Technology Assisted Review

EDRM Duke Law Sets Sights on Technology Assisted Review Guidance

By Michael Greene

An organization that develops models and standards for electronic discovery has set its sights on developing guidance on technology assisted review (TAR)—a process that involves using machine learning models to help classify documents.

EDRM (Electronic Discovery Reference Model), which is part of the Duke Law Center for Judicial Studies, is working with with members of the plaintiffs' and defense bar, as well the judiciary, in drafting, editing, and reviewing best practices or protocols that are intended to provide more direction on how to use TAR.

TAR can potentially help companies in litigation more efficiently identify responsive documents and reduce their discovery costs. Yet, nearly five years after the first written court decision approving the use of TAR for managing document production, many attorneys and their clients still view the the process as a too much of a "black box" and lack confidence that it will timely produce reliable results.

"There seems to be an unnatural fear of TAR by some parties but it has been proven as a very effective tool given case size, timeline and financial resources," said Jim Waldron, Director of EDRM Duke Law, told Bloomberg BNA in email. EDRM's hope is "that adding our presence in the dialog would be helpful to a better understanding of TAR," he added.

TAR, which is sometimes referred to as computer-assisted review or predictive coding, can be a much more sophisticated way of searching for responsive materials than just using key word searches and manually reviewing. Instead, it is a process in which software-based tools are used to cull and prioritize a large data set by training the tool on what type of documents are being sought.

More guidance on TAR could narrow the often wide knowledge gap between those who have experience using the process and those who don't. It also could alleviate concerns that TAR might increase the uncertainly over what must be produced, and reduce the likelihood that using it leads to more discovery disputes.

"Clearly there are differing opinions (even among 'experts') when it comes to when, how, and how much TAR is best utilized," Magistrate Judge Douglas Arpert of the U.S. District Court for the District of New Jersey told Bloomberg BNA in an email.

"Proponents of TAR must be knowledgeable and able to clearly communicate its benefits and limitations," he said. "A central concern is that a dispute over whether to use TAR or which protocols to use has the potential to create another tier of litigation, adding (rather than reducing) costs and delay."

Challenging Task Generating a consensus among practitioners about what TAR protocols or best practices should look like won't be an easy task. EDRM held a conference Sept. 7-8 in Arlington, VA, where attorneys, judges, and industry experts voiced differing opinions on TAR-related topics.

During EDRM's Sept. conference, concerns were raised over whether any specific standards could become obsolete before they are even published. Technology related to e-discovery is developing rapidly, and electronically stored information is more commonly appearing in non-text formats that most current TAR models aren't able to process, such as audio, video, or symbols like emojis.

Conversely, other speakers at the event suggested that guidance could be helpful if it is simple, practical, and flexible so that it could be applied to the demands of different types of cases.

(The event was held under the "Chatham House Rule," whereby participants agreed not to attribute statements to speakers by name.)

Among the topics that EDRM could take up in its guidance include when TAR should be used, what type of disclosures need to be made, how a TAR-based production can properly checked by the recipient for accuracy, and how it can be used in a cost-effective and timely manner. It also could address whether litigants receiving documents should be allowed an "escape hatch" or a "second bite at the apple" if they aren't satisfied with TAR-based production.

"I think there is general consensus on how to define TAR as a process. Different views may exist on which form of TAR and which algorithm does TAR better. But I think those differences may be explained in fairly simple terms," Michael Quartararo, Director of Litigation Support Services at Stroock & Stroock & Lavan LLP told Bloomberg BNA in an email. Quartararo is a co-leader of the EDRM TAR project.

"The larger sticking point is going to be the issue of transparency. I think both the plaintiffs' bar and the defendants bar are going to need to accept some compromise on this issue," Quartararo said.

Need for Best Practices Meanwhile, others suggested the need for agreement on a uniform definition of TAR and on identifying which methods are more useful.

"One thing that I took away from the conference is the need to reach consensus on terminology," Robert Keeling, a Washington-based partner at Sidley Austin LLP, who is co-chair of his firm's e-discovery task force, told Bloomberg BNA. "The use of the word 'TAR' is one example. I would define TAR to mean the application of supervised machine learning (whether passive or active) to a data set—what is commonly referred to as predictive coding."

Others define TAR to include predictive coding plus other types of analytics, which "risks muddying the waters unnecessarily," he said. "But the legal community (including the vendors servicing the legal community) have not yet reached consensus on this rather fundamental point."

"Some of the commonly used TAR methods have been shown to be effective; others have not," Maura Grossman, who is a research professor at the University of Waterloo and an eDiscovery attorney and consultant in New York, told Bloomberg BNA in an email.

Grossman's and Waterloo professor Gordon Cormack's seminal research on TAR has been widely cited as showing that the process can provide an effective way to review documents.

"It would be useful to the bench and bar to identify those TAR methods that have been shown to be most effective according to sound empirical evidence and to describe potential protocols for their use, but it would be premature to prescribe guidelines or best practices in an area where the technology is still in the process of development and there is limited research to support such prescription," Grossman said

"There is inherent difficulty with creating static 'best practices' in the application of TAR because specific best practices may change depending on the particular matter and/or the data at issue," Keeling said. "And using a slightly less efficient TAR approach does not necessarily make that approach wrong."

"Because of this, continuing to educate the bar and the courts about TAR will likely result in the greatest benefit. In that regard, general guidelines would be useful to help further educate the legal community," he added

To contact the reporter on this story: Michael Greene in Washington at mGreene@bna.com

To contact the editor responsible for this story: S. Ethan Bowers at sbowers@bna.com