Emerging Technology

Technology-Assisted Review Panels Draw Crowds at Advanced eDiscovery Institute

George-town Law Continuing Legal Education presented its 10th Annual Advanced eDiscovery Institute Nov. 21 and 22 in McLean, Va. The two-day conference explored hot topics in the eDiscovery arena, such as information governance, ESI and privilege, and the proposed amendments to the Federal Rules of Civil Procedure. The event also looked to the future of eDiscovery with several in-depth panels on predictive coding and technology-assisted review.

TAR Application Issues

The first panel, titled “TAR 201: Determining When Technology-Assisted Review Is or Isn’t the Right Solution,” was moderated by Thomas Gricks, of Schnader Harrison Segal & Lewis LLP, who led the panelists in a lively conversation about TAR in front of a packed audience. University of Waterloo Professor Gordon V. Cormack, Scott A. Kane, of Squire Sanders, and Dera J. Nevin, of TD Bank Group, spoke regarding the dos and don’ts of using TAR.

Gricks opened the session by asking the speakers to consider the various application issues surrounding the use of TAR. Gricks said that while TAR is generally a good product for culling documents, many people question whether it is usable in the area of privilege review.

“Since TAR focuses mostly on text, how does it work with privilege-based documents which focus mostly on who is speaking?” Gricks asked the panel.

Does using TAR constitute taking reasonable steps towards preventing disclosure under Federal Rule of Evidence 502(b)?

Cormack said he uses TAR frequently for privilege review, noting that in terms of culling, the legal teams he works with do not let any documents out the door unless they also take a look at them, so, according to Cormack, they may as well use TAR to review the documents for privilege prior to that eyeballs-on review.

But how does this play out in the event of an unintentional disclosure of privileged information? Gricks asked Kane if he believed using TAR constituted taking reasonable steps towards preventing disclosure under Federal Rule of Evidence 502(b).

Kane said in terms of “reasonable steps,” the devil is in the details. According to Kane, using TAR to find privileged documents is likely going to be as effective or more effective than using contract lawyers.

“However, the technology cannot distinguish between an internal memo that analyzes your exposure and a letter to the other side that discusses those very same issues,” Kane noted, pointing out the difficulty that can be encountered when using predictive coding.

When asked if she advocated the use of TAR for privilege review, Nevin stated, “I prefer when people use all available resources, including TAR.”

TAR and Early Case Assessment. Gricks then asked the panelists to discuss whether TAR is an ideal tool for early case assessment. Kane spoke to the value of using TAR for ECA because the technology creates a ranking of documents that is useful in determining the risks and benefits in the early stages of case management. According to Kane, crucially important documents “bubble up to the top” and “hot documents” are found more efficiently and quicker because all of the unwanted documents have been eliminated by the system.

Nevin touted the benefit of TAR as a great quality control tool, and also stated that TAR has helped her inform her strategy in 26(f) conferences in a few cases.

Situational Issues

Gricks then switched gears, asking the speakers to consider various situational issues. He began by asking all three what they considered to be a minimum and maximum document population in terms of collection size for TAR.

Kane said 75,000, noting that TAR users get greater efficacy the larger the document population. Gricks seconded that statement.

“From an in-house perspective, there is no bench mark we’ve set, but my number tends to be the same,” Nevin said, adding that she had never used TAR on a document set lower than 90,000.

In terms of the high end and how well TAR scales, Cormack said the biggest obstacle to using TAR is money.

The speakers then addressed the use of TAR to find relevant documents that contain mixed text and non-text.

“When we do TAR, the first step we do is generally get rid of documents that are not text,” Gricks said. “But what about patent numbers, and metadata, can TAR be used for that too?”
Cormack explained that TAR is not necessarily restricted to text, as the tool requires you to essentially identify various features of each document. “You can include distinguishing words like the ‘to’ and ‘from’ in an address, as well as numbers,” Cormack said. “You can also include voice transcription sequences of phone-ins.”

The technology is not just text-based, and, according to Cormack, he would not want to use a tool that could not index metadata information such as the “to” and “from” data.

However, despite the myriad uses for TAR, Cormack did state that he would never use it for information governance. Kane, on the other hand, said ‘I think it works better than the alternatives.’

Which Training Method? Gricks asked the speakers what method of training the tool they preferred, offering the three choices of random, seed set, and multimodal, which includes a combination of seed sets, exemplars, and random selection.

Kane said he has a philosophical objection to seed sets, which he believes is a method for simply telling the tool what to do, rather than training it. Kane acknowledged that a while ago he would have advocated for the use of random training, but now he prefers multimodal.

“I want a strong element of randomness, but our multi-modal wouldn’t start with a seed set, but rather with a quality control sampling,” Kane explained.

Cormack disagreed with using a seed set as well, saying he would always try to find as diverse a set of exemplars as possible, including non-relevant documents.

Sets With Poor Richness. Gricks also asked the panelists how they felt about using TAR on document sets with poor richness. Richness (or prevalence as it is often called) relates to how many relevant documents are contained within a data collection—the greater the richness, the more relevant documents are in a collection. In addition, richness can affect the ease in which relevant documents can be found, depending on the search methodology.

Gricks advocated for increasing richness prior to using TAR, so as to make TAR more effective. However, Cormack disagreed, arguing that TAR works well on low richness data sets, and that there is no need to increase the richness because pre-TAR actions risk leaving behind relevant documents that the tool would have otherwise identified. Cormack argued that if a TAR tool does not work on low richness data sets, the tool should be replaced.

Disclosure Issues. Gricks then opened the discussion to the topic of disclosure, asking Nevin if she believes she is required to disclose the use of TAR to the opposite party. “We were using TAR years before Da Silva Moore, and we didn’t know we were supposed to tell people,” Nevin said. “I would not always disclose its use all the time, but it has to be consistent with 26(f) strategy.”

The session closed with a discussion of preparation and training, in which Gricks asked Kane if he prefers to de-duplicate documents across the entire universe or just within particular custodians. Kane said he prefers to de-dupe across custodians and then capture the additional custodian information on the front end for later production spec.

Cormack assured Gricks that even if duplicates come up in the process of trying to train the tool, the duplicates will not cause the over-emphasis of documents within the model.

The TAR Plenary Session
The second full session devoted to TAR was moderated by Magistrate Judge James C. Francis of the Southern District of New York and was titled Technology-Assisted Review One Year After Da Silva Moore.” Tracy Greer, of the Antitrust Division at the Department of Justice, Jerome J. English, Director of eDiscovery at Intel Corp., and Martin T. Tully, of Katten Muchin Rosenman LLP, were the panelists.

While the session focused more on the future of the tool than the first session did, the panel opened with a discussion of how TAR is most useful, including addressing Cormack’s statement that he would never use the tool for information governance.

“I don’t agree with Gordon’s statement that you shouldn’t use it for information governance,” Tully said. “Jason Baron [former Director of Litigation at the U.S. National Archives and Records Administration] said TAR holds the most promise for records management rather than for relevance . . . If you think about the type of principles that go into TAR, its not a big leap to consider how you can use the tool to help organizations wrestle with the burgeoning amounts of data that multiply every day.”

Tully also noted that as Baron previously commented, the use of TAR for creating defensible destruction plans is likely less challenged than it is in the context of culling documents.

Judge Francis asked the speakers if the tool can be used in the front end of litigation, such as during the early stages of an investigation.

Tully said absolutely in terms of ECA, and, in response to the question of whether it could be used for privilege reviews, English gave a response echoing Nevin’s in the previous day’s session.

“If you think about the type of principles that go into TAR, its not a big leap to consider how you can use the tool to help organizations wrestle with the burgeoning amounts of data that multiply every day.”

Martin T. Tully
Katten Muchin Rosenman LLP

“I would use TAR in any circumstance where I would normally put non-trial team eyeballs on documents,” English said. “An average custodian has 20,000 to 200,000 documents, and we have reached the point where we can no longer have the trial team do the review.”

English said he has to train individuals to understand the litigation and privilege issues. According to English, he has as much faith in the tool as he does the trained individuals. Nonetheless, English did state that at some
point, he does use human review as part of his quality control process.

Judge Francis then asked the speakers to consider how transparent parties should be in the use of TAR. English said it is enough to disclose the process—i.e. the use of TAR—and that he would not necessarily feel obligated to disclose seed sets or non-responsive documents. Greer said she asks opposing parties for a look at the statistically significant sample of the non-relevant data, noting this is a “big ask.”

Tully joined the discussion by opining that the use of TAR has possibly raised the standard on how discovery has been executed all along.

“In order for my opponent to be comfortable with my use of TAR, I’m going to have to be more translucent to get to an agreement,” Tully said. “If I’m ultimately not able to reach an agreement and I have to prove to a judge that what I did was defensible, I may have to turn over a little more than I want to.”

Baron, as an audience member, spoke up on the issue of transparency by stating that parties are under no obligation to share, but that he advocates for cooperation.

TAR in Front of Today’s and Tomorrow’s Judges. The next topic of discussion—whether TAR tools are required to pass a Daubert test—garnered responses from the panel and the audience.

Tully said he didn’t believe the tool’s validation need come as close as a Daubert test. Magistrate Judge David Waxse, of the District of Kansas, also commented from the audience, disagreeing with Tully.

“Federal Rule of Evidence 702 applies to the use of expert testimonies on these kinds of issues, and there is no exception for discovery,” Judge Waxse said. “While some say this isn’t evidence, so we don’t have to follow the Rules, the rules don’t focus on evidence but on the presentation of evidence to a judge.”

Judge Waxse explained that parties should not ignore the Rules, which are set up to promote validity and to allow judges to determine if parties are acting appropriately during the trial process.

The session concluded by discussing the future of TAR and where attorneys can expect TAR to go technologically. English said that the idea that TAR is going to “go back in the box” is false and that it is the only way lawyers will be able to manage ever-increasing ESI demands.

Tully proposed that the future use of TAR will require attorneys to become more comfortable with the tool, and that a few specific modules will become the most widely used.

BY TERA E. BROSTOFF

To contact the reporter on this story: Tera Brostoff in Washington at tbrostoff@bna.com.

To contact the editor responsible for this story: Carol Eoannou at ceoannou@bna.com.