

CORPORATE COUNSEL

From the Experts: Document Review with No Pain, All Gain

Text analytics increases review speed, reduces costs, and improves quality

Jay Leib

It is said that document review costs are "the forgotten 80 percent." This implies that a strategy to accelerate and organize the actual document review should be contemplated up front, prior to review, and not forgotten until the end.

For projects of any size, document review costs share a direct correlation with volume of information, as more data requires more hours to locate the important, relevant documents. Additionally, human document review is not a process that benefits from economies of scale when data is not strategically organized. Using technology specifically designed to organize documents, combined with a sound review workflow, reduces review costs and compresses the time needed for review.

Text analytics search technology enables smarter review strategies that improve both the speed and quality of the expensive review process. This search technology, which evaluates unstructured text, enables users



Jay Leib

to automatically group documents into conceptually related subsets. Using text analytics to structure a review allows litigation professionals to significantly reduce discovery costs by increasing review speeds.

Basic Linear Review vs. Text Analytics

Case teams navigate data sets in a linear review by reading each document in order, often by date or control number. Reviewers move through the collection of documents one at a time, identifying responsive documents interspersed with non-responsive ones.

This type of linear organization presents challenges that can impact both the speed and quality of the review. For example, a reviewer may be responsible for reading thousands of e-mails and documents from one custodian, knowing that the majority of the documents are non-responsive. Since the material is organized by date, though, the subject matter of each document in sequence will vary greatly. Due to this ambiguity, reviewers have to slow down for each new document in order to understand its unique context.

Text analytics allows reviewers to organize documents by concept, separating sets of potentially responsive documents from sets of non-responsive documents. Reviewers are able to build momentum by applying knowledge gained from each document to the entire set, like a mountain biker muscling through the uphill climb and then reaping the benefits of a downhill journey. A re-

viewer also gains insight into a custodian's involvement well before completing the review of an entire document set. One conceptually similar document group could be all that is needed to gain a quick understanding of the facts for more informed decision-making.

In the context of document review, text analytics offers two primary features for speeding up review: searching by concept and classification.

Using Concept Search and Classification to Go Faster

Let's say we have a senior attorney who needs to build the facts of a case, moving quickly and trying to follow an investigative, non-linear pattern of thought. From the perspective of that attorney, the most important aspect of review is finding the most relevant documents for a particular issue and linking them together.

Using text analytics to search by concept rather than keyword is a key component of finding connections between documents. To execute this search, a user selects anything from a few sentences to an entire document that best represents an important issue in the case. The search engine analyzes the text and returns similar documents.

For example, if a reviewer is searching for information about the recall of a product, that reviewer can select a paragraph

from a document that discusses the recall and can retrieve related documents that may not specifically contain indicators such as the product's name.

Additionally, automatic classification—or clustering—groups similar documents together without requiring any user input. By using clustering early in the review, reviewers can gain insight into their data to better prioritize documents and focus on the ones most likely to be relevant. Reviewers can continue to code documents the same way they always have, but can also focus on groups of related documents to improve both speed and accuracy.

Beyond clustering, attorneys can train the system on relevant issues by providing key documents as examples. This type of user-trained classification—called categorization—allows case teams to magnify the efforts of subject matter experts and senior reviewers by using their knowledge to quickly identify the documents of consequence in a case. Combining categorization with sound statistical sampling-based quality control methodology can bring about exponential cost savings, as well as increase the quality of the review.

Cutting the Cost of Review

It is becoming increasingly important for legal professionals to present cost-effective solutions,

with proper quality control measures intact, to their clients. While the hourly labor costs of document review remain high, the new technologies now available offer the opportunity to achieve dramatic improvements over standard linear methods of organizing the review process.

Jay Leib, chief strategy officer of kCura Corporation, has more than a decade of experience in litigation support technology. His specialties include e-discovery workflow, document management, software development, emerging technologies, and database design. Before joining kCura, he was a senior manager with Ernst & Young's Fraud Investigation & Dispute Services practice. Prior to joining Ernst & Young, Jay founded Advocate Solutions, Inc., where he built Discovery Cracker(tm), the world's first shrink-wrapped e-discovery application. His software applications have been recognized by numerous surveys and awards in the litigation support and e-discovery industry. Jay has also worked in various roles on some of the largest e-discovery projects in the world.