

InsideCounsel

Technology: Search Terms Are the Bane of My Existence

Reprinted with permission from the 01/20/2012 Edition of *InsideCounsel*. © 2012 ALM Media Properties, LLC. Further duplication without permission is prohibited.

By David J. Walton



January 20, 2012

When confronted with an electronically stored information (ESI) issue, the first thing lawyers do—almost reflexively—is exchange a list of search terms and apply them broadly against all types of media. Many judges use the same approach even though, conservatively, search terms result in 90 percent false positives. These false hits are very expensive to cull and review. Thus, a key to controlling ESI discovery costs is to use search terms carefully.

A great illustration is *I-Med Pharma Inc. v. Biomatrix, Inc.*, 2011 U.S. Dist. LEXIS 141614 (D. N.J. Dec. 9, 2011). The key issue involved the use of unallocated space and search terms. Unallocated space is the area on the hard drive where “deleted” files are stored. Although there are files in this space, the computer sees the space as open and available.

At the beginning of the case, the plaintiff agreed to an ESI protocol where the defendants were allowed to hire an expert to conduct a “forensic investigation” and keyword search of the plaintiff’s entire computer system, using more than 50 search terms. This search was not limited to specific custodians or time periods. In essence, it was a disaster waiting to happen. The unallocated space alone yielded an estimated 65 million hits, or 95 million pages of files.

Faced with reviewing million of documents for privilege and relevancy, the plaintiff sought relief from the court, even though it had agreed to the protocol and search term list. Fortunately for the plaintiff, the court modified the protocol so they did not have to produce data from the unallocated space, noting that even a cursory review of the information would mean copious

amounts of time and millions of dollars. The court also recognized that the defendants did not show the likelihood that relevant and non-duplicative information was stored in the unallocated space.

Finally, the court addressed the proposed search terms and discussed five factors to consider when analyzing whether those terms were reasonable:

The scope of documents searched and whether the search is restricted to specific computers, file systems, or document custodians

Any date restrictions imposed on the search

If the search terms contain proper names, uncommon abbreviations, or other terms unlikely to occur in irrelevant documents

If operators such as “and,” “not” or “near” are used to restrict the universe of possible results

If the number of results obtained could be practically reviewed given the economics of the case and the amount of money at issue

The moral of this story is clear: Never agree to collect and produce documents based on search terms unless you test the terms first to see how many potential hits there are. Many times I am asked to get involved in cases where search terms have already been agreed upon but no one has run the terms first to see how many hits the terms actually trigger. This is huge mistake.

Although predictive coding and concept searching are developing quickly and may one day replace search terms, for now, search terms are a necessary evil. Yet that doesn't mean that you have to blindly agree to them. Test the terms first by running a hit report. Share the report with the other side. Work with your adversary to limit the scope for each term. Be creative, but never agree to terms with too many hits. This is one of the keys to controlling ESI costs.

About the Author

Dave Walton

Dave Walton is a member in Cozen O'Connor's labor & employment practice group and co-chair of the firm's e-discovery task force. He has extensive experience in litigating matters involving restrictive covenants, trade secrets, fiduciary duties and defending employers targeted by discrimination lawsuits. He can be reached at dwalton@cozen.com.