LSP DATA SOLUTIONS

What Not to Do While Self-Collecting and Preserving ESI Data

By Shawn Huston

We see many clients forego forensic collection by data collection specialists in order to save money. While we understand the necessity to limit the overall expense during discovery, doing a poor job preserving and collecting discovery data can cost more money in the long run if you need to recollect data or spend billable hours dealing with issues raised by opposing counsel about the veracity of the process.

Knowing clients and their IT personnel will still seek to control the preservation process we are providing the below three tips to help avoid common collection problems that continue to reappear in the self-collected data sets we see.

Don't Trample the Metadata

By far, the most common problem we deal with concerning self-collected data is the corruption of metadata. When improperly copying data from one where it resides on the original computer or server you risk altering the underlying information about the files, particularly the original created and modified dates.

It's common to see clients forwarding emails directly to their counsel. In this example, the original sent and received metadata can become lost forcing you to rely solely on the information you see in the body of the email itself. On top of that, you may now need to redact the email header portion where your client forwarded the email to her counsel, adding more billable hour expense.

The way around these issues is to ensure your clients and their technical staff are using copy and backup solutions that capture the entirety of the metadata without altering it in any way. File copy tools such as Robocopy and Safecopy can help with this task directly to your delivery media. Do NOT use the copy and paste functionality provided within Windows. If you do, you can guarantee at least a minimum amount of date corruption.

If you need to transfer the compiled data after it has been compiled into a single location, such as to transfer via ftp or other file transfer method, be sure to enclose the files in a container file (.zip, .rar, forensic container, etc.) so all that work you did to properly capture the files and associated metadata wasn't for naught.

Along these lines, make sure you capture the appropriate custodian information along with each set of data collected. Having a group of emails is great, but you will want to know whose inbox they came from. If you haven't accounted for the capture of this information during collection, then it will be nearly impossible to trace back without significant effort.

The key here is to ask yourself whether the tools and methods you have use qualify as forensically sound.

Don't Over Preserve

Some years ago, we received a call from the General Counsel of a financial company that was facing bet-the-company litigation. They had already taken steps to preserve all data potentially related to the pending litigation, which was the correct thing to do. Unfortunately, the volume of data they preserved began as a backup of 30TB of data, and they were adding an additional 1TB to it daily.

By the time they picked up the phone to ask for help in limiting the data they had already preserved and craft a preservation strategy that was more in line with the specifics of the matter, they had already preserved 100TB of total data.

It is imperative to preserve all the necessary data, but ask questions early and often as to what data exists, where it is stored and who has access to it in order to limit the total volume of data to only what is necessary. The cost to sort through it after the fact could be massive if you don't plan appropriately in the beginning.

Don't Miss Important Data

An even more common situation than over-collection when relying on client resources is under collection. Missing relevant documents and data during the preservation and collection process can lead to the need to recollect data, or potentially inadvertent deletion of important data.

When we are engaged to assist in gathering data, one of the first orders of business is to aide in identifying any additional sources of data that may not have been already identified. This often leads us to external media, mobile devices, network shares, cloud file storage or new email sources which were never previously considered. In more than one-third of all data collections we perform, new data sources to collect are identified through this process.

While these sources may not be end up related to the issue at hand, it is better to have performed due diligence in the event the thoroughness of the collection is ever questioned. Document this information with a data map to identify and record what data existed, which custodians had access to it, and why it was not relevant.

Just because your situation may not be amenable to engaging with a forensic expert to perform your client's preservation and collection, it's still important to do it right the first time. Every situation is different when it comes to collecting data, so while we hope these tips aid in that effort, don't hesitate to reach out to us or your preferred provider for advice and guidance.