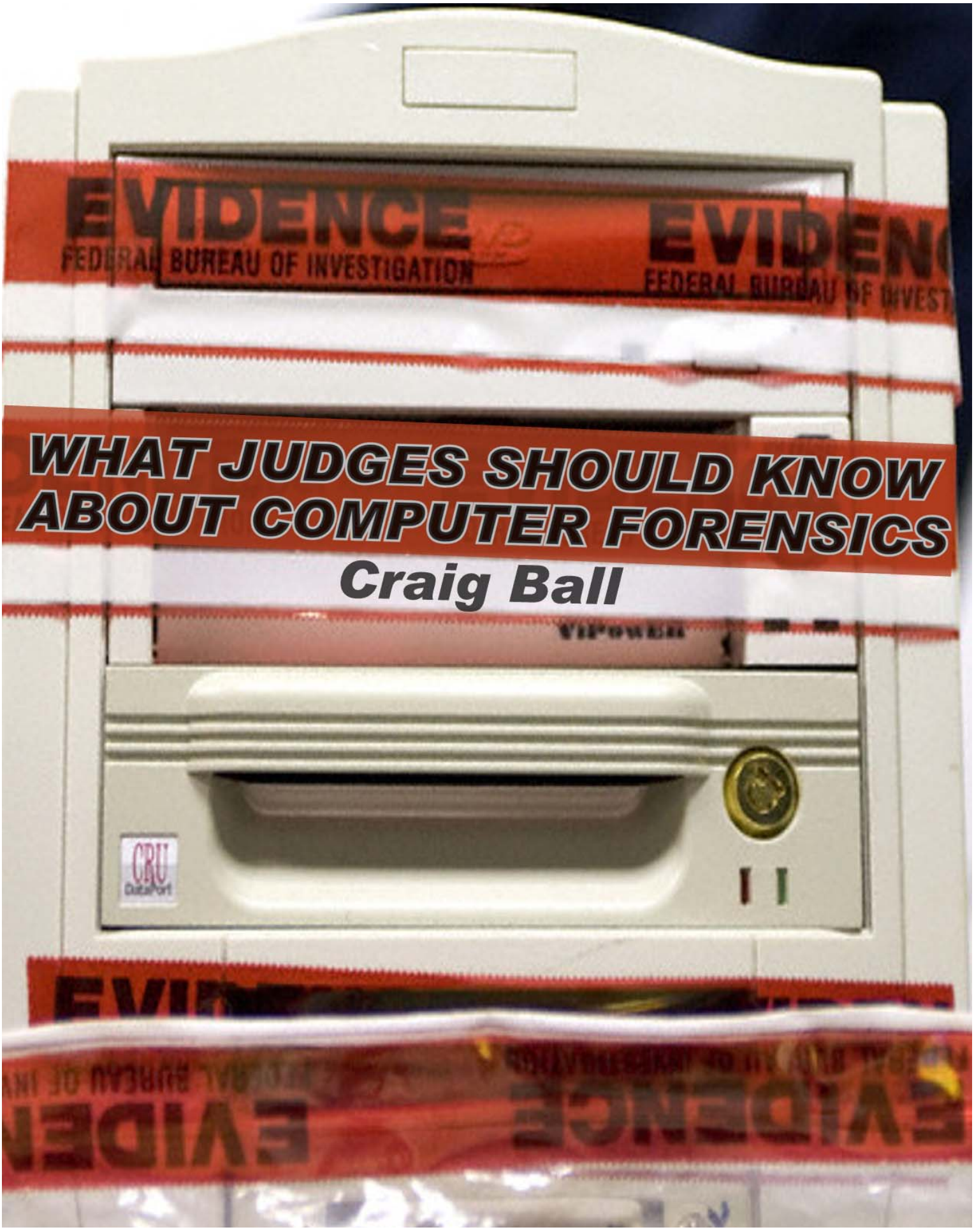
EVIDENCE
FEDERAL BUREAU OF INVESTIGATION
WHAT JUDGES SHOULD KNOW ABOUT COMPUTER FORENSICS
Craig Ball

What Judges Should Know About Computer Forensics

Craig Ball[1]

Courts increasingly see motions by litigants seeking access to an opponent's computers for the purpose of conducting a computer forensic examination. The impetus may be allegations of discovery abuse, stolen intellectual property, spoliation, forgery, network intrusion, child pornography, piracy, discrimination or a host of other claims.

When bits and bytes are involved, it can be hard to know if the proposed examination is reasonable and necessary or an abusive fishing expedition.

This article looks at some of the fundamentals of computer forensics to help judges weigh the need and burden of acquisition and examination. It addresses, *inter alia*, what computer forensics can and cannot accomplish and flags common errors made by parties and the courts in ordering such examinations.

Table of Contents

[1] The author gratefully acknowledges the invaluable editorial contributions of his spouse, Diana Ball, and of esteemed colleagues, Sharon Nelson and John Simek of Sensei Enterprises, Inc., for their helpful suggestions.

What is Computer Forensics?

A computer's operating system or **OS** (e.g., Windows or Vista, Mac or Linux) and installed software (**applications**) generate and store much more information than users realize. Some of this unseen information is **active data** readily accessible to users, but requiring skilled interpretation to be of value in illuminating human behavior. Examples include the data *about* data or **metadata** tracked by the OS and applications, but not displayed onscreen. For example, Microsoft Outlook records the date a Contact is created, but few of us customize the program to display that "date created" information.

Other active data reside in obscure locations or in coded formats less readily accessible to users, but enlightening when interpreted and correlated. Log files, hidden system files and information recorded in non-text formats are examples of **encoded data** that may reveal information about user behavior.

Finally, there are vast regions of hard drives and other data storage devices that hold **forensic data** even the operating systems and applications can't access. These "data landfills," called **unallocated clusters**[2] and **slack space**[3], contain much of what a user, application or OS discards over the life of a machine. Accessing and making sense of these vast, unstructured troves demands specialized tools, techniques and skill.

Computer forensics is the expert acquisition, interpretation and presentation of the data within these three categories (**Active**, **Encoded** and **Forensic** data), along with its juxtaposition against other available information (e.g., credit card transactions, keycard access data, phone records and voice mail, e-mail, documents and instant message communications and texting).

In litigation, computer forensics isn't limited to personal computers and servers, but may extend to all manner of devices harboring electronically stored information (**ESI**). Certainly, external hard drives, thumb drives and memory cards are routinely examined. *When relevant*, information on cell phones, cameras and even automobile navigation systems and air bag deployment modules may be implicated. The scope of computer forensics—like the scope of a crime scene investigation—should be reasonably tailored to the available evidence and issues before the court.

How Does Computer Forensics Differ from Electronic Discovery?

Computer forensics is a non-routine subcategory of "e-discovery." In simplest terms, electronic discovery addresses the ESI accessible to litigants; computer forensics addresses the ESI accessible to forensic experts. However, the lines blur because e-discovery often requires litigants to grapple with forms of ESI—like backup tapes—traditionally regarded as inaccessible, and computer forensics relies on information readily accessible to litigants, such as file modification dates.

The principal differentiators are **expertise** (computer forensics requires a unique skill set), **issues** (most cases can be resolved without resorting to computer forensics, though some will hinge on matters that can only be resolved by forensic analysis) and **proportionality** (computer forensics injects issues of expense, delay and intrusion). Additionally, electronic discovery tends to address evidence as discrete information items (documents, messages, databases), while computer forensics takes a more systemic or holistic view of ESI, studying information items as they relate to one another

[2] Unallocated clusters are storage areas flagged by the file system as available to hold data. When these have been previously used for data storage, their former contents linger until overwritten by new data.

[3] File slack space is the excess storage space between the end of a file and the end of the final cluster in which the file is stored. Slack space may hold fragments of deleted files.

and in terms of what they reveal about what a user did or tried to do. And last, but not least, electronic discovery deals almost exclusively with existing ESI; computer forensics tends to focus on what's gone, how and why it's gone and how it might be restored.

When to Turn to Computer Forensics

Most cases require no forensic-level computer examination, so courts should closely probe whether a request for access to an opponent's machines is grounded on a genuine need or is simply a fishing expedition. When the question is close, courts can balance need and burden by using a neutral examiner and a protective protocol, as well as by assessing the cost of the examination against the party seeking same until the evidence supports reallocation of that cost.

Certain disputes fairly demand forensic analysis of relevant systems and media, and in these cases, the court should act swiftly to support appropriate efforts to preserve relevant evidence. For example, claims of data theft may emerge when a key employee leaves to join or become a competitor, prompting a need to forensically examine the departing employee's current and former business machines, portable storage devices and home machines. Such examinations inquire into the fact and method of data theft and the extent to which the stolen data has been used, shared or disseminated.

Cases involving credible allegations of destruction, alteration or forgery of ESI also justify forensic analysis, as do matters alleging system intrusion or misuse, such as instances of employment discrimination or sexual harassment involving the use of electronic communications. Of course, electronic devices now figure prominently in the majority of crimes and many domestic relations matters, too. It's the rare fraud or extramarital liaison that doesn't leave behind a trail of electronic footprints in web mail, online bank records and cellular telephones. For further guidance on circumstances justifying direct access to an opponent's ESI, see, e.g., *Ameriwood Ind., Inc. v. Liberman*, 2006 WL 3825291 (E.D. Mo. Dec. 27, 2006).

Balancing Need, Privilege and Privacy

A computer forensic examiner sees it all. The Internet has so broken down barriers between business and personal communications that workplace computers are routinely peppered with personal, privileged and confidential communications, even intimate and sexual content, and home computers normally contain some business content. Further, a hard drive is more like one's office than a file drawer. It may hold data about the full range of a user's daily activity, including private or confidential information about others. Trade secrets, customer data, e-mail flirtations, salary schedules, Internet searches for escort services, bank account numbers, medical records and passwords abound.

So how does a court afford access to the non-privileged evidence without inviting abuse or exploitation of the rest? With so much at stake, courts need to approach forensic examination cautiously. Granting access should hinge on demonstrated need and a showing of relevance, balanced against burden, cost or harm. It warrants proof that the opponent is either untrustworthy or incapable of preserving and producing responsive information, or that the party seeking access has some proprietary right with respect to the drive or its contents. Showing that a party lost or destroyed ESI is a common basis for access, as are situations like sexual harassment or data theft where the computer was instrumental to the alleged misconduct.

The parties may agree that one side's computer forensics expert will operate under an agreed protocol to protect unwarranted disclosure of privileged and confidential information. Increasingly,

courts appoint neutral forensic examiners to serve as Rule 53 Special Masters for the purpose of performing the forensic examination *in camera*. To address privilege concerns, the information developed by the neutral is first tendered to counsel for the party proffering the machines for examination, which party generates a privilege log and produces non-privileged, responsive data. Use of a Special Master largely eliminates the risk of privilege waiver in unrelated litigation, a potential addressed in *Hopson v. Mayor of Baltimore*, 232 F.R.D. 228 (D. Md. 2005)

Whether an expert or court-appointed neutral conducts the examination, the order granting forensic examination of ESI should provide for handling of confidential and privileged data and narrow the scope of examination by targeting specific objectives. The examiner needs clear direction in terms of relevant keywords and documents, as well as pertinent events, topics, persons and time intervals. A common mistake is for parties to agree upon a search protocol or secure an agreed order without consulting an expert to determine feasibility, complexity or cost. Generally, use of a qualified neutral examiner is more cost-effective and ensures that the court-ordered search protocol is respected.

Who Performs Computer Forensics?

Computer forensics is a young discipline, so the most experienced examiners may be largely self-taught. Experienced examiners still tend to emerge primarily from law enforcement, but this is changing as a host of computer forensics certification courses and even college degree plans have appeared. Unfortunately, though the ranks of those offering computer forensics services are growing rapidly, there is inadequate assessment or regulation of the profession. No universally recognized standard exists to test the training, experience and integrity of forensic examiners. A few states require computer forensic examiners to obtain private investigation licenses, but don't demand that applicants possess or demonstrate expertise in computer forensics.

Computer experts without formal forensic training or experience may offer their services as experts, but just as few doctors are qualified as coroners, few computer experts hold forensic qualifications. Programming skill has little practical correlation to skill in computer forensics.

Selecting a Neutral Examiner

Ideally, the parties will agree upon a qualified neutral. When they cannot, the court might:

1. Require the parties to designate examiners they deem qualified, then have the partisan examiners agree upon a third party neutral examiner;
2. Seek recommendations from other judges before whom well-qualified examiners have appeared; or,
3. Review the *curriculum vitae* of examiner candidates, looking for evidence of training, experience in court, credible professional certification, publications, bench references and other customary indicia of expertise. Checking professional references is recommended, as CV embellishment is a great temptation in an unregulated environment.

A computer forensic analyst must be able to grasp the issues in the case and, where indicated, possess a working knowledge of privilege law.

What Can Computer Forensics Do?

Though the extent and reliability of information gleaned from a forensic examination varies, here are some examples of the information an analysis can uncover:

1. Manner and extent of a user's theft of proprietary data;

2. Timing and extent of file deletion or antiforensic (e.g., wiping software) activity;
3. Whether and when a thumb drive or external hard drive was connected to a machine;
4. Forgery or alteration of documents;
5. Recovery of e-mail and other ESI claimed not to exist or to have been deleted;
6. Internet usage, online research and e-commerce transactions;
7. Intrusion and unauthorized access to servers and networks;
8. Clock and calendar manipulation;
9. Image manipulation; and
10. Second-by-second system usage.

What *Can't* It Do?

Notwithstanding urban legend and dramatic license, there are limits on what can be accomplished by computer forensic examination. To illustrate, an examiner generally cannot:

1. Recover any information that has been completely overwritten—even just once—by new data;
2. Conclusively identify the hands on the keyboard if one person logs in as another;
3. Conduct a thorough forensic examination without access to the source hard drive or a forensically-sound image of the drive;
4. Recover data from a drive that has suffered severe physical damage and cannot spin;
5. Guarantee that a drive won't fail during the acquisition process; or
6. Rely upon any software tool to autonomously complete the tasks attendant to a competent examination.

Supervision of Examination

A party whose systems are being examined may demand to be present throughout the examination. This may make sense and be feasible while the contents of a computer are being *acquired* (duplicated); otherwise, it's an unwieldy, unnecessary and profligate practice. Computer forensic examinations are commonly punctuated by the need to allow data to be processed or searched. Such efforts consume hours, even days, of "machine time" but not examiner time. Examiners sleep, eat and turn to other cases and projects until the process completes. However, if an examiner must be supervised during machine time operations, the examiner cannot jeopardize another client's expectation of confidentiality by turning to other matters. Thus, the "meter" runs all the time, without any commensurate benefit to either side except as may flow from the unwarranted inflation of discovery costs.

One notable exception is the examination of machines believed to house child pornography. As possession of child pornography is itself a crime, the government requires that examinations be conducted on government premises and under close supervision.; refusing to allow data to be processed in the examiner's lab.

Forensic Acquisition & Preservation

Courts are wise to distinguish and apply different standards to requests for forensically-sound *acquisition* versus those seeking forensic *examination*. Forensic *examination* and analysis of an opponent's ESI tends to be both intrusive and costly, necessitating proof of compelling circumstances before allowing one side to directly access the contents of the other side's computers and storage devices. By contrast, forensically duplicating and preserving the status quo of electronic evidence is relatively low-cost and can generally be accomplished without significant intrusion upon privileged or confidential material. Accordingly, the court should freely allow forensic preservation upon a bare showing of need.

Acquisition guards against both intentional spoliation and innocent spoliation engendered by continued usage of computers and intentional deletion. It also preserves the ability to later conduct a forensic examination, if warranted.

During the conduct of a forensic acquisition:

1. Nothing on the evidence media may be altered by the acquisition;
2. Everything on the evidence media must be faithfully acquired; and,
3. The tools and processes employed should authenticate the preceding steps.

These standards cannot be met in every situation, but the court should require the party deviating from the accepted criteria to justify the departure.

Exemplar Acquisition Protocol

An exemplar protocol for acquisition follows, adapted from the court's decision in *Xpel Techs. Corp. v. Am. Filter Film Distribs*., 2008 WL 744837 (W.D. Tex. Mar. 17, 2008):

> The motion is GRANTED and expedited forensic imaging shall take place as follows:
>
> A. The Forensic Examiner's costs shall be borne by the Plaintiff.
>
> B. Computer forensic analysis will be performed by _____ (the "Forensic Examiner").
>
> C. The Forensic Examiner must agree in writing to be bound by the terms of this Order prior to the commencement of the work.
>
> D. Within two days of this Order or at such other time agreed to by the parties, defendants shall make its computer(s) and other electronic storage devices available to the Forensic Examiner to enable him to make forensically-sound images of those devices, as follows:
>
> i. Images of the computer(s) and any other electronic storage devices in Defendants' possession, custody, or control shall be made using hardware and software tools that create a forensically sound, bit-for-bit, mirror image of the original hard drives (e.g., EnCase, FTK Imager, X-Ways Forensics or Linux dd). A bit-stream mirror image copy of the media item(s) will be captured and will include all file slack and unallocated space.
>
> ii. The Forensic Examiner should photographically document the make, model, serial or service tag numbers, peripherals, dates of manufacture and condition of the systems and media acquired.
>
> iii. All images and copies of images shall be authenticated by MD5 hash value comparison to the original hard drive(s).

iv. The forensic images shall be copied and retained by the Forensic Examiner in strictest confidence until such time the court or both parties request the destruction of the forensic image files.

v. Without altering any data, the Forensic Examiner should, as feasible, determine and document any deviations of the systems' clock and calendar settings.

E. The Forensic Examiner will use best efforts to avoid unnecessarily disrupting the normal activities or business operations of the defendants while inspecting, copying, and imaging the computers and storage devices.

F. The Defendants and their officers, employees and agents shall refrain from deleting, relocating, defragmenting, overwriting data on the subject computers or otherwise engaging in any form of activity calculated to impair or defeat forensic acquisition or examination

Forensic Examination

There is no more a "standard" protocol for forensic examination than there is a "standard" set of deposition questions. In either case, a good examiner tailors the inquiry to the case, follows the evidence as it develops and remains flexible enough to adapt to unanticipated discoveries. Consequently, it is desirable for a court-ordered protocol to afford the examiner discretion to adapt to the evidence and apply their expertise.

Although the goals of forensic examination vary depending on the circumstances justifying the analysis, a common aim is recovery of deleted data.

The Perils of "Undelete Everything"

Even if the parties agree, be wary of issuing an order directing the examiner to, in effect, "undelete all deleted material and produce it." This was the court's approach in *Ameriwood Ind., Inc. v. Liberman*, 2006 WL 3825291 (E.D. Mo. Dec. 27, 2006), where the forensic examiner was ordered to recover:

> [A]ll available word-processing documents, incoming and outgoing email messages, PowerPoint or similar presentations, spreadsheets, and other files, including but not limited to those files that were "deleted." The Expert shall provide the recovered documents in a reasonably convenient and searchable form to defendants' counsel, along with, to the extent possible, the information showing when any files were created, accessed, copied, or deleted, and the information about the deletion and the contents of deleted files that could not be recovered. Id. at 13.

Although it may seem sensible at first blush, a directive that speaks in terms of all "other" files and all "deleted" files—especially one that seeks detailed information about "the contents of deleted files that could not be recovered"--creates unrealistic expectations and invites excessive cost. Here's why:

Historically, libraries tracked books by noting their locations on index cards in a card catalog. A computer manages its hard drive in much the same way. The files are the "books" and their location is tracked by a card catalog-like index called the **file table**. But there are two key differences between libraries and computer file systems. Computers employ no Dewey decimal system, so

electronic "books" can be on any shelf. Further, electronic "books" may be split into chapters and those chapters stored in multiple locations across the drive. This is called "**fragmentation**." Computers depend on their file tables to keep track of all those file fragments

When a user hits "Delete," nothing happens to the file targeted for deletion; only the file table changes. It's as if someone tore up a card in the card catalogue. Like its literary counterpart, the deleted file is still on the "shelf," but now it's a needle in a haystack, lost among millions of unallocated clusters.

To recover deleted files, a computer forensic examiner employs three principal techniques:

1. File Carving by Binary Signature

Because most files begin with a unique digital signature identifying the file type, examiners run software that scans each of the millions of unallocated clusters for particular signatures, hoping to find matches. If a matching file signature is found and the original size of the deleted file can be ascertained, the software copies or "carves" out the deleted file. If the size of the deleted file is unknown, the examiner designates how much data to carve out. The carved data is then assigned a new name and the process continues.

Unfortunately, deleted files may be stored in pieces as discussed above, so simply carving out contiguous blocks of fragmented data grabs intervening data having no connection to the deleted file and fails to collect segments for which the directory pointers have been lost. Likewise, when the size of the deleted file isn't known, the size designated for carving may prove too small or large, leaving portions of the original file behind or grabbing unrelated data. Incomplete files and those commingled with unrelated data are generally corrupt and non-functional. Their evidentiary value is also compromised.

File signature carving is frustrated when the first few bytes of a deleted file are overwritten by new data. Much of the deleted file may survive, but the data indicating what type of file it was, and thus enabling its recovery, is gone.

File signature carving requires that each unallocated cluster be searched for each of the file types sought to be recovered. When a court directs an examiner to "recover all deleted files," that's an exercise that could take weeks, followed by still more weeks spent culling corrupted files. Instead, the protocol should specify the *particular* file types of interest based upon how the machine was used and the facts and issues in the case.

2. File Carving by Remnant Directory Data

In some file systems, residual file directory information revealing the location of deleted files may be strewn across the drive. Forensic software scans the unallocated clusters in search of these lost directories and uses this data to restore deleted files. Here again, reuse of clusters can corrupt the recovered data. A directive to "undelete everything" gives no guidance to the examiner respecting how to handle files where the metadata is known but the contents are suspect.

3. Search by Keyword

Where it's known that a deleted file contained certain words or phrases, the remnant data may be found using keyword searching of the unallocated clusters and slack space. Keyword

search is a laborious and notoriously inaccurate way to find deleted files, but its use is necessitated in most cases by the enormous volume of ESI. When keywords are not unique or less than about 6 letters long, many false positives ("**noise hits**") are encountered. Examiners must painstakingly look at each hit to assess relevance and then manually carve out responsive data. This process can take days or weeks for a single machine.

Better Practice than "Undelete" is "Try to Find"

The better practice is to eschew broad directives to "undelete everything" in favor of targeted directives to use reasonable means to identify specified types of deleted files. To illustrate, a court might order, "Examiner should seek to recover deleted Word, Excel, PowerPoint and PDF files, as well as to locate potentially relevant deleted files or file fragments in any format containing the terms, 'explosion,' 'ignition' or 'hazard.' If the examiner finds evidence of deletion of other files satisfying these criteria but which prove unrecoverable, the examiner shall, as feasible, identify such files and explain the timing and circumstances of their deletion."

Eradication Challenges

When confidential or proprietary data ends up where it doesn't belong, courts may order a computer forensic expert to eradicate it. Redaction of data from a hard drive is more challenging than most lawyers and judges appreciate. In fact, it's harder than some forensic examiners realize.

Files sought to be deleted may exist in multiple iterations, versions and fragments within the active and/or the deleted areas of the hard drive. There's rarely just one copy of any file that must be found and destroyed. As discussed above, the target file may be fragmented (stored in segments separated by unrelated information). If fragmented files were deleted, the pointers to the fragments may be lost, complicating the examiner's ability to gather all the pieces needing to be erased. Reduced to manual examination of thousands or even millions of clusters, the task quickly becomes infeasible.

When framing an eradication protocol, the court should assess the lengths a party may go to in an effort to recover and use the data. Relegating accessible copies and drafts of files irrevocably to the digital trash heap may be sufficient to forestall use by ordinary users. In other circumstances, the sensitivity of the data or the sophistication and resources of the user may dictate the data be eradicated in a manner impervious to forensic recovery.

There are three principal areas where the data resides: Allocated Clusters (active data), Unallocated Clusters and File Slack Space. Importantly, the last two are not needed by users for proper function of their machines so selective eradication within these forensic areas is wasted effort (except insofar as may be required to gather evidence of the misconduct or establish damages). Instead, a sufficient eradication protocol may require the examiner to first overwrite or "double delete" (deletion followed by emptying of the recycle bin) the contraband data and then to thoroughly overwrite the entire contents of unallocated clusters and slack space. Alternatively, the protocol may provide for the examiner to relocate only benign data to a new drive and destroy or sequester the drive holding the contraband data. Either method is reasonably effective in preventing the user from regaining access to the contraband data using the sterilized drive.

Before any changes are made to the evidence drive to effect data eradication, the court should assess whether the contents of the drive with contraband data must first be preserved by forensic imaging for use as evidence in the case.

Exemplar Examination Protocol

Computer forensics examinations are often launched to resolve questions about the origins, integrity and authenticity of electronic documents. The processes employed are specialized and quite technical. Following is a list of exemplar steps that might be taken in a forensic examination to assess the alleged authoring dates of particular Excel and Word documents and e-mail:

1. Load the authenticated image into an analysis platform and examine the file structures for anomalies.
2. Assess the integrity of the evidence by, e.g., checking Registry[4] keys to investigate the possibility of drive swapping or fraudulent reimaging and looking at logs to evaluate BIOS date manipulation.
3. Look at the various creation dates of key system folders to assess temporal consistency with the machine, OS install and events.
4. Look for instances of applications that are employed to alter file metadata and seek to rule out their presence, now or in the past.
5. Gather data about the versions and installation of the software applications used to author the documents in question and associated installed hardware for printing of same.
6. Seek to refine the volume snapshot to, e.g., identify relevant, deleted folders, applications and files.
7. Carve the unallocated clusters for documents related to Excel and Word, seeking alternate versions, drafts, temp files or fragments.
8. Look at the LNK[5] files, TEMP directories, Registry MRUs[6] and, as relevant, Windows prefetch area[7], to assess usage of the particular applications and files at issue.
9. Look at the system metadata values for the subject documents and explore evidence, if any, of alteration of the associated file table entries.
10. Run keyword searches against the contents of all clusters (including unallocated clusters and file slack) for characteristic names, contents of and misspellings in the source documents, then review same.
11. Sort the data chronologically for the relevant Modified, Accessed and Created (MAC) dates to assess the nature of activity proximate to the ostensible authoring dates and claimed belated authoring dates.
12. Run a network activity trace report against, inter alia, the index.dat[8] files to determine if there has been research conducted at pertinent times concerning, e.g., how to change dates, forge documents and the like.
13. Examine container files for relevant e-mail and confirm temporal consistency. If web mail, look at cache data. If not found, carve unallocated clusters in an effort to reconstruct same.
14. Gather the probative results of the efforts detailed above, assess whether anything else is likely to shed light on the documents and, if not, share conclusions as to what transpired.

[4] The system Registry is a complex database used by the Windows operating system to record configuration and other data pertaining to the file system and installed applications..
[5] LNK (pronounced “link’) files are shortcut files which Microsoft Windows automatically creates for the operating system’s use each time a user accesses a file or storage device.
[6] MRU stands for “Most Recently Used.” Entries called “keys” within the system Registry record the files used most recently by applications.
[7] To optimize performance, Microsoft Windows stores data revealing program usage patterns.
[8] Index.dat files store records of a user’s Internet activity, even if the user has deleted their Internet history.

Problematic Protocols

Though the preceding is actually a simplified and focused examination protocol, it details activities clearly beyond the ken of most lawyers and judges. Effective protocols demand technical expertise to design and describe. Not surprisingly, court-ordered examination protocols seen in reported cases are frequently forensic examinations in name only or simply gloss over the actions permitted to the examiner. ***See*** Appendix A: An Illustration of Problematic Examination Protocols for a discussion of two recent decisions that exemplify the multitude of problems that can result from misguided examination protocols.

To safeguard against time- and money-wasting examinations, the court and counsel must either avail themselves of expert assistance or become conversant about the technical issues presented in order to craft examination protocols that are feasible, cost-effective and calculated to achieve the desired ends. Proposed protocols should be drafted by an expert, or at least reviewed by one before maturing into an order.

Crafting Better Forensic Examination Orders

In framing a forensic examination order, it's helpful to set out the goals to be achieved and the risks to be averted. By using an aspirational statement to guide the overall effort instead of directing the details of the expert's forensic activities, the court reduces the risk of a costly, wasteful exercise. To illustrate, a court might order: "The computer forensic examiner should, as feasible, recover hidden and deleted information concerning [relevant issues and topics] from Smith's systems, but without revealing to any person(s) other than Smith's counsel (1) any of Smith's personal confidential information or (2) the contents of privileged attorney-client communications."

The court issued a clear, succinct order in **Bro-Tech Corp. v. Thermax, Inc., 2008 WL 724627 (E.D. Pa. Mar. 17, 2008)**. Though it assumed some existing familiarity with the evidence (e.g., referencing "the Purolite documents"), the examiner should have had no trouble understanding what was expected and conducting the examination within the confines of the order:

> (1) Within three (3) days of the date of this Order, Defendants' counsel shall produce to Plaintiffs' computer forensic expert forensically sound copies of the images of all electronic data storage devices in Michigan and India of which Huron Consulting Group ("Huron") made copies in May and June 2007. These forensically sound copies are to be marked "CONFIDENTIAL--DESIGNATED COUNSEL ONLY";
>
> (2) Review of these forensically sound copies shall be limited to:
> (a) MD5 hash value searches for Purolite documents identified as such in this litigation;
> (b) File name searches for the Purolite documents; and
> (c) Searches for documents containing any term identified by Stephen C. Wolfe in his November 28, 2007 expert report;
>
> (3) All documents identified in these searches by Plaintiffs' computer forensic expert will be provided to Defendants' counsel in electronic format, who will review these documents for privilege;
>
> (4) Within seven (7) days of receiving these documents from Plaintiffs' computer forensic expert, Defendants' counsel will provide all such documents which are not privileged, and

a privilege log for any withheld or redacted documents, to Plaintiffs' counsel. Plaintiffs' counsel shall not have access to any other documents on these images;

(5) Each party shall bear its own costs;

Of course, this order keeps a tight rein on the scope of examination by restricting the effort to hash value, filename and keyword searches. Such limitations are appropriate where the parties are seeking a small population of well-known documents, but would severely hamper a less-targeted effort.

Hashing

In the order just discussed, the court referenced MD5 hash value searches. Hashing is the use of mathematical algorithms to calculate a unique sequence of letters and numbers to serve as a "fingerprint" for digital data. These fingerprint sequences are called "message digests" or, more commonly, "hash values." It's an invaluable tool in both computer forensics and electronic discovery, and hashing is deployed by courts with growing frequency.

The ability to "fingerprint" data enables forensic examiners to prove that their drive images are faithful to the source. Further, it allows the examiner to search for files without the necessity of examining their content. If the hash values of two files are identical, the files are identical. This file-matching ability allows hashing to be used to de-duplicate collections of electronic files before review, saving money and minimizing the potential for inconsistent decisions about privilege and responsiveness for identical files.

These are the most important things for a jurist to know about hashing:

1. Electronically stored information of any type or size can be hashed;
2. The algorithms used to hash data are not proprietary, and thus cost nothing to use;
3. No matter the size of the file that's hashed, its hash value is *always* a fixed length;
4. The two most common hash algorithms are called MD5 and SHA-1;
5. No one can reverse engineer a file's hash value to reveal anything about the file;
6. The chance of two different files having matching MD5 hash values is one in 340 *trillion trillion trillion*.

A court may order the use of hash analysis to:

1. Demonstrate that data was properly preserved by recording matching hash values for the original and its duplicate;
2. Search data for files with hash values matching hash values of expropriated data alleged to be confidential or proprietary;
3. Exclude from processing and production files with hash values matching known irrelevant files, like the Windows operating system files or generic parts of common software; or,
4. Employ hash values instead of Bates numbers to identify ESI produced in native formats. Much ESI no longer lends itself to printable, page-like forms. Hash values offer a low-cost, reliable way to uniquely identify and authenticate these new forms.

Hashing is often a pivotal tool employed to conclusively identify known contraband images in prosecutions for child pornography.

Although hashing is a useful and versatile technology, it has a few shortcomings. Because the tiniest change in a file will alter that file's hash value, hashing is of little value in finding contraband data once it's been modified. Changing a file's name won't alter its hash value (because the name is generally not a part of the file), but even minimally changing its contents will render the file unrecognizable by its former hash value. Another limitation to hashing is that, while a changed hash value proves a file has been altered, it doesn't reveal how, when or where within a file changes occurred.

Frequently Asked Questions about Computer Forensics

How do I preserve the status quo without ordering a party to stop using its systems?

The ongoing use of a computer system erodes the effectiveness of any subsequent computer forensic examination and presents an opportunity to delete or alter evidence. Where credible allegations support the need for forensic examination, the best course is to immediately require that a forensically sound image of the machine or device be secured by a qualified technician and authenticated by hash value calculation. Alternatively, the party in control of the machine may agree to replace the hard drive and sequester the original drive such that it will not be altered or damaged.

A party wants to make "Ghost" images of the drives. Are those forensically sound?

No. only tools and software specially suited to the task collect every cluster on a drive without altering the evidence. Off-the-shelf software, or the failure to employ write protection hardware devices, will make changes to the evidence and fail to collect data in all of the areas important to a thorough forensic examination.

The use of Ghost imaging methods may be entirely sufficient to meet preservation duties when issues requiring computer forensics issues aren't at stake.

Do servers need to be preserved by forensically sound imaging, too?

Though forensic examiners may differ as to when exceptions apply, as a general rule, forensically sound imaging of servers is unwarranted because the manner in which servers operate makes them poor candidates for examination of their unallocated clusters. This is an important distinction because the consequences of shutting down a server to facilitate forensic acquisition may result in severe business interruption consequences to a party. Live acquisition of the server's active data areas is usually sufficient and typically doesn't require that the server be downed.

What devices and media should be considered for examination?

Though computer forensics is generally associated with servers, desktops and laptops, these are rarely the only candidates for examination. When they hold potentially relevant ESI, forensic acquisition and/or examination could encompass external hard drives, thumb drives, media cards, entertainment devices with storage capabilities (e.g., iPods and gaming consoles), online storage areas, optical media, external media (e.g., floppy and ZIP disks), co-located data centers, cell phones, personal digital assistants, automobile air bag modules, incident data recorders ("black boxes"), backup tapes and any of a host of other digital storage devices. Moreover, machines used at home, legacy machines sitting in closets or storage rooms and machines used by secretaries, assistants family members and other persons serving as proxies for the user must be considered as candidates for examination.

How intrusive is a computer forensics examination?

The intrusion associated with acquisition is a temporary loss of access to the computer or other device. To enable an examiner to make a forensically sound image, the user must surrender his or her computer(s) for several hours, but rarely longer than overnight. If a user poses no interim risk of wiping the drive or deleting files, acquisition can generally be scheduled so as not to unduly disrupt a user's activities.

A properly conducted acquisition makes no changes to the user's data on the machine, so it can be expected to function exactly as before upon its return. No software, spy ware, viruses or any other applications or malware are installed.

The intrusion attendant to forensic examination flows from the fact that such examination lays bare any and all current or prior usage of the machine, including for personal, confidential and privileged communications, sexual misadventure, financial and medical recordkeeping, storage of proprietary business data and other sensitive matters. Though it may be possible to avoid intruding on such data within the orderly realm of active data, once deleted, relevant and irrelevant data cannot easily be segregated or avoided. Accordingly, it's important for the court to either impose strict limits on the use and disclosure of such information by the examiner, or the examination should be conducted by a neutral examiner obliged to protect the legitimate discovery and privacy concerns of both sides.

What does it cost?

Though the forensic acquisition and preservation of a desktop or laptop machine tends to cost no more than a short deposition, the cost of a forensic examination can vary widely depending upon the nature and complexity of the media under examination and the issues. Forensic examiners usually charge by the hour, with rates ranging from approximately $200-$500 per hour according to experience, training, reputation and locale. Costs of extensive or poorly targeted examinations can quickly run into five- and six-figures. Nothing influences cost more than the scope of the examination. Focused examinations communicated via clearly expressed protocols tend to keep costs down. Keyword searches should be carefully evaluated to determine if they are over- or under inclusive. The examiner's progress should be followed closely and the protocol modified as needed. It's prudent to have the examiner report on progress and describe work yet to be done when either hourly or cost benchmarks are reached.

Further Reading

Other Articles by Craig Ball

Four on Forensics, available without cost at **http://www.craigball.com/CF4_0807.pdf**
This collection of articles includes, "***Computer Forensics for Lawyers Who Can't Set a Digital Clock*,"** an in-depth but accessible look at the nuts-and-bolts of computer forensics, written for the non-technical reader. Also included are, "*Meeting the Challenge: E-mail in Civil Discovery*," "*Finding the Right Computer Forensics Expert*" and "*Cross-examination of the Computer Forensic Expert*."

Eight on EDD, available without cost at **http://www.craigball.com/EDD8_May_2008.pdf**
This collection focuses on a wide range of electronic discovery topics including:

- **Musings on E-Discovery—Ball in Your Court: April 2005 through June 2008:** The award winning electronic discovery column from Law Technology News.

- **Hitting the High Points of the New e-Discovery Rules**: A thumbnail summary of the e-discovery Amendments to the Federal Rules and some of the ways they change the landscape of litigation.
- **What Judges Should Know About Discovery from Backup Tapes:** The e-discovery wars rage in the mountains of e-mail and flatlands of spreadsheets, but nowhere is the battle so pitched as in the trenches of back up tapes. Here's why, and how not to end up a casualty.
- **The Plaintiff's Guide to Meet and Confer**: Learning to navigate the Rule 26(f) conference and its ilk--asking the right questions and being ready with the right answers—is an essential advocacy skill. This article instructs requesting parties how to make the most of meet and confer in ways that balance the need for information against the requisite costs and burden.
- **Metadata: Beyond Data About Data**: What's metadata, and why is it so important? It's the electronic equivalent of DNA, ballistics and fingerprint evidence, with a comparable power to exonerate and incriminate. All sorts of metadata can be found in many locations. Some is crucial evidence; some is digital clutter. But because every active file stored on a computer has some associated metadata, it's never a question of whether there's metadata, but what kinds of metadata exist, where it resides and whether its potential relevance demands preservation and production.
- **The Perfect Preservation Letter**: This article looks at what is usually the requesting party's first foray into EDD: the letter demanding preservation of electronic evidence. A well-drafted preservation letter serves as the e-discovery blueprint, and the considerations that go into drafting the "perfect" preservation letter reveal much about the power and perils of EDD. An exemplar letter is included.
- **The Plaintiff's Practical Guide to E-Discovery**: This two-part article focuses on the needs of the requesting party. Part I addresses challenges unique to EDD, elements of a successful e-discovery effort and steps to compel preservation of e-evidence. Part II looks at the pros and cons of production formats, explores common e-mail systems and offers tips for getting the most out of your e-discovery efforts and budget.
- **Discovery of Electronic Mail: The Path to Production**: This article outlines issues and tasks faced in production of electronic mail—certainly the most common and perhaps the trickiest undertaking in electronic discovery. It's a guide to aid attorneys meeting and conferring with opposing counsel, working with e-discovery service providers, drafting production requests and explaining the cost and complexity of e-mail production to clients and the court.

Published by the Federal Judicial Center

Managing Discovery of Electronic Information: A Pocket Guide for Judges by Barbara J. Rothstein, Ronald J. Hedges, and Elizabeth C. Wiggins
Available at **http://www.fjc.gov/public/pdf.nsf/lookup/eldscpkt.pdf/$file/eldscpkt.pdf**

Published by the U.S. Department of Justice, National Institute of Justice

Electronic Crime Scene Investigation: A Guide for First Responders, Second Edition, available at **http://www.ncjrs.gov/pdffiles1/nij/219941.pdf**

Appendix A: Problematic Protocols: Two Recent Decisions

A well-crafted protocol is the key to a successful forensic examination—one that employs the most efficient and cost-effective tools and methods for analyzing the particular type and quantity of ESI presented. Two recent decisions exemplify the problems that flow from inadequate examination protocols.

Consider this protocol from **Ferron v. Search Cactus, L.L.C., 2008 WL 1902499 (S.D. Ohio Apr. 28, 2008):**

> 1. Within seven days of the date of this Opinion and Order, Plaintiff's forensic computer expert shall mirror image both of Plaintiff's computer systems' hard drives and Plaintiff shall preserve this mirror image.
> 2. Plaintiff's forensic computer expert shall then remove only Plaintiff's confidential personal information from the mirror image of Plaintiff's computer systems' hard drives. Plaintiff's expert shall provide Defendants with the protocol he utilized to remove the confidential information.
> 3. Plaintiff shall then provide Defendants' computer forensic expert access to his computer systems' hard drives.
> 4. Defendants' forensic computer expert shall mirror image Plaintiff's computer systems' hard drives in approximately four to eight hours for each system. If the expert finds that this is not enough time, Plaintiff is expected to be reasonable in allowing some additional time. Defendant is expected to be considerate with regard to scheduling times that are less intrusive to Plaintiff and his business.
> 5. Defendants' expert shall review his findings in confidence with Plaintiff prior to making any findings available to Defendants.
> 6. Plaintiff shall identify for deletion any information that is irrelevant and create a specific privilege log of any relevant information for which he claims privilege. The computer forensic expert shall remove the information claimed as privileged and provide all other information to Defendants.
> 7. Defendants' expert shall provide Plaintiff with the protocol he utilized to remove the privileged information.
> 8. Forensic computer experts [omitted] shall act as officers of this Court. Defendants shall be responsible for remunerating [Defendant's expert] and Plaintiff shall be responsible for remunerating [Plaintiff's expert].

It's unclear whether the plan is for plaintiff's expert to sterilize drive images before making *the images* available to the other side's examiner or whether the unsterilized source *hard drives* will be made available to the defendant's expert for imaging and examination. A literal reading supports the latter conclusion, but then what's the point of plaintiff's sterilization effort? Moreover, the order ignores the Herculean challenge faced in thoroughly cleansing a drive of particular confidential information in anticipation of forensic examination. If, for example, the plaintiff's e-mail included both confidential and discoverable messages, the examiner would be obliged to obliterate and reconstitute the plaintiff's e-mail container file. Additionally, the unallocated clusters typically carry tens or hundreds of gigabytes of commingled, undifferentiated data. Finally, the order offers no guidance as to what the examiners are allowed or expected to do, or whether the defendant's expert is limited in what he can share with defense counsel. The order appears detailed, but offers practically no pertinent guidance.

Another recent case, **Coburn v. PN II, Inc., 2008 WL 879746 (D. Nev. Mar. 28, 2008),** exemplifies the difficulties attendant to programming a forensic examination in a court order. The court modeled its order on the protocol in **Playboy Ent., Inc. v. Welles, 60 F.Supp.2d 1050 (S.D. Cal. 1999).** While there is much to commend the order in terms of addressing the choice of expert, privilege concerns, confidentiality and convenience, the order is silent as to whether or how the forensic expert will recover information or analyze the data.

> 1. The parties shall meet, confer and agree upon the designation of a computer expert who specializes in the field of electronic discovery to create a "mirror image" of the relevant hard drives. If the parties cannot agree on an expert, they shall submit suggested experts to the court by April 18, 2008. The court will then select and appoint a computer specialist. The services of the expert will be paid by defendants.
>
> 2. The court appointed computer specialist will serve as an officer of the court. To the extent the computer specialist has direct or indirect access to information protected by the attorney-client privilege, such "disclosure" will not result in a waiver of the attorney-client privilege. Defendants herein, by requesting this discovery, are barred from asserting in this litigation that any such disclosure to the court designated expert constitutes any waiver by Coburn of the attorney-client privilege. The computer specialist will sign a protective order stipulated to by the parties. Lastly, any communications between defendants and/or defendants' counsel and the computer specialist as to the payment of fees and costs pursuant to this order will be produced to Coburn's counsel.
>
> 3. The parties shall agree on a day and time to access Coburn's computer. Defendants shall defer to Coburn's personal schedule in selecting this date. Representatives of both parties shall be informed of the time and date, but only Coburn and her counsel may be present during the hard drive recovery.
>
> 4. After the computer specialist makes a copy of Coburn's hard drives, the "mirror image" (which the court presumes will be on or transferred to a disk(s)) will be given to Coburn's counsel. Coburn's counsel will print and review any recovered documents and produce to defendants those communications that are responsive to any earlier request for documents and relevant to the subject matter of this litigation. Such discovery shall include, but not be limited to, information pertaining to defendants' contention that Coburn misappropriated their trade secrets. While no counterclaim for misappropriation of trade secrets has been made, such information is relevant to, and is reasonably calculated to lead to admissible evidence concerning, Coburn's alleged damages and emotional distress, and for impeachment purposes. All documents that are withheld on a claim of privilege will be recorded in a privilege log.
>
> 5. Coburn's counsel will be the sole custodian of and shall retain this "mirror image" disk(s) and copies of all documents retrieved from the disk(s) throughout the course of this litigation. To the extent that documents cannot be retrieved from Coburn's computer hard drives or the documents retrieved are less than the whole of data contained on the hard drives, Coburn's counsel shall submit a declaration to the court together with a written report signed by the designated expert explaining the limits of retrieval achieved.

> 6. The "mirror image" copying of the hard drives, and the production of relevant documents, shall be completed by May 30, 2008.

Notably, all the expert does is duplicate the drive and hand it over to counsel for printing and review of any “recovered” documents; but the expert isn’t permitted to *recover* any documents, and presumably Coburn’s counsel is ill-equipped to conduct a forensic examination or recover any hidden or deleted data without expert assistance. The “forensic” nature of the process is illusory because the examiner isn’t permitted to do more than duplicate the hard drive. The party seeking forensic examination is in no wise aided because the process isn’t calculated to expose new evidence. Coburn already had access to the contents of her hard drive, and Coburn’s counsel was already under an obligation to search same for responsive ESI. It’s an empty, expensive exercise.