 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. With a large number of backup tapes this can be an expensive and time-consuming process.

Let's take the next step downwards into the bowels of a computer – and this is where you can often find the "smoking guns" – where there are what are known as the "residual data." This is information that appears to be gone but is still recoverable from the computer system. It includes "deleted" files still extant on hard drives and data existing in other system hardware such as buffer memories of printers, copiers and fax machines.

How are deleted data recoverable? In most operating systems, the term "deleted" does not mean destroyed. Rather, when a file is "deleted" the computer simply makes the space occupied by that file available for new data. Reference to the deleted file is removed from directory listings and from what is known as the file allocation table, a kind of internal master index of all files on a hard drive. But the bits and bytes that make up the file remain on the hard drive until they are overwritten by new data or "wiped" (truly deleted with a bunch of zeroes) through use of utility software.

The result is that to the user a file appears to have been deleted, but it may still be recovered from the hard drive. 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, 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 -- 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. This usually means that the specialist will gather this information by literally creating a mirror image copy of the drive at issue. This process can be pricey, but without it you may miss out on information the opposition had hoped to hide from you.

Finally, you should know that there's not only content to be ferreted out of a computer, computers also contain *information about the information* they store. For example, every computer contains its own kind of audit trail. All files are date and time-stamped, and networked computers contain computer logs about who, when, where and how long a user was on the system. That information can be golden in the face of a witness's denial that he or she had no knowledge of certain facts, or was off in Hawaii on the date the document was created and discussed by corporate head honchos.

Also recorded within a computer's own self-maintenance system 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.

And you thought *1984* was a work of fiction? Big Brother -- your computer -- is here now, watching and recording your every move.

You may be getting more of a primer on what computers are capable of than you care to know. But hold on, we are almost done.

Networked computers allow a large number of people to share information and keep it all in one central place, the Numero Uno computer known as the network server. Big companies will have more than one server, with backup servers to jump into action if any of the others should malfunction. Networks have their own logic that goes beyond any one single workstation. For example, there's what is known within networks as "access control." Access control lists limit users' rights to access, view and edit various

files otherwise available on a network. 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 access rights. Thus, the accounting department may have read-and-write access to some or all files, whereas managers may have read-only access. They can look but they can't touch.

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.

All right, we've seen some of the electronic jungle, now what to do with it as a lawyer?

That depends on whether your client is in litigation or hopes to keep clear of it. And if you are in litigation, it depends on whether your client is the plaintiff or the defendant.

Let's look at litigation first, since this is usually when the business lawyer is first made aware of problems with discovery and his or her clients' computers.

If your prosecuting a case, you have to be concerned that the adversary is going to cover its tracks (or at least those of its employees with jobs at stake being so motivated), and you can be sure that defense counsel will raise arguments about how intrusive, over-reaching and burdensome it is to "shut down" a company's computers just so you can forage around in them, often in areas that are confidential or privileged.

From the plaintiff's perspective, you and your forensics expert have to meet the challenge that making mirror image copies of hard drives is not intrusive, and that the process is technically an easy and straightforward task. You must be prepared to enter into a protective order for your client that will allow the opposition to claim (and have reviewed *in camera* by the judge if necessary) that certain identified files are off limits. This is nothing more than the usual arduous process associated with protective orders for printed documents, and in fact, with electronically stored media, the process is a much easier task than it is with paper.

In many cases, the hardest part about getting a protective order is educating the judge. He or she may want to keep computers out of the discovery process out of fear that motions relating to computers will be too technical or difficult to decide. The prudent lawyer comes to the discovery hearing with his or her forensics expert in tow to overcome judicial bias or insecurity about the potentially intimidating aspects of computers.

The cheapest way for the plaintiff's lawyer to preserve evidence is to send the opposition a written notice to preserve that evidence. How effective such a notice is, however, is another matter. A more potent approach is to immediately serve a Request for Production of Documents and Things to the opposing party requesting that you wish to "copy" all the adversary's equipment at a time least disruptive to the adversary's business, then to follow that up as soon as possible with a Rule 30(b)(6) deposition of the person most knowledgeable about the opposing party's computer system. Anything destructive done by the opposing party following the initiation of such discovery efforts will almost certainly result in the imposition of severe sanctions by the judge.

The Request for Production of Documents and Things should be detailed and specific. In it you should explain that the information sought may exist actively 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, floppy diskettes, tapes and other removable electronic media.

The Request should also specify that no potentially discoverable data should be deleted or modified and that procedures that may affect such data should not be performed unless all potentially discoverable data have 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.

If you are truly concerned that critical data may be removed before you have a chance to conduct discovery, it may be necessary to obtain a restraining order to prevent any destruction of evidence, followed by a hearing to obtain a more longer-lasting preliminary injunction. To obtain that kind of relief, you normally have to make a showing that the harm you fear is imminent; that it will cause irreparable harm to your client's detriment if the data are purged; and that you have a likelihood of success on the underlying claims you are prosecuting on behalf of your client.

It may take such extraordinary injunctive relief to guarantee the integrity of the computers you want to examine. With respect to system users that may have discoverable information on their computers, the restraining order should state that no new software should be loaded and no data compression and 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 upgrades.

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

For defense lawyers seeking to prevent a wholesale onslaught on their clients' information systems, the best defense is a comprehensive (and hopefully agreed to) protective order that allows for an orderly inspection of both sides' computers (assuming such reciprocity is relevant). The protective order should provide defense counsel with adequate opportunity to remove from "mirrored" copied hard disks those files that are work product, privileged or confidential trade secrets of the client. Defense counsel should also be entitled to conduct discovery as to the methods and practices of any expert hired by plaintiff counsel to assure data integrity and chain of custody of any evidence acquired from his or her client.

For litigants with limited resources, with only one discovery arrow to fire at the opposing party's computer data, by all means focus on the e-mail. E-mail is now easy and ubiquitous. A recent survey by the Electronic Marketing Association projected that between by 2000 the number of e-mail users will double, surpassing the 108 million mark. A recent survey of e-mail users conducted by Georgia Tech University's Graphic Visualization & Usability Center showed that 25% of respondents had home e-mail and 56% of respondents had more than one e-mail account.

As a business tool, e-mail is even more commonplace. Statistics show that 90% of organizations having over 1,000 employees use e-mail and that 40% of all workers now use e-mail on the job.

E-mail has several characteristics that make it an excellent source of evidence. Many people believe that e-mail messages are ephemeral. But e-mail is more difficult to get rid of than most users believe. For most e-mail systems, permanently deleting messages is usually a two-step process and many users only complete only the first step. Also, e-mail is easily copied and forwarded thus making distribution of a message nearly impossible to control. And finally, undeleted e-mail may be captured on system backups even if the user later decides to erase incriminating messages.

It's on e-mail that people let their guards down, where gossip is spread, sexist and racist jokes shared, where defamation is commonplace, and where people confide things to friends, thinking they are unobserved. E-mail is where people tell the unadorned truth, often with embarrassing results later if it pops up in court.

For the lawyer who simply wishes to advise his/her client to be aware of potential liabilities down the road, it is an almost impossible task to control the flow and proliferation of e-mail. Still, an employer can and should implement and post prominently a policy that prohibits private use of e-mail, and to remind employees that all e-mails are like postcards, available for all to read, including the employer. From time to time, employee e-mail should be monitored for compliance. Old e-mail should be routinely deleted and expunged.

Programs, such as MIMESweeper by Content Technologies, Inc. (info@mimesweeper.com), can routinely scan all employee e-mail for offensive or illegal e-mail content.

While such policies and tools may seem obtrusive, there is a growing body of case law in the U.S. and elsewhere that can come to haunt employers if they do anything to create -- or acquiesce in -- the impression that an employee has a right of privacy to his or her e-mail or other electronic files.

Nor is prevention of misconduct going to be helped much with passwords and encryption. Passwords usually apply only to the user's entry into a computer. Once the computer is up and running, the files created are not ordinarily password-protected. A court can order a person to open up his or her computer, just as it could order a garage or warehouse unlocked. Further, in many instances it is not that difficult to bypass passwords through software that can "hack" a system.

Encryption software and hardware exist to make files less easily accessible, but again a user can be compelled by court order to decrypt files so that they can be read and used in pre-trial discovery.

Thus, businesses should know that security measures to keep out electronic trespassers do not afford any shield in the context of litigation.

A good resource for helping businesses to limit their liability exposure is Michael R. Overly's *E-policy - How to Develop Computer, E-Mail, and Internet Guidelines to Protect Your Company and Its Assets*, published by the American Management Association, 1601 Broadway, New York, NY 10019, 1999, ISBN 0-8144-7996-0, www.amanet.org.

Whether advising a client about discovery options in litigation or how it can reduce its risks, it's no longer an option for today's business lawyer to be ignorant about computers, not as to how they work, but about what kinds of information they can potentially hold, for better or worse.

###