



# DIGITAL DISCOVERY & E-EVIDENCE



**VOL. 9, NO. 4**

**REPORT**

**APRIL 1, 2009**

Reproduced with permission from Digital Discovery & e-Evidence, 09 DDEE 04, 04/01/2009. Copyright © 2009 by The Bureau of National Affairs, Inc. (800-372-1033) <http://www.bna.com>

## TRENDS

EMC'S Jake Frazier examines the necessity of making forensic copies, now that forensically sound collection practices have been widely accepted.

### Are We Entering the Post-Forensic Era?

By JAKE FRAZIER

**“F**orensics” is a term that is thrown around in e-discovery on a daily basis, by those with a firm grasp on the issues and by others as well. Terms like “forensic copy” and “forensically sound” populate almost every one of the hundreds of company brochures offering “end-to-end” e-discovery services and tools. However, as diluted as this terminology has become, the necessity for thorough, verifiable, and repeatable methods to collect electronically stored information (ESI) is not in question.

So how then can I suggest that we are entering a “post-forensic” era? The issue turns more on the necessity for a “forensic copy” of a piece of media, and not the use of forensically sound methods for collecting evidence. Thus, put correctly, the question becomes: “Are we entering a post ‘forensic copy’ era?”

*Jake Frazier, Esq. is director of the Compliance and E-discovery Practice in EMC's Content Management and Archiving division. Mr. Frazier can be reached at [frazier\\_jake@emc.com](mailto:frazier_jake@emc.com)*

**Guidance from Sedona.** In all issues relating to e-discovery, it is best to first consult The Sedona Conference®, whose contribution to this difficult intersection of law and technology cannot be overstated. In *The Sedona Conference® Glossary For E-Discovery & Digital Information Management*, the definition of “forensics” is as follows:

*Forensics:* The scientific examination and analysis of data held on, or retrieved from, ESI in such a way that the information can be used as evidence in a court of law. It may include the secure collection of computer data; the examination of suspect data to determine details such as origin and content; the presentation of computer based information to courts of law; and the application of a country's laws to computer practice. Forensics may involve recreating “deleted” or missing files from hard drives, validating dates and logged in authors/editors of documents, and certifying key elements of documents and/or hardware for legal purposes.

More context is gained from this definition when it is juxtaposed next to the entry in the same glossary for “Forensic Copy.”

*Forensic Copy:* A forensic copy is an exact copy of an entire physical storage media (hard drive, CD-ROM, DVD-ROM, tape, etc.), including all active and residual data and unallocated or slack space on the media. Compresses and

encrypts to ensure authentication and protect chain of custody. Forensic copies are often called “image” or “imaged copies.” See Bit Stream Back-up and Mirror Image.

It is the difference between these two definitions that is the focus of my inquiry. Very clearly, evidence must be collected in a manner that can be demonstrated to ensure its accuracy, a chain of custody, and its authenticity.

Either purposefully or accidentally, it is very easy to alter ESI; in fact, it’s much easier to alter ESI than to alter a physical document.

For example, the deletion or addition of the word “not” could control an entire case. Imagine the importance of the word “not” in an employment case, where an e-mail reads, “I am directing you not to consider race in your hiring decisions.”

If this same document were printed, the word “not” could certainly be whited out and copied over, but then an obvious gap would be left in the sentence, making it clear that someone has altered the document; not so for ESI.

Also, thousands of pages of evidence could be deleted altogether with one keystroke, by deleting a \*.PST file that contains thousands of e-mails. The same cannot be said for boxes of paper documents, which would need to be shredded or burned, a laborious and easily detected process. Hence, the need for the heightened sensitivity around forensics and ESI becomes obvious.

**Forensic Copying to the Rescue.** Early tools such as En-Case and Forensic Toolkit helped solve this problem by creating “forensic copies” of pieces of media. With this kind of copy made early in a case, and perhaps without the custodian’s knowledge, the largest amount of potential evidence was then safely quarantined.

Forensic experts could then make a working copy of this same image, and take their time recreating the actions of the user of the machine such as their browsing history or determining whether they made copies of documents using the USB ports or other copying mechanisms. These consultants could also recover or recreate deleted files from the fragments left on the magnetic hard disk inside a machine, defeating attempts at evidence destruction and intentional spoliation.

**Storage Capacity Explosion.** It was also at this time that the typical storage amounts on individual workstations meant users could store thousands of documents and e-mails very easily on their laptops or PCs and have complete and unsupervised access to these ESI stores, often while they were away from the office.

The prospect of custodians spending the night before they were ordered to turn in their computers for evidence harvesting reading and selecting and deleting any incriminating documents they chose struck panic in those charged with evidence collection.

Law enforcement also began using these tools in prosecutions of suspects in crimes ranging from bank fraud to child pornography possession.

Accidental spoliation of ESI that becomes evidence can occur as well. The creation of forensic copies helped those who wished to do nothing more than comply with their preservation obligations, but had so many IT systems running (like virus scans, autodelete or autopurge routines, enterprise search applications, etc.) that the “last accessed date” metadata field might change.

Any case where a “knowingly” or “intentional” mental state requirement was at issue was made more difficult if it were not clear when a custodian “knew” a given fact, as established by ascertaining when he accessed a file.

Without doubt, the forensic copy has been and continues to be a key asset in the evidence collection arsenal.

**Have We Gone Too Far?** Working at an e-discovery service bureau prior to the revisions to the Federal Rules of Civil Procedure, my company was hired by a defendant-corporation in a civil case. A team of more than a dozen forensic technicians was mobilized to over 30 cities to create and collect over 1,000 forensic images of individual workstations, servers, and the like.

Forensic copies were made, so every system file and every bit of “empty” space on each hard drive was collected, and put into a proprietary compressed evidence file. The bill was well over \$1 million to do the collection alone and pay for travel expenses.

Once the processing started, technicians certified by the software company that owns the collection software were required to open and extract the documents from these copies. This was a painstaking process and took weeks. Only then could the documents be put into a review tool such that the e-discovery process could resume.

To my recollection, there was no suspicion of wrongdoing or intentional deletion of files, no was recreating the actions of the users particularly important. However, this was the “best” way to collect evidence at the time.

It is likely that this example and many others like it were the impetus for revising the Federal Rules of Civil Procedure, especially Fed. R. Civ. P. 26(b)(2)(b), which now states:

(B) *Specific Limitations on Electronically Stored Information.* A party need not provide discovery of electronically stored information from sources that the party identifies as not reasonably accessible because of undue burden or cost.

Without diving deeply into the issue of what constitutes “undue burden,” it is not unreasonable to assume that objection to the creation of forensic copies of hard drives—where there has been no showing of the special need for doing so, is at least a reasonable argument to be made during the Fed. R. Civ. P. 26 (f) meet and confer process, or if before that process, perhaps as a best practice, or to be discussed in preservation communications earlier on in a case.

Otherwise, the only conclusion is that making a forensic copy of any hard drive that contains ESI that “may be relevant to future litigation” is in order, and that is clearly cost prohibitive.

The forensic tools do offer network versions of the same products with one enhancement: the ability to create forensic images and collect them remotely. While negating the need for travel, this does, however, still result in proprietary file formats that are difficult to leverage with other e-discovery tools and processes.

**Current Storage Configurations.** That today’s network storage of documents is different than the environment described above is changing the landscape further. Many relevant documents and messages are now stored in centrally accessible systems. For example, the advent of e-mail archiving and enterprise content management

systems means much of the critical information sought is accessible via access to these platforms.

Also, many organizations are now synchronizing the “My Documents” folder located on a user’s hard drive with a corresponding folder on that user’s “U drive” or whatever the nomenclature is for that user’s dedicated portion of central fileshare storage.

In addition, with the proliferation of Sharepoint, eRoom, Wikis, blogs, and other collaborative spaces, there is another network accessible point of evidence collection.

Virtualization has also complicated the “forensic copy” method as today the pendulum is swinging back to more of a “terminal” model and less of a local storage model.

**Newer Tools.** Today, next generation automated search and retrieval tools can index and retrieve evidence from all of these sources, as well as fileshares and even the very laptops and workstations that gave rise to the forensic copy. True, these solutions can only obtain “active” files that have not been deleted, but that is likely an issue to be discussed in light of the “reasonably accessible” language of Fed. R. Civ. P. 26(b)(2)(b), or the spirit of the same.

In addition, these collection tools operate “forensically” or in a “forensically sound manner,” meaning that they do not change metadata and maintain the hash value of each file collected with robust audit trails.

These tools also account for other realities of today’s e-discovery environment, namely the creation of “data topologies,” which are helpful in Fed. R. Civ. P. 26(a)(1)(b) disclosures as well as Fed. R. Civ. P. 26(f) meet and confers, and to satisfy the “identifies” requirement in Fed. R. Civ. P. 26(b)(2)(b).

Furthermore, because they maintain a full text index of ESI in the environment, they can also be used for prospective litigation holds (where not only relevant ESI that exists “today” is in scope, but also everything created going forward) instead of having to recreate the work of re-imaging the hard drives every few weeks or few months.

Rolling productions are also made easier with these tools, helping those complying with government investigations to satisfy accelerated deadlines.

**Best Practices?** As part of the working group that drafted the Preservation node of the Electronic Discovery Reference Model, this issue was discussed and published at length. The central theme of the group’s recommendations centered on the designation of “key players” or primary witnesses, “secondary witnesses,” and “tertiary” or peripheral witnesses—expressed as the witnesses’ logical proximity to the transactions or occurrences.

For example, in a case about the hiring practices of a division, it is logical to assume that the forensic copying of the ESI of the division president, the head of HR for that division, and perhaps two or three other key players is reasonable and not terribly expensive. Individual licenses of a forensic tool should always be kept on hand for these limited purposes. However, for secondary and peripheral witnesses, next generation active data collection tools can save organizations significant amounts of money, all while maintaining the spirit of “forensics.”

---

*Care to comment on the views expressed by the author? Send an e-mail to [ceannou@bna.com](mailto:ceannou@bna.com)*