InsideCounsel.com
BUSINESS INSIGHTS FOR LAW DEPARTMENT LEADERS
May 11, 2012
TECHNOLOGY


Technology: Forensically sound collection of ESI

Options for gathering electronic information without altering metadata

BY THOMAS LIDBURY, MICHAEL BOLAND

The phrase "forensic collection" often is associated in our minds with a bit-by-bit copy of a computer's entire hard drive. This may be crucial in cases where we might expect authentication issues or where investigation of slack and fragmented space may be important. Criminal cases are a prime example. But a collection of electronically stored information (ESI) may be limited to only certain files that are likely to be relevant and still be forensically sound.

What makes a collection forensically sound, whatever its scope, is not that the entire storage media has been copied bit by bit, but that the files that have been collected can be shown to be exact copies of what was on the source, including associated metadata. This requires that the collection method not alter the files or their metadata It also usually includes some way of ensuring non-alteration after collection, which generally means taking a digital fingerprint in the form of a hash value that can be securely stored and used later to verify that the document still is exactly like it was at the time of collection

There are several commercially available tools that can collect specific files or entire hard drives in a forensically sound manner. With some of those, it is possible to narrow the collection by date ranges, search terms and other parameters. Any good collection vendor has these in his tool box. Some companies also have these tools in house.

Reasons for collecting in a forensically sound manner may be varied. It may be done because of a high expectation of challenges to authenticity, as is common in criminal cases. On the other hand, it may be done simply because a forensic tool is readily available and using it minimizes potential risk, permitting the affected employees to return to normal document management after collection has occurred. But it is not just a defensive measure. Many benefits of collecting in a forensically sound manner inure to the benefit of the party doing the collecting.

A forensically sound collection including the associated metadata allows for more robust data analytics and culling. For example, we can use email threading to see who has been communicating with whom about what by using a metadata field that identifies related emails in a chain even if the subject line has been changed along the way. This can be a valuable tool as we identify key players and begin to separate the potentially relevant documents from the rest. Of course other metadata fields also can be useful for sorting and searching.

But a completely forensically sound collection is not always necessary. Much of the benefit of a forensically sound collection can still be obtained without using specialized collection software. Some simple methods for collecting ESI do alter some metadata fields, which usually are less important and may be unnecessary in many cases, such as the creation date, last modified date, last accessed date, source path, etc.

For the best results:

- Do not attach loose documents to emails
- Do not copy documents to a new location
- Do not PDF or TIFF them (which essentially turns them into a digital version of paper)
- Instead, ZIP them up at the folder level from their original location. Once ZIPed, the files can be emailed, FTPed or placed on portable media for overnight shipment. The ZIP file can easily be password protected for security in transit. This method is not perfectly forensically sound, but it may be enough in many situations.

ZIP (or RAR) files are just container files. The programs can easily be downloaded from the web using applications like Winzip or Winrar. Think of these as electronic bankers boxes that hold loose files.

However, unlike a banker's box 2,500 page limit, a .ZIP file can be very large and hold many gigabytes of data. A ZIP folder is a wrapper around the documents that protects most of the metadata, while at the same time compressing the data into a smaller size, making it easier to copy and transmit.

For email:

- Create a PST (Outlook) or NSF (Lotus Notes) file with relevant email insidePST and NSF files are just container files for emails
- Do not forward the relevant email to your outside counsel or e-discovery vendor
- Do not attach the relevant email to another email for transmission
- And, again, do not PDF or TIFF them: That will destroy the ability to search and sort by custodian, sender, recipient, date, etc., while also breaking apart family relationships with attachments

To collect email, create a specific folder within a user account and then create a PST or NSF file of those emails. Then these PST or NSF files can be FTPed or copied onto portable media for over night shipment with password protection if necessary.

If a truly forensically sound collection is called for, there is often a challenge with remote employees who rarely connect to the network, or do so on small or unreliable pipelines. Several software companies have created a portable, plug-and-play version of their collection tools. These are small devices that look like a thumb drive or a small external hard drive. They ship out to the remote employee and plug into the USB port. They can be pre-programmed to execute the collection protocol desired for the case in question, e.g., full bit-by-bit image or a targeted collection of certain files. When the collection is complete, the remote employee ships back the device to the lawyers or the e-discovery vendor.

Another option is remote collection. Again, multiple vendors have remote collection capabilities. They may be able to conduct small and mid-sized collection remotely and during off hours. This reduces the intrusion while also cutting on travel costs and other incidental charges.

These services are available through many ESI collection vendors and, often, outside law firms.