THE LAWYER'S DAILY

Information Technology

Experts serve up primer on technology assisted review

By **Carolyn Gruske**

(November 23, 2017, 2:34 PM EST) -- If you're not already using technology assisted review (TAR), now is the time to start learning about it.

(l-r) Constantine Pappas, Maura Grossman and Andrew Jay Peck

That's the best way to sum up the panel discussion "A Non-Linear Approach to Change" held at the Technology in Practice conference Nov. 8 in Toronto. The panellists were the University of Waterloo (UW) research professor dubbed the "Queen of TAR" Maura Grossman, Magistrate Judge Andrew Jay Peck of the United States District Court, Southern District of New York and Constantine Pappas, manager of Relativity Solutions at e-discovery software producer Relativity.

Grossman, who is a lawyer in addition to being a member of UW's computer science department, is the co-creator — along with another UW professor Gordon Cormack, who was working in the field of spam filtering — of the TAR concept. That happened when they published the paper "Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review" in a 2011 edition of the *Richmond Journal of Law and Technology*. They followed that up with a subsequent paper in 2013 to further define TAR.

Even though she originated the approach, Grossman said users have to be aware of its limitations and understand what it really is.

"TAR is not synonymous for good. It also doesn't mean any technology that anybody uses in connection with a review. You have to be very careful," she said.

"You want to match your tool with your data and your problem. So if you've got lots of photographs, TAR won't be the answer."

According to Grossman, programs only fall under the TAR banner if they either categorize or prioritize large sets of documents. Technology that only performs threading or clustering doesn't qualify under her definition.

Additionally, TAR "software learns to distinguish between the relevant document and the not relevant documents in the same way a spam filter distinguishes between spam and mail. And it gets better as it continues to apply what it learns," she said.

Grossman explained that TAR tools could be active or passive. Passive TAR means a human suggests the initial group of documents that the software uses to learn from, while an active TAR system lets the algorithm pick the starter set. TAR tools can also be divided into what people call TAR 1.0 (a

simple system) and TAR 2.0 (a continuous system). Grossman said under a simple 1.0 system there is a stop and a separation between the training phase and the review phase (during which the system learns what is and what isn't a relevant document). The continuous 2.0 system doesn't stop, and instead performs an ongoing cycle of review, feeds it back into the system, review, feeds it back into the system.

"You keep reviewing and training. From day one the system is learning and it never stops."

Grossman said lawyers and judges need to have some basic familiarity with the types of TAR systems because they "have important workflow implications ... some allow you to get started right away and some take a couple of weeks before you train the algorithm. It really is important to know it's not just one thing."

Grossman's articles formed the basis for Judge Peck to issue what is considered to be the first opinion on TAR in the world in 2012 when he released his opinion on *Da Silva Moore v. Publicis Groupe et al.* In it, he approved the use of TAR and "sided for the defendant's protocol for the defendant's review" working under the assumption that "the producing party is in the best position to know how to do a review." Peck also issued another key TAR judgment in 2015 for *Rio Tinto PLC v. Vale S.A* and offered some guidelines on what he referred to at the time as predictive coding. Grossman acted as a special master in the case.

While Peck was willing to delve deeply into how TAR works, he cautioned that not everybody has the same level of familiarity with the system, including judges. That means sometimes lawyers need to perform some rudimentary education, and Peck suggested starting with familiar applications.

"Talk about Pandora. Forget the code. You go to Pandora the first time and you say 'I like Billy Joel' and Pandora will play some Billy Joel songs for you and then it will throw in an Elton John song and you give it a thumbs up or down and Pandora will learn your musical taste from that. That's in essence what predictive coding is doing with responsive and non-responsive documents."

Pappas said at Relativity, he and his colleagues often find lawyers putting up barriers in front of their ability to learn about TAR.

"They tell us: 'We want this, but we want to wait for the right case.' But why aren't you gaining experience and expertise on some of these smaller cases where it is a no-brainer? There is no right case as long as the documents and the text are amenable to the technology. It's the right case," he said.

"If it's your first time using the tool, it is great for prioritizing. It gives you the most relevant documents."

Pappas added that if firms are engaged in "rolling production" (sharing documents with opposing counsel on a timed release schedule as opposed to doing a massive dump at the end of the discovery process) TAR ensures the most important documents are delivered at the start. He also said that there is nothing to prevent firms from still having reviewers look over every document before it goes out.

"You can say to the court 'we've had eyes on everything.' You save the most money when you don't do that, but that's a simple way to get more comfortable with it."

He added that by having a TAR system serve up the most relevant documents first, it's going to help lawyers not get bored with junk documents right from the start. They will be more engaged, which will help generate buy-in for future TAR projects.

Grossman too said there are no size barriers when it comes to starting to use TAR. She has used it on cases as small as 9,000 documents and cases as large as 54 million.

"For most commercial products you are going to go out and pay for in a cloud model. I've heard the number 50,000 or so is where it starts to get economical but there are lower end tools where you can do 15,000 documents," she said.

As for the legal professionals who feel too intimidated to try TAR, the panellists all feel that's not a legitimate excuse.

"I think the competence rules, at least in the United States, say that's not acceptable. You can't really say 'I don't understand default swaps' or 'I don't really understand trusts and estates.' If you're doing discovery, you can't say that," said Grossman.

"One of the takeaways is you do your homework before you're in a matter where you want to use this. You spend some time playing with the tool so you're not seeing it for the first time in a moment of crisis."