David P. Hathaway, Esq.

Making Sense of E-Discovery: 10 Plain Steps for Producing ESI – Part 2

Have you ever been handed a detailed flowchart on electronically stored information (ESI) and wondered what to make of it? Have you read case law summaries about ESI and thought, "How do I actually apply this holding to my practice?" And when you do produce ESI, is there any concern about whether the methods you used would pass muster if critiqued by a federal judge?

In March's issue of *The Briefs*, we discussed the necessity of having a plan to make sure you do everything right. The article below picks up where we left off, with the remaining five of the ten tips that will help increase your knowledge as you prepare to produce ESI

6. Collect the Data. Collecting data is not the same as filtering or searching; here you are simply corralling all of the sources of ESI so that they can be searched later in their entirety. The key is to figure out where everything is located. For example, there may be servers in California holding data that one of the custodians inputs from time to time for a particular project involving the West Coast. If you are collecting old data, don't confuse the term "backup" with "archive." A typical backup application takes periodic images of active data, usually only for a few days or weeks, to provide a method of recovering records that have been deleted or destroyed, such as to facilitate a disaster recovery. An archive is designed to provide ongoing access to decades of business information. The collection of data should be fairly inexpensive, but communication is key. If a self-collection error is made, for instance by collecting in such a way that metadata is lost, the entire process will have to be restarted, which could be extremely expensive. One of the critical points in collecting data is to understand that this is your obligation as counsel; you cannot leave it entirely up to the client. Whether you use a vendor at this stage, later in the process, or not at all depends largely on the size of the document request. If you ask detailed questions about pricing plans and "scalability" of the process, you might find an affordable option with a vendor. However, no matter who is going to collect the data, the attorney must actively assist in the process. Your goal is to collect the entire universe of data from the sources that you have identified within your client's network. Of course this raises concerns about personal information that may be captured as well. Be careful about taking physical items (e.g., thumb drives) and the chain of custody issues that follow; naturally you want to lessen the chance of having to be called to the stand as a witness. When the process is complete, you should be able to check off every source from the outline you created and ensure that each was captured in its entirety.

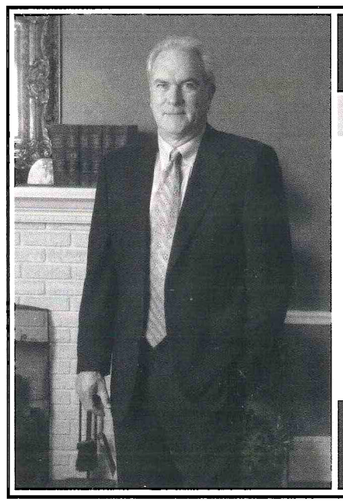
7. Filter into Groups. This stage is sometimes called "culling" or "processing," where you transfer all the collected data into a software program, such as Nuix or Symantec's Clearwell eDiscovery Platform, so that you can filter it by keyword, data range, or some other way. You are not yet reviewing the documents themselves, but only running data sets to see how many hits come back. Courts have rejected attorneys' requests for opposing parties to search the entire universe of collected data prior to filtering. When you think you have the right filters in place, you will need to get written confirmation with opposing counsel that the data sets you intend to search are acceptable, because you do not want to have to search the same material twice. Programs should have "deduplication" capabilities to eliminate duplicates and near-duplicates from the data set. When you filter into groups and then see the number of hits, you will need to get a sense of how many documents might be generated by those hits. One megabyte of email generally translates to around 60 pages of paper. Forty megabytes would be 2,400 pages, which is roughly one banker's box. One gigabyte is 1,024 megabytes, which translates to 61,440 pages, or almost 25 banker's boxes. There are a number of variables to these estimates, however, such as how many of the documents are emails versus word-processed documents, spreadsheets, presentations, PDF's, images, text files, or other types of files. For example, one megabyte of email may be 60 pages, but one megabyte of spreadsheets may be 150 pages. Another issue to consider is whether any of the files have been compressed, such as in ZIP, RAR, or other formats (some types of files compress more efficiently than others). Whereas one megabyte of text emal could be 60 pages, one megabyte of compressed email could be hundreds more. You may want to consult with an expert at this point because the page estimates will be important for budgeting. The review costs for relevancy and privileges are usually driven by the quantity of page equivalents.

8. Search for Relevancy. Now that you have a finite number of documents to search, you have two choices: either perform a "linear review," which means going through them manually at your billable rate, or use technology assisted review (TAR) which will apply predictive coding to determine which documents are most likely to be relevant. You may notice that online retailers already use predictive coding to offer products you are likely to purchase, based on past sales. Statistics indicate that TAR is actually more accurate than human review, and, in fact, some studies show that two attorneys who search the same data sets for relevancy end up with vastly different document productions. However, if a computer is going to search for relevancy, you will probably need a subject matter expert (SME) to provide the search strings, because a computer won't know the facts of the case or the applicable law. There is also the question of whether using one SME is enough and whether the SME created the right search strings. Linear review theoretically avoids those issues, but clients who choose linear review may find more than 70% of their e-discovery legal fees are spent on this step (more than all others combined). As a practical matter, if there are hundreds of boxes of documents to review, any human being is going to have trouble focusing for days or weeks on end, and there is an argument to be made that someone with a law degree should not be doing that kind of work. The same is true for reviewing documents that have been produced to you by opposing parties. Again, an e-discovery vendor can explain more about the pros and cons as to your particular case It may be worthwhile to speak to a few different vendors on the phone, as some have very different approaches, and the material they post online can look like a physics report from NASA. Be sure to provide your client with enough information to make a reasoned decision, and confirm it in writing.

9. Audit Your Results. This is an important and often overlooked step, sometimes called "sampling" or "quality control sampling." It allows you to test a small, random set of documents to make informed decisions about the entire production. After searching based on relevancy, whether by linear review or through TAR, you will be left with a responsive set and a non-responsive set. You need to randomly sample the non-responsive set and ensure they don't include anything relevant. Document your findings carefully in your journal. If you find relevant information in your non-responsive set, you obviously need to revisit the previous step, improve your approach, and run another audit. Since the relevancy search can be such an expensive endeavor, you should think very carefully from the start about how to catch every potentially relevant document and thereby avoid failing the audit. The ultimate goal is to be sure your ESI production is defensible in the court of law, and a number of legal opinions around the country indicate that sampling and other quality assurance techniques must be employed to provide reassurance to the court The party selecting the methodology must be prepared to explain the rationale for the search method selected, demonstrate that it was appropriate for the task at hand, and show that it was properly implemented. Without an audit, you will never really know whether all your hard work paid off.

10. Review for Privileges. The final step usually requires human review. When you look over your responsive set of documents and get ready to produce them, you need to review them for privileges. This is not as simple as it seems; you will need to know how the case developed through time, when the legal problems started to occur, who was dealing with those problems from which locations, and how the issues morphed into litigation. At a minimum you will need a complete list of in-house and outside counsel who were communicating with your client during the entire time period. You cannot simply look at the "To," "From" and "Cc:" lines on an email, nor can you just search for attorney or law firm names, because you might find an email from your client's CFO to the CEO stating, "I spoke with our Orlando counsel yesterday and they told me. " You should also consider the actual elements necessary to claim the attorney-client privilege, in particular that the communication was for the purpose of securing legal advice, and consider whether it was waived by including outsiders on the communication or otherwise by disclosing it to others. If you pull every document between lawyer and client, such as "I'll meet you in the lobby of our building at 2 p.m.," your privilege log is going to be as thick as a dictionary. On the other hand, many lawyers will tell you that privileged documents are almost always inadvertently produced. As discussed above, there are clawback provisions in Federal Rule of Evidence 502 that can be incorporated into an order or court-approved agreement prior to the production. Rule 502 also limits a potential waiver to the particular issue that was the subject of the communication. When reviewing for privileges, also consider the work product doctrine, which is owned by the attorney, as well as the joint defense privilege/common interest rule and the accountant-client privilege under Florida law. You also need to be on the lookout for any potential trade secrets or other confidential business information that may be contained in the documents, but the older the documents are, the less likely they will divulge trade secrets. After you make your privilege log, Bates stamp the rest of the responsive set, and produce them as a proper, defensible production of ESI.

David P. Hathaway, Esq., Dean, Mead, Egerton, et al., has been a member of the OCBA since 2002.



Lilburn R. Railey III

Mediation Services

30 Years Experience
As Civil Trial Lawyer

Circuit-Civil Mediator

FL Supreme Court
Qualified Arbitrator

National Academy of
Distinguished Neutrals

+ Contract Disputes

+ Construction

+ Commercial | Business

+ Architects | Engineers
Professional Liability

Railey Harding + Allen | PA
15 North Eola Drive | Orlando FL 32801 | 407.648.9119
schedule via: kschaffner@raileyharding.com or raileyharding.com/mediation