

## In E-Discovery, Look Before You Search

---

September 2009

Just when you think you have heard it all, a case comes along to remind you how vulnerable you and your client could be when it comes to electronically stored information. Imagine e-discovery costs of \$6 million. Now imagine that same cost but, instead of being a party to the litigation, you are a non-party responding to a subpoena. Feeling sympathy for this non-party? Feeling like this could never happen to you? Well, don't be so certain. The central theme – the validity and appropriateness of document search and review methods – could create a potential risk for any litigator during the discovery process.

The amendment to the Federal Rules of Civil Procedure regarding electronic discovery was a watershed event for complex commercial litigation. The amendment was intended to lessen the cost and risk associated with discovery of electronically stored information. However, in many cases it had the opposite effect. Since becoming effective on December 1, 2006, the e-discovery landscape has been evolving at a rapid pace as litigants adapt to the amendments. On January 1, 2009, similar amendments in Michigan went into effect to guide attorneys through the complexities of e-discovery.

One of the more recent horror stories involves a third party government agency responding to a subpoena requesting over 30 categories of documents. The agency spent over \$6 million, more than 9% of its annual budget, responding to a subpoena. The agency agreed to search its entire network and its backup tapes by utilizing keyword searches with the keywords to be supplied by the requesting party. In *re Fannie Mae Securities Litigation*, 552 F.3d 814 (D.C. Cir. 2009). The agency also stipulated to the process before seeing the requesting party's keyword search list consisting of over 400 terms. Even though the search resulted in approximately 660,000 documents, the court ordered the agency to comply with the parties' stipulated order. The agency was obligated to complete a document review that required hiring 50 contract attorneys.

So how can costs be controlled and the risk of e-discovery lessened? Since 2006, federal courts and litigators have been grappling with those questions but the ground keeps shifting. Tried and true technologies and processes, such as keyword searches, which have been applied to routine e-discovery tasks are being challenged. At the same time, new (or, probably more appropriately, less utilized) technologies and processes promise to reduce costs and make e-discovery more efficient, but the unknowns – how opposing counsel will challenge the use of these technologies in court – coupled with unfamiliarity with the technologies make many attorneys and clients hesitant to use them.

### **Keyword searches - the staple of electronic discovery**

Keyword searches - familiar to any attorney who has ever spent time conducting legal research electronically or searching Yahoo or Google - is a tool used by litigants and courts to locate responsive and privileged documents. Virtually all review software programs allow searching by keywords, albeit each with their own eccentricities related to how to actually undertake the search. Yet, the same familiarity which makes attorneys comfortable with using keyword searches also may unwittingly allow them to forget that keyword searches are imperfect. As Magistrate Judge Grimm has written, "while it is universally acknowledged that keyword searches are useful tools for search and retrieval of ESI, all keyword searches are not created equal." *Victor Stanley, Inc. v. Creative Pipe, Inc.*, 250 F.R.D. 251 (D. Md. 2008).

## Continued

---

More federal courts are adopting tests to assess the adequacy of a party's keyword searches, usually applied in determining if a waiver occurred with the inadvertent production of privileged documents. Forming the basis of most tests is an evaluation of the qualifications of the person(s) who selected the keywords to use and an assessment of the sampling methods used to determine if the searches are reliable and are locating the most responsive documents. *Victor Stanley and William A. Gross Construction Associates, Inc. v. American Manufacturers Mutual Insurance Co.*, 256 F.R.D. 134 (S.D.N.Y. 2009) are two cases applying the test. As an alternative, some parties negotiate the search terms and any searching logic that should be used, making it more difficult for either side to challenge that methodology later in court.

### What Else Is Out There?

Concept searching? Auto-categorization? Is this the "new" lingo of e-discovery? These are different types of search methodologies that have been available for several years, although not as widely used in litigation as keyword searches to identify responsive and privileged documents.

Without going in to too much detail, here is a quick – and admittedly incomplete – layman's explanation of each. Conceptual searching relies on mathematical algorithms to find documents with similar language use or concepts. It is not dependent, like keyword searches, on matching identical words or phrases. Rather, after a reviewer identifies words, phrases, or paragraphs that are relevant in their case, the program performs a search and evaluates the relationships of the meaning of the words within each document to determine if it is conceptually similar to the relevant words. For instance, if you were searching for the word "rabbit," these searching tools would also bring up documents with "bunny."

Auto-categorization takes the conceptual search a step further. The program "learns" from a reviewer's actions. In its most simple form, a reviewer tags documents as either responsive and should be produced or not responsive and should not be produced. Once a review set is completed, the program scans the remaining unreviewed documents and locates more documents like those that are identified as responsive and pulls them forward for review. The non-responsive documents are pushed to the bottom of the review pile to be confirmed non-responsive as part of the review closeout. Visualize cream rising to the top – the most responsive documents are brought forward early in the review process allowing attorneys to grasp the relevant issues in the case sooner.

Vendors offering programs utilizing these search methods promise faster, cheaper, more efficient and more reliable document reviews than typical keyword searches. However, these search methodologies are not as tried and true in the litigation context as keyword searches, and their potential traps not as well known. Nonetheless, steps are being taken to encourage litigants to employ them. The notes to the recently amended Federal Rule of Evidence 502 suggest that federal courts might be warming to the validity of conceptual and auto-categorization search tools. In the context of determining if a party has waived privilege by inadvertently producing a privileged document, the notes state that, "Depending on the circumstances, a party that uses advanced analytical software applications and linguistic tools in screening for privilege and work product may be found to have taken 'reasonable steps' to prevent inadvertent disclosure." Explanatory Note on Evidence Rule 502 Prepared by the Judicial Conference Advisory Committee on Evidence Rules (Revised 11/28/2007). However, the qualifying phrase "depending on the circumstances" may still cause attorneys to shy away from non-keyword search methodologies.

Continued

---

**Lessons to Learn**

What went wrong in Fannie Mae? In trying to be cooperative and resolve disputes over the discovery requests, the agency relinquished control of its search methodologies to the requesting party. It agreed to the requesting party's search term list without having reviewed the terms beforehand. In doing so, it was unable to test whether those terms would actually identify responsive documents rather than simply documents unrelated to the litigation but still containing those terms. It could not negotiate over the search terms, or question the qualifications of those who crafted the search list. The agency no longer had any objections as to the reliability of the search methodology which sampling would have provided. The agency lost control over its collection and review costs.

Contact Kimberly Scott at +1.734.668.7696 or [scott@millerandcanfield.com](mailto:scott@millerandcanfield.com).