

The Metropolitan Corporate Counsel®

www.metrocorpcounsel.com

Volume 19, No. 1

© 2011 The Metropolitan Corporate Counsel, Inc.

January 2011

Intelligent Discovery Provides Cutting-Edge Technology

The Editor interviews Scott Kane, Partner, Squire, Sanders & Dempsey L.L.P. Mr. Kane is a litigator and the Co-chair of the firm's E-Discovery & Data Management Team.

Editor: We are witness to an emerging wave of change in electronic discovery. What is driving these changes?

Kane: Primarily cost, and to a lesser extent, fear of sanctions. Cost is driven by the size and nature of document populations in modern discovery. Electronic document populations are typically much larger than those at issue historically in paper discovery cases. The needle is buried in a larger haystack. The other cost driver is the ephemeral nature of electronic data, i.e., it is not fixed, making it harder to deal with. The significant cost of collecting and reviewing electronically stored information in modern litigation is spurring approaches that seek to limit rising discovery costs, but which do so in a reasonable and defensible manner.

Editor: Would it be fair to say that e-discovery as a business process and as currently conducted is actually broken?

Kane: I don't know if I would say it is broken, but it's far from ideal. Current practices are not cost efficient. It's a challenge to keep electronic discovery costs from becoming a determining factor in litigation, particularly in cases where the amount in controversy does not justify "bet the company" type efforts. The current process has to catch up to the demands of modern practice – not just the demand to meet discovery obligations, but also the demand to do so in a cost-effective manner.

Editor: Squire Sanders has announced a new approach to electronic discovery – Intelligent Discovery. What is behind your approach?

Kane: Our approach seeks to reduce client costs by limiting, to the greatest extent possible, human review of large data populations. Modern litigation delivers a tsunami of data.

The cost of reviewing it in traditional fashion – linear page-by-page review by teams of human reviewers – is becoming prohibitively expensive. We combine review by experienced trial lawyers with automated analysis and prioritization of documents, as well as sampling and other best practices, to reduce document review costs. You can never completely eliminate human review, nor would we advocate that. But we can significantly reduce its scope and still deliver reliable, defensible, high-quality review results.

Editor: Is Intelligent Discovery mainly technique or mainly technology, or both?

Kane: It's both. One without the other is not sufficient to achieve the desired result. Certainly, there's a strong technology component insofar as you are using technology to leverage the efforts of experienced human reviewers. But it's not sufficient only to rely on technology – you need to make educated, reasoned decisions about how best to apply the technology as part of a broader workflow. You also need to ensure that you're employing the technology in a way that is reasonable and defensible. Combining the technology with a well thought out process lets you deliver enhanced review quality in a cost-effective manner.

Editor: Please describe the technology you are using and how it works.

Kane: We are not wed to any particular technology. The best solution can vary from case to case and often does. A product that is well-suited to one type of case may not be the best option for another case with different needs. One technology that fits very well with our Intelligent Discovery approach is Equivio>Relevance. We have had great results with it. It's a powerful tool that uses predictive coding and classification rather than discrete keyword searches. Unlike keyword searches, Equivio>Relevance takes into account all the words in a document along with the relationship of the words to one another to determine what is and what is not likely to be relevant. Equivio>Relevance lets

us leverage the results of review by experienced trial attorneys across very large document populations.

Editor: Is it organized around a compound group of algorithms that elicit the right kind of information?

Kane: That is exactly right. It's organized around algorithms that allow classification by concept, rather than specified keywords, through interaction with an experienced human reviewer. The technology is "trained" by a human reviewer telling it what is relevant in samples of the data set. This interactive training process continues in iterative fashion through the review of multiple sample sets. This allows the importance of document terms, and their relationship with one another, to be identified and refined through the search algorithms. The technology then leverages this training across the entire document population to rank and prioritize documents according to predicted relevance.

Editor: How much time does it normally take to train the system?

Kane: It can vary from case to case, depending on the complexity and the types of documents. It might be as little as ten to fifteen hours or as many as thirty or forty hours, depending on the pace of the document review and the "richness" of relevant documents within a data set. Equivio>Relevance generates sample sets in batches of forty documents. We haven't tried to determine exact mathematical averages, but we're seeing forty-five passes of forty documents as a rough benchmark. The technology eventually reaches a point where the review is stabilized, essentially telling you that the technology can't "learn" anymore from the review of additional sample sets.

Editor: How do you know when you do have a statistically significant group of documents reviewed?

Kane: With the Equivio>Relevance tool, the technology will tell you when the review has

Please email the interviewee at skane@ssd.com with questions about this interview.

stabilized, indicating that the review of additional sample sets will not materially influence the review results. The point of stabilization is always after review of a statistically significant number of documents. Because the review of the sample sets is iterative, the technology is refining the predicted importance of terms. When no further refinement is possible, the review of sample batches ends.

Editor: Do you have your technology staff present at the same time the lawyers are doing the review and doing the initial sample for the machine?

Kane: We routinely involve technology professionals in e-discovery projects in addition to the lawyers to assist with the overall workflow, but not to participate in the actual document review. The Equivio>Relevance system is easy to use so that you don't need an IT person or a litigation support professional looking over your shoulder, but certainly we want them involved in other aspects of large e-discovery projects.

Editor: I assume Equivio does cut your costs by eliminating some personnel in the review process.

Kane: Absolutely. It cuts costs by dramatically reducing the number of reviewer hours. You can find and prioritize the relevant documents much quicker than through other methods.

Editor: What are the benefits of this approach over the legacy keyword approach?

Kane: Traditional keyword searching is both under-inclusive and over-inclusive. You can't identify all the relevant terms on the front end, and the ones you do identify will generate false positives – i.e., documents that contain the keyword but are not relevant. Equivio>Relevance and similar technologies look at all the terms in a document as well as their relationship to each other. If you have a dispute that centers around a commercial transaction called "Project Maple," then "Maple" will be an important term. A traditional keyword search would locate documents containing that phrase, but would also return documents addressed to someone on "1000 Maple Street." Equivio and other predictive coding technologies go beyond looking at individual terms and would "learn" that Maple is an important concept, but only in relation to "Project" and not in relation to "Street." This is a very simplistic example, but helps understand what Equivio does with all the words in all the documents. Keyword searching can't replicate that.

Editor: The billable hour has reigned supreme in litigation for as long as law firms have existed. Do you think this will diminish the importance of the billable hour?

Kane: I do. I am reluctant to say that the billable hour is ever coming to an end because, although we all roundly criticize it, we haven't identified a fully translatable alternative. Whatever its other faults, the billable hour is easy to understand and measure. In the discovery arena, with tools like Equivio>Relevance, we're moving to a process that allows us to review in a way that is not so dependent on the number of hours expended. If you deploy these tools effectively, a client is buying expertise and results rather than just hours of review time. We like it because it aligns the interests of the client and counsel to deliver high-quality results at a relatively lower cost.

Editor: Cost is, of course, one thing, but is the use of predictive coding technology defensible? I understand it's never really been tested in the courts.

Kane: It is defensible. There is no reported decision that explicitly blesses the use of predictive coding. But the ideas behind it are well-recognized. For example, the Advisory Committee Notes to Rule 502 of the Federal Rules of Evidence favorably reference the use of advanced analytical and linguistic tools to identify potentially privileged documents. Numerous cases and other authorities, including the Advisory Committee Notes to the 2006 amendments to the Federal Rules of Civil Procedure, discuss the importance of sampling data populations. Predictive coding is essentially a very powerful and sophisticated method of sampling large document sets to determine in an efficient way what is likely to be relevant. Our Intelligent Discovery approach does not rely on the use of technology alone. If you apply technology in an informed manner, in a way that is reasonable and is intended in good faith to identify relevant documents, it is just as defensible as human review if not more so. Empirical studies will tell you that human review locates fewer relevant documents from a data set than the well-applied use of technological approaches.

Editor: How is the market responding to this new approach?

Kane: I would say the market remains in a nascent stage. Everyone is keenly interested in this technology and its potential and it's hard to find a vehement critic. But the interest in the technology has not translated into its actual employment in the e-discovery market. Lawyers are inherently risk adverse and they worry about possible challenges to new

approaches. My view is that if you employ the technology effectively, and also a broader process to make sure that your overall review is diligent and reasonable, then you shouldn't worry about a potential challenge. We have considered the issues carefully and we think the use of these approaches is defensible and in our clients' best interest now, particularly with respect to reducing discovery costs.

Editor: How did you test it initially?

Kane: We had a class action case that involved review in traditional fashion of a very large population of documents. The case settled after review was complete and just before documents were to be produced. Using this sample, we tested Equivio's technology by applying it to a portion of that same document population to see how it would measure up against our previous review. An attorney with an extensive knowledge of the case did the Equivio-assisted review, which involved the review of 42 passes of 40 documents each. We found that Equivio matched up well to the prior human review and located relevant documents effectively in a fraction of the time. There were instances where the prior human review team and Equivio differed as to whether a document was likely to be relevant. Focusing on those documents, we found that Equivio was just as likely to be "right" concerning relevance as the human reviewer. Subsequent reviews and tests have borne this out. No review approach is perfect. But the use of advanced technology, employed in a reasonable and good faith overall process, is as good or better at finding relevant documents as traditional human review.

Editor: In ten years' time, we might look back and view firms like yours as the firms that took the lead in changing the history of e-discovery. Who would you say are the chief beneficiaries of this change – the law firm, the client, or maybe even the justice system?

Kane: I think clients are the ultimate beneficiaries of this approach. They're paying the bills for the cost of litigating commercial disputes. We want to offer them an approach that allows the cost of discovery and document review to be as low as reasonably possible. To the extent we can employ advanced technologies to do so, my clients are better served because we are meeting discovery obligations in a less expensive and more productive way. I think you can make a good argument that this improves the entire civil justice system. Regardless of differences in any particular matter, clients as a whole – both mine and my opponent's – benefit if we can resolve commercial disputes more efficiently and less expensively.