

New Technology Can Prioritize Documents

Benefits include speed, monetary savings and flexibility for law departments

By **MARY ANN BENSON** and
CHRIS JANAK

The last few years have seen the acceptance of early data assessment tools designed to reduce the volume of data that is collected, processed, reviewed and ultimately produced. Subsequent successes in this area have led to wide, mainstream usage and their adoption continues to reduce costs in the most expensive segment of the process, document review.

These technologies prioritize documents based upon trained relevance ratings and allow lawyers to estimate the value of those documents and organize them in a way that optimizes the review process, while simultaneously giving the case team early access to key information. These applications represent one of the most transformative trends in e-discovery technology. Whether you are responding to a second request or reviewing documents for litigation, you should make it a point to familiarize yourself with what is happening in this area.

The past 12 to 18 months have been a time of gaining confidence in the use of these new technologies, but 2011 will be the year in which their utilization will become more prolific in the legal community. These software applications not only provide a way of substantively assessing the collected material by identifying which documents are most likely to be responsive and relevant to a case (and correspondingly, identifying more likely non-responsive material), but they can also provide the means to structure the review workflow in order to maximize benefits.

There are different terms used to describe

Mary Ann Benson is director of consulting for Epiq Systems. She also spent nine years with Price Waterhouse's Law Firm Services Group. Attorney Chris Janak is director of client Services for Epiq Systems eDiscovery group.

the various applications, including prioritization, predictive coding and automated or suggested coding, and while many people use these terms interchangeably, there are significant differences in approach and outcome. The key distinction is how the documents are evaluated and scored/coded.

Predictive coding, automated or suggested coding processes typically result in a coding "decision" – i.e., this document is responsive or not, related to this issue or not. Prioritization, on the other hand, does not make a determinative call. Rather, it provides a numeric score for each document and the review workflow is created based on these scores (the higher the score, the more likely it is that the document is responsive). The intent of this article, though, is not to discuss the different technologies in detail but to highlight the benefits of prioritization technology.

Benefits Of Prioritization

First, the earlier relevant information is available to the decision makers, the more quickly a well-informed strategic case plan can be developed and executed. True assessment of the merits of a case, by both in-house and outside counsel, cannot be done without having access to the key material.

Second, the speed with which that information is reliably identified is critical. This is particularly important when responding to second requests and conducting investigations where timelines to respond to court-ordered deadlines are short and must be met.

Third, significant cost savings can be achieved. Utilizing early data assessment tools provides for cost reductions while maintaining full defensibility. By inserting the prioritization step into the review workflow, clients are able to realize cost savings by matching the different priced resource pools (contract attorneys, paralegals and firm associates) to the appropriate set of documents. Higher value documents can be directed to high value reviewers quicker, while lower value documents

may be directed to less expensive reviewers such as contract attorneys or even paralegals.

Fourth, corporate law departments are increasingly directing outside counsel not to conduct first pass review utilizing their own associates. As a result, some law departments are becoming involved in the management of the first pass review effort themselves. This involvement may manifest itself in one of two primary ways. A

legal department may either establish or expand an in-house capability to manage and conduct first pass review or, alternatively, the company may establish a relationship with a third party that can provide managed document review supervised by outside counsel.

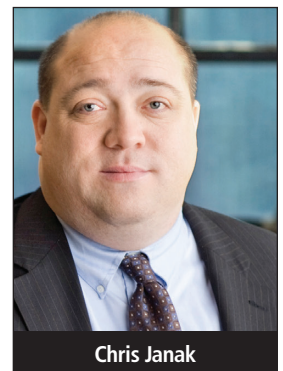
Under either scenario, the intent is that the basic task of determining initial responsiveness is handled by lower-cost resources. Technology that provides a means of determining which material should be sent to these lower cost resources while not delaying the case team's access to information is essential to the success of these cost reduction efforts.

Success Stories

Our experience over the last year working with dozens of law firms and corporate clients



Mary Ann Benson



Chris Janak

to implement this technology has been exciting and informative. While all of the projects in which the clients utilized prioritization technology had positive results, many had results with rather dramatic outcomes. Examples of these efforts include:

- In response to a second request, a law firm was able to prioritize 1.4 million documents in just a few days and was subsequently able to review the material and meet its tight production deadline to the federal government.
- Another law firm, also involved in a second request, was able to prioritize nearly 700,000 documents and similarly review and meet its production deadline to a federal agency.
- For a manufacturing company, the legal team was preparing for a multiple-month review. After prioritizing the document set and finding the relevant documents within the first week of the case, the legal team concluded that they needed to shift to a settlement position. The team estimated that without prioritization, it would have taken months to sort through the documents before arriving at the same decision.
- Similarly, in a large pharmaceutical litigation, outside counsel prioritized almost one million documents and was able to identify the relatively small responsive set of documents within days, saving its corporate client hundreds of thousands of review dollars.

In all cases, the correlation between the prioritization scores assigned to the documents (which serve as the basis for prioritizing the documents) and the review team determinations of responsiveness provided validation for the use of the scores.

Lawyers in more modest-sized matters also benefit from using these technologies. Whether there are 10,000 or 10 million documents, prioritizing documents on the basis of their responsiveness to the issues benefits all concerned both from a strategic and cost perspective.

Question Of Defensibility

As the use of these technologies becomes more routine, one issue that must be ad-

ressed is the belief that every record collected from the custodians' computers and shared servers needs to be individually reviewed. While one must approach document review in a reasonable and defensible manner, the obligation is to review the responsive material – there is no corresponding obligation to review non-responsive material.

How one accomplishes this objective varies. For example, in a paper-based discovery effort, if a box of travel expense reports and accompanying receipts for a relevant custodian is located in the file room or warehouse, but the matter at hand had nothing to do with travel or expenses, it is a fair certainty that other than possibly lifting the lid on the box and confirming content with a quick flip through the paper, not many people would spend any further effort on that box and feel that their responsibilities had been shirked. The new technologies are able to facilitate a similar sort of assessment of electronic documents, and prioritize the least likely material to the end of the review queue, while simultaneously ordering the most likely responsive material to the front.

As we all know, removing documents from the data collected for a matter is something that is done routinely – but it is done at the beginning of the data funnel. Typically when dealing with electronic data, both in-house and outside counsel apply keyword searches to filter data during or after collection and prior to processing or review begins. Traditionally, there is little testing of the results of those searches to validate the results.

While some will review the “hits,” few people, if any, look at the documents that were not “hits.” Ironically, despite this approach with the top of the data funnel, there is reticence to apply even more defensible methodologies, such as prioritization, to data after it has been processed and is available to the review team. The way in which the documents in a collection are evaluated and scored for prioritization will prove instrumental in establishing reliability of the process.

Successful prioritization requires that one single expert trains the software to identify relevant and non-relevant material. The consistency of decision making, coupled with the subject matter expertise of one individual, will ensure that the software builds its learning on

the best available information. Once the software is trained – a process that will take just a few days, and which can be done in parallel to traditional review if time constraints exist – all of the documents are assessed and scored, typically within hours.

Moreover, all of the documents are scored in a uniform manner. The computer never tires, its attention never wanders and it is able to apply what it has learned consistently over any size document population. Another feature to look for in this type of technology is whether it provides statistical assessments of the outcomes. If these are available, the process becomes very compelling.

Ultimately there is the question of defensibility and whether these new technologies will be accepted by the courts. While none of these technologies have been tested in court – and keep in mind that relatively few cases have even tested the validity of using key words to reduce the document population – the courts are clearly interested in utilizing technology to facilitate the discovery process and maintain proportionality. Indeed, some judges have voiced the opinion that attorneys should thoughtfully consider use of such tools and not be constrained by the fact that these tools are new to the world of e-discovery.

Conclusion

It is in no one's best interest to spend unnecessary time and money to review documents not material to the case. These new e-discovery technologies truly represent the intersection of a single expert attorney's judgment with computerized assessment. The benefit is not only an accelerated review process for a prioritized document collection, but also the availability of a level of accuracy and consistency that is typically not found in a traditional human review.

At a minimum, the law firms which adopt technology early will strategically differentiate themselves as they will be able to provide more accurate and insightful direction to clients earlier in the case, thereby reducing overall litigation costs. As this is a primary goal of most corporations in today's financial climate, those law firms embracing defensible, emerging technology will likely enjoy a strategic advantage in retaining and expanding their client base. ■