E-DISCOVERY

AI: E-DISCOVERY GETS SMARTER

E-discovery does not sit still. To provide high-level service, practitioners necessarily deal with legal technology at the bleeding edge of development. This involves the embrace of nascent artificial intelligence (AI) in combination with other analytic tools and techniques to tackle increasingly challenging discovery projects. As ever-expanding volumes and sources of information strain the capacity of counsel to manage discovery, AI is coming just in time.

AI is the subject of much hype and misunderstanding. Some companies refer to all of their software offerings as AI, making it no more than a marketing term. At base, however, the term refers to "technologies that can mimic and enhance human thought processes and capabilities," says John Davis, senior counsel at Crowell & Moring and co-chair of the firm's E-Discovery & Information Management Group. While there is no true thinking machine with self-awareness, there are educable tools that perform a fair imitation. Used by experienced practitioners, he says, "AI can be a real boon to discovery in litigation and investigations as well as transactional inquiries, leading to quicker, more accurate, and defensible results."

Two types of AI that are having a significant impact on e-discovery are machine learning and natural language processing (NLP). Machine learning, as the name suggests, uses mathematical models to assess enormous datasets and "learn" from feedback and exposure to additional information. This enables the models to uncover hidden patterns and make predictions or determinations on their own about targeted data. NLP enables computers to effectively communicate in the same language as their users, advancing the ability of the machines to understand written and spoken human language and more closely approximate human cognitive patterns.

Increasingly powerful analytics have also expanded the scope of tasks that can be automated, as well as the types of possible searches and analyses. Today's e-discovery and compliance tools can tease out hidden patterns in the text fragments and disassociated communications of millions of electronic files to categorize and cluster documents by concepts, content, or topic. For example, AI-fueled "sentiment analysis" goes beyond term searches to look for indicators of relevant behavior, such as concealment, deceit, panic, or concern. "AI is reaching the point where the technology can even identify facial expressions and voice patterns in videos and recordings that point to certain sentiments. This, in conjunction with analyses of subjects' writings and transactional data, can form a fuller picture of individual and group conduct," says Davis.

AI systems can also search for anomalies—"irregular occurrences or omissions, things that are or are not there, contrary to expectations," says Davis. "People are now more guarded about how they communicate in emails. They may avoid emailing about a sensitive subject or use a different terminology or channel. These analytics help you look for out-of-character communications, code language, or patterns that point toward underlying meaning. For example, if someone who is usually chatty in texts suddenly sends one saying, 'Just call me on my cell,' the system can flag that." It can also find suspicious gaps in communication frequency that can raise red flags for further inquiry or signal failures of production or destruction of evidence.

Even at this relatively early stage, AI has a proper place in the discovery tool kit. "It's not yet the stuff of science fiction, where sentient robots are going to replace all the lawyers," says Davis. Instead, "AI gives counsel and clients more leverage with large sets of data. It extends their reach and allows them to work faster and more efficiently, with higher confidence in quality."

AUTHORITIES ARE SIGNALING ACCEPTANCE

Courts and regulators continue to be open to the use of advanced technology in e-discovery, with some preferring it to conventional review. For example, in *U.S. ex rel. Proctor v. Safeway, Inc.*, the plaintiff objected to Safeway's production of 575,000 documents based on a keyword screen and produced without review. In March 2018, the U.S. District Court for the Central District of Illinois agreed that Safeway's document dump failed to meet its Rule 26(g) obligation to make a reasonable inquiry and certify the production as complete and responsive, but declined to require a document-by-document review. Instead, the court ordered Safeway to use a technology assisted review process to identify likely responsive documents and then review them for production. The Antitrust Division of the Department of Justice has issued guidance similarly noting its preference for TAR over keywords.

Also significant was the Northern District of Illinois's decision in *In Re Broiler Chicken Antitrust Litigation* in January 2018, where the court adopted a detailed process for validating the use of machine learning-based TAR in identifying likely relevant documents in massive datasets. "This is a robust protocol

that, while probably more than is needed for many cases, predictably will be influential in the courts. It gives comprehensible direction for acceptable workflows and levels of transparency, so courts and parties won't have to think as hard about a technical topic," says Davis. "We see now that the debate has moved from whether these technologies are acceptable or not to how TAR should best be implemented to assure reliability."

That question is likely to be a key issue going forward, as AI becomes more prevalent and sophisticated. While the traditional use of search terms in discovery is well understood, AI technology is a "black box" to most observers. It can be nearly impossible to reconstruct how the machine makes decisions about data. Even knowledge of the code in abstract would not be revealing, as the algorithms react to input (the dataset and human feedback), which is different for every matter and provokes adaptation through the learning process. "We've gotten to the point where few people, including many experts, really understand the math and the technology underlying these AI search capabilities," says Davis. However, the stakes are high, and courts and parties will continue to seek clarity—and counsel will need to be there with answers. "I can see a push from industry circles and experts toward more transparency and standardization in AI operations," says Davis. Expansions of unmonitored AI applications and concerns about potential bias in AI decision-making are likely to fuel that trend. "Validation exercises alone may not be sufficient. We may see AI methodology being subject to something like the *Daubert* standard, requiring expert testimony."

Meanwhile, AI will continue to progress. For example, says Davis, "next-generation AI will aid in integrating disparate types of information, such as audio, video, and transactional, and be better able to recognize languages and dialects through natural language processing. This will enable attorneys to ask the machine more semantically complex questions and receive nuanced responses organized across information types." The ability to tie differentiated datasets together into a comprehensible whole is becoming more important as attorneys work with more streams of information, under more exacting standards and timelines.

"Advances in AI will enable the software to anticipate and suggest complex questions that may be applied to the data for a variety of circumstances. It will permit better search and understanding of discovery information and will get attorneys closer to the answers that matter," Davis continues. "AI technology will save money and provide better results. It needs to be considered for any complex e-discovery strategy."

E-DISCOVERY MEETS DATA PRIVACY

Since the EU's General Data Protection Regulation went into effect in May 2018, its impact has been felt in everything from sales and marketing to finance and compliance—and the legal department.

The GDPR imposes restrictions on the use of the personal information of EU data subjects. For companies with operations and data in Europe, the GDPR creates challenges for discovery in U.S. courts. For example, the GDPR in many ways encourages controllers and processors to restrict the amount of personal information processed to only that which is needed, and to justify such use. "This raises the difficulty level in transferring personal data from the EU to the U.S., which is not considered to offer comparable protections," says Crowell & Moring's John Davis. "Although the GDPR did not significantly change pre-existing restrictions and exemptions for transfer, the enhanced process and potential penalties for non-compliance have really focused attention. We are likely to see an accelerating GDPR impact in terms of reduced amounts of data coming from Europe through the discovery process."

While keeping an eye on European regulations, companies also have to comply with U.S. discovery orders—which can be something of a balancing act. "Counsel should be sure to educate courts and requesting parties about the particular burdens and barriers involved in sourcing data from overseas, and get them involved in creative solutions. Certainly, GDPR effects are relevant for proportionality arguments as well as in discussing the scope and staging of discovery," Davis says.

Managing such issues in cross-border matters "can be intensely complicated," he adds. The U.S. also has its share of information restrictions, and more can be expected at both the federal and state level. The California Consumer Privacy Act of 2018 is already influencing other authorities to act similarly. "These developments have raised the bar for counsel. Now more than ever, it is important to be thoughtful in dealing with personal information in discovery," Davis says.