Trial Briefs

The newsletter of the Illinois State Bar Association's Section on Civil Practice & Procedure

AI in Discovery? It's Called TAR

BY GEORGE BELLAS

Although artificial intelligence ("AI") has recently entered the mainstream conversation, AI has had a place in the realm of e-discovery since 2005.[1] Technology assisted review ("TAR")[2] uses a technique called predicative coding, which is the ranking of documents based on how potentially relevant they are to the matter. The process involves training computer software to classify documents based on inputs from expert reviewers and expedites the organization and prioritization of document collections.[3]

Discovery of electronically stored information ("ESI") has become a primary focus in discovery as electronic data proliferates. E-discovery tools have likewise been developed to include advanced search technologies to include keyword, concept, predictive coding generally referred to as TAR. TAR is used to supplement and extend human review to identify, review and disclose ESI. TAR is a form of artificial intelligence in which the artificial intelligence machine learning assists in the review process by automatically reviewing, sorting, and classifying large volumes of documents based on whether the documents are likely responsive to discovery requests. TAR allows the human reviewers to examine the most likely to be relevant documents first and can avoid the need to review at all documents predicted by the computer to not be relevant. As the Sedona Conference has stated:

In just a few short years, the use of technology-assisted review (TAR) for the exploration and classification of large document collections in civil litigation has evolved from a theoretical possibility to an essential tool in the litigator's toolbox.[4]

Despite its flaws, many senior lawyers (and some clients) still consider manual review to be the "gold standard" against which other review techniques are compared. While the volume of electronically stored information (and concomitant expense) has largely eliminated manual review as the sole method of document review, manual review remains used along with, for example, keyword screening.

Considering whether manual review is the gold standard, two recent research studies clearly demonstrate that computerized searches are at least as accurate, if not more so, than manual review and concluded that "[o]n every measure, the performance of the two computer systems was at least as accurate (measured against the original review) as that of human re-review."[5] Professors Maura Grossman and Gordon Cormack, while using data from the Text Retrieval Conference Legal Track, concluded that "[t]he idea that exhaustive manual review is the most effective – and therefore the most defensible – approach to document review is strongly refuted. Technology-assisted review can (and does) yield more accurate results than exhaustive manual review, with much lower effort."[6] Grossman and Cormack noted that "not all technology-assisted reviews. . . are created equal" and that future studies will be needed to "address which technology-assisted review process(es) will improve most on manual review."[7]

Studies have shown that technology-assisted document review is at least 50 times more efficient than human only or manual review.[8] Moreover, TAR is accepted and used by governmental agencies such as the Department of Justice, which issued a recommended protocol for the use of TAR,[9] provided for in the Manual for Complex Litigation, recognized in the Resource Guide for Managing Complex Litigation[10] and contemplated as a means of review and production in this Court's Mandatory Initial Discovery Pilot which provides:

2. ESI. a. Duty to Confer. When the existence of ESI is disclosed or discovered, the parties must promptly confer and attempt to agree on matters relating to its disclosure and production, including:
i. requirements and limits on the preservation, disclosure, and production of ESI;
ii. appropriate ESI searches, including custodians and search terms, **or other use of technology-assisted review**; and
iii. the form in which the ESI will be produced.[1] (emphasis added)

Accordingly, courts have approved and promoted the use of TAR for many years.

As one magistrate judge explained in 2012:

What the Bar should take away from this Opinion is that computer-assisted review is an available tool and should be seriously considered for use in large-data volume cases where it may save the producing party (or both parties) significant amounts of legal fees in document review.[2]

Today, litigants in complex litigation routinely use TAR for e-discovery.[3] Though no court has yet needed to specifically order a party to use TAR over its formal litigated objection, courts frequently encourage and assist parties in reaching agreement on the use of TAR. "[I]t is widely recognized that 'TAR is cheaper, more efficient and superior to keyword searching.'" *In re Mercedes-Benz Emissions Litig.*, Civ. Ac. No. 2:16-cv-881 (KM)(ESK), Dkt. No. 281 (Jan. 9, 2020) (quoting *Hyles v. New York City*, No. 10-CIV-3119, 2016 WL 4077114 at *2 (S.D.N.Y. Aug. 1, 2016)). TAR is widely recognized as the way to address concerns about the scope and burden of e-discovery, and parties that decline to follow this established approach in major litigation do so at their own peril.

In the *Broilers Chicken* anti-trust litigation, for example, the court confirmed that discovery reaches all non-privileged matters that is relevant to any party's claim or defense and proportional to the needs of the case, even when that entials significant electronic information. At the urging of the court, and with the assitance of a special master, the parties used a TAR protocol to address concers about the cost of discovery, which was approved by the Court.[4]

Similarly, where a party refused to use TAR, Retired District Judge Cavanaugh, as a special master, warned the parties that he would not be receptive to complaints about "burden of discovery requests, specifically cost and proportionality" after the refusal to use TAR.[5]

Validation of Keyword Culling

TAR works best and most efficiently without using keywords to narrow the documents. In fact, a major advantage of TAR is to avoid the laborious and flawed system of using keywords to search documents.[6] A limited keyword search alone may not constitute a reasonable inquiry

under Rule 26(g).[7] Studies have shown that keyword culling may leave behind responsive documents.[8] Reported examples of emails with the following comments have been overlooked:

- "you can't lie if you don't have to talk"
- "We really don't need them getting in our chili about deadlines"
- "it was like dogs watching TV (and me too)"
- "fix one thing, break 3 others"
- "it's such a s**t show"
- "we have wasted so much time and money on this. And it was completely avoidable"
- "I still haven't been forgiven by god for the covering up I did last year"
- "I just received a shovel to starty my journey to the hotter place…"
- "think of the $$$ wasted on those useless bags of #$@% the last 3 years"
- "this airplane is designed by clowns, who in turn are supervised by monkeys"
- "I just jedi mind tricked this fool."
- "I don't know how to refer to the very, very few of us on the program who are interested only in truth But it's mostly depressing that it's so few"

Nor can the use of keywords be justified by any argument about the "richness" of the documents (i.e., that the full document set includes too many irrelevant documents and too few relevant documents for TAR to work). If keywords are used, the process must be validated to avoid concealing masses of relevant documents.

Stopping Criteria

The stopping criteria is simply the metric the parties use to determine when to stop review and check the effectiveness of the TAR process (which is "validation"). After stopping, if the TAR process is valid, then the review is complete. If the TAR process is not sufficient upon validation, the parties must then determine how to address any identified deficiencies.

It makes no sense to stop the review when one or more of every 10 documents (i.e., 10% or more documents) are still relevant, especially if they are meaningful. This is not a disproportionate effort. Studies have

found that when the proportion of relevant documents at the end of the TAR process is 1/6 of the proportion of the documents at the start of the TAR process, it becomes a good indicator that you can stop the TAR review. These are both very intuitive and easy methods to apply, which require no additional effort or burden beyond the regular review process. They are supported by multiple empirical studies involving hundreds of RFPs and a dozen data sets. These methods have been used successfully by Grossman & Cormack on hundreds of TAR reviews since 2014. TAR provider Catalyst (now OpenText) uses a very similar measure for its stopping criteria.

Validation of TAR

Validation is the mechanism used to check the effectivenss of the TAR protocol. Validation is a question of the estimated share of responsive documents identified, and the quantity and quality of the documents likely missed. In *City of Rockford v. Mallinckrodt Ard Inc.*, District Court Judge Johnston analyzed plaintiffs' proposal under Rule 26(g) reasonableness and under Rule 26(b)(1) proportionality standards and decided a a random sample of the null set was both reasonable and proportional, weighing the factors under the proportionality standard.[9]

Summary

Although TAR has been recognized by courts to be a reasonable method for document review, courts have not yet taken the leap forward of ordering a party to use TAR in its document review and production. However, as the AI technology evolves, it is increasingly likely that TAR will become the standard and a very useful tool in the litigator's toolbox. And litigators are well advised to catch up on this technology and consider using TAR in any large document case. ∎

George Bellas served as chair of the ISBA Artificial Intelligence Committee and has been appointed to the Illinois Supreme Court's Judicial Conference Task Force on the use of generative AI in Illinois

courts. George served on the 7ᵗʰ Circuit e-Discovery Committee and has served as a panelist at the Sedona Conference.

1. Gricks, Thomas C. III; Ambrogi, Robert J. (November 17, 2015). *"A brief history of technology assisted review".* Law Technology Today." American Bar Association; Sedona Conference. TAR Case Law Primer Public Comment Version August 2016 Retrieved August 17, 2016, https://thesedonaconference.org/publication/tar-case-law-primer-public-comment-version-august-2016.

2. TAR is also referred to as "predictive coding" or "computer assisted review".

3. https://edrm.net/resources/frameworks-and-standards/technology-assisted-review/.

4. The Sedona Conference, *TAR Case Law Primer, Second Edition*, 24 Sedona Conf. J. 1 (2023).

5. *"Document Categorization in Legal Electronic Discovery: Computer Classification vs. Manual Review,"* Journal of Am. Society for Information Science & Technology, 61(1):70-80 (2010).

6. *"Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review,"* Richmond J. of Law & Tech., Vol. XVII, Issue 3, 1-48 (2011).

7. *Search, Forward: Will manual document review and keyword searches be replaced by computer-assisted coding?,* by Judge Andrew Peck, Law Technology News, October 1, 2011.

8. Maura Grossman & Gordon MacCormack, *Efficient E-Discovery*, ABA Journal (April 2012).

9. See, *https://www.justice.gov/atr/page/file/1096101/download*

10. See "Technology-Assisted Review for Discovery Requests, A Pocket Guide for Judges," Federal Judicial Center, 2017.

11. Amended Standing Order Regarding Mandatory Initial Discovery Pilot Project (2020), https://www.ilnd.uscourts.gov/Pages.aspx?jYyawIFLXKMJrmXzxFk8lw== .

12. *Da Silva Moore v. Publicis Groupe*, 287 F.R.D., 182, 193 (S.D.N.Y. 2012). See also Zach Warren, Judge Peck: There's Still a Misconception That Eyes-On Review is the 'Gold Standard,' LegalTech news (Jan. 6, 2020) (explaining that the judiciary still does not understand "that done right, technology assisted review (TAR), also known as predictive coding, is at least just as good as eyes-on review if not better, and at a much lower cost").

13. *In re Mercedes-Benz Emissions Litig.*, Civ. Ac. No. 2:16-cv-881 (KM)(ESK), Dkt. No. 281 (Jan. 9, 2020); *U.S. ex rel. Proctor v. Safeway, Inc.*, No. 11-CV-3406, 2018 WL 1210965 (N.D.Ill. March 8, 2018); Stipulation and Order Re: Revised Validation and Audit Protocols for the Use of Predictive Coding in Discovery, *Rio Tinto* v. *Vale*, 14 Civ. 3402 (RMB) (AJP), Dkt. No. 338 (Sept. 8, 2015); *See William A. Gross Constr. Assocs.* v. *Amer. Mfrs. Mutual Ins. Co.*, 256 F.R.D. 134 (S.D.N.Y. 2009) (Peck, M.J.); *City of Rockford* v. *Mallinckrodt* ARD, 326 F.R.D. 489 (N.D. Ill. 2018);, *Rio Tinto* v. *Vale*, 14 Civ. 3402 (RMB) (AJP), Dkt. No. 338 (Sept. 8, 2015);

14. *In re Broiler Chicken Antitrust Litigation,* No. 16 C 8637, 2018 WL 3586183 (N.D. Ill. July 26, 2018).

15. *In re Mercedes-Benz Emissions Litig.*, Civ. Ac. No. 2:16-cv-881 (KM)(ESK), Dkt. No. 281 (Jan. 9, 2020) (inviting Plaintiffs to renew their request for TAR if the defendants' non-TAR approach was deficient).

16. *Id, FN 4.*

17. See, *U.S. ex rel. Proctor v. Safeway, Inc.*, No. 11-CV-3406, 2018 WL 1210965 (N.D.Ill. March 8, 2018) (Judge Schanzle-Haskins ordered the defendants to use a TAR exercise to identify relevant documents from its production.)

18. *See, e.g.*, David.C. Blair and M.E. Maron, *An Evaluation of Retrieval Effectiveness for a Full-Text Document Retrieval System*, 28 Commc'ns ACM 289 (1985) (searchers using keywords found only 20% of the relevant documents when they thought they had found at least 75%); Stephen Tomlinson, et al., *Overview of the 2007 TREC Legal Track* (Apr. 30, 2008) (lawyers using keywords found no more than 22% of the relevant documents in a large collection); Douglas W. Oard, et al., *Overview of the TREC 2008 Legal Track* (Mar. 17, 2009) (lawyers using keywords found no more than 24% of the relevant documents in a large collection).

19. *City of Rockford v. Mallinckrodt Ard, Inc.*, 360 F. Supp. 730 (N.D. Ill. 2019).