

---

Portfolio Media, Inc. | 111 West 19th Street, 5th floor | New York, NY 10011 | www.law360.com  
Phone: +1 646 783 7100 | Fax: +1 646 783 7161 | customerservice@law360.com

---

# A Proposed Technology-Assisted Review Framework

By **Christine Payne and Michelle Six** (April 27, 2020, 5:22 PM EDT)

It's time to have a frank conversation about technology-assisted review in the context of civil litigation. For orientation, technology-assisted review is often called TAR, and that's the acronym we will use here.

TAR is a broad term used to describe a variety of computer-assisted methods of document review. Some approaches are algorithmically driven, others have a linguistic or "rules-based" element. TAR serves to amplify human judgment about document categorization within large data sets. In other words, lawyers review a small number of documents, and the technology helps review the rest.

So, what is up for discussion? First, there are inconsistent standards applied to those litigants that choose to use TAR and those that opt for attorney review. [1] Second, the bar should recognize that TAR has become weaponized to the detriment of both litigants and courts. Finally, we propose a technique to reduce cost and contention, addressing legitimate concerns about the accuracy of large-scale responsiveness reviews while also mitigating unhelpful and expensive arguments over process.

Specifically, we propose that parties skip the costly preemptive negotiations about TAR protocols in favor of a report card system that would require post-hoc sampling to establish objective quality control, thus permitting parties and courts to objectively assess whether a given review methodology actually worked. Parties may select and implement any review methodology they like, and at the end of the process they must complete and sign a report card that contains objective metrics demonstrating results, as well as information about sampling methodology.

The signature would function as an attestation and could be made either by the litigant (like an interrogatory verification) or by the attorney under a Rule 26(g) standard under the Federal Rules of Civil Procedure (or state-law equivalent). Certain courts might require both. If the "grades" on the report card are not good enough, the parties and court can craft a remedy.

In evaluating this proposal, it's helpful to understand how views on TAR have evolved, for both practical and commercial purposes.

## Critical and Not-so-Critical Analysis of TAR and Its Benefits

Depending on the data set, TAR can often be faster and cheaper than manual review. We have used TAR in numerous cases, and the evolution of the technology over the past several years is impressive.

For a decade now, there has been a great deal of excitement over the possibilities of TAR and how it might change the way in which civil litigation is handled and resolved. Most litigators have heard the oft-repeated claim that TAR is more accurate than attorney review.

The findings of existing studies, however, are quite measured. For example, in 2010, Herbert



Christine Payne



Michelle Six

Roitblat, Anne Kershaw and Patrick Oot published the results of a head-to-head comparison between TAR and manual coding, "Document Categorization in Legal Electronic Discovery: Computer Classification vs. Manual Review."<sup>[2]</sup> That study concluded that the performance of the TAR systems tested was "at least as accurate (measured against the original review) as that of a human re-review."<sup>[3]</sup>

The next year, Gordon Cormack and Maura Grossman published a law review article, "Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review."<sup>[4]</sup> That study, which is arguably the leading research on the topic, reported significant advantages for TAR in terms of efficiency, and for accuracy reported mixed results. In that review, TAR showed higher rates of precision, but for recall, "the measurements [suggested] that the technology-assisted process may yield better recall, but the statistical evidence is insufficiently strong to support a firm conclusion to this effect."<sup>[5]</sup>

Many in the e-discovery community seized upon these papers, making progressively more enthusiastic claims about the studies' conclusions. In a blog post, one author reported that Cormack and Grossman's study "illustrated that TAR offers results with higher recall and precision" than human review.<sup>[6]</sup> Another said Cormack and Grossman found that TAR is "more effective than human review at finding relevant documents."<sup>[7]</sup>

TAR's superior accuracy became as frequently presumed as Napoleon Bonaparte's short stature. "[T]he accuracy of TAR is also higher than that of human review,"<sup>[8]</sup> read one blog. TAR can "return more accurate and complete results than any human review team," said another.<sup>[9]</sup>

Excited litigants sought to use TAR in real-world applications. Any skepticism about TAR's ability to accurately identify responsive documents was met with the generous offer of unprecedented transparency.

For example, in the seminal case *Da Silva Moore v. Publicis Groupe*, Magistrate Judge Andrew Peck of the U.S. District Court for the Southern District of New York noted that the producing party's "transparency in its proposed search protocol for electronically stored information made it easier for the court to approve the use of predictive coding."<sup>[10]</sup>

Judge Peck encouraged future litigants to be similarly transparent: "such transparency allows the opposing counsel (and the Court) to be more comfortable with computer-assisted review, reducing fears about the so-called 'black box' of the technology. This Court highly recommends that counsel in future cases be willing to at least discuss, if not agree to, such transparency in the computer-assisted review process."

In some cases, such as *Progressive Casualty Insurance Co. v. Delaney*, transparency became a threshold for the ability to use TAR at all.<sup>[11]</sup>

The presumed but unproven superiority of TAR in accurately identifying responsive documents has led us to a strange place. Litigants that decline to use TAR are cast as profligate Luddites, notwithstanding the fact that attorney review — when managed well and strategically — is still sometimes the best choice. Simultaneously, litigants that want to use TAR are subject to attack by opposing parties claiming that the process is unreliable and demanding oversight to ensure adequate results.

More to the point, there are now different standards for "transparency and cooperation" for TAR versus traditional attorney review. Specifically, attorney review offers a simple (though perhaps expensive and time-consuming) procedure: review the documents and produce what is responsive and nonprivileged.

Using TAR, on the other hand, frequently requires an additional level of transparency, resulting in heavily negotiated, fear-based protocols that can be as expensive as they are cumbersome — without any sort of guarantee of the promised increase in accuracy or decrease in costs. Litigants pay the price.

### **Fighting Over TAR Is Expensive and Unhelpful**

Serious concerns exist about TAR's adoption in large-scale litigation. For example, John Rabiej of Duke University posited that some parties are discouraged from using TAR because of confusion over the extent of disclosure that would be required if they did so.

In a recent email proposal to Judge John D. Bates, chairman of the Advisory Committee on Civil Rules, Rabiej suggested additional clarification, stating "[t]he ambiguity between 'aspirational' and 'obligatory' cooperation has resulted in parties avoiding TAR rather than incur the costs of extended negotiations and satellite litigation resolving the ambiguity." [12]

Similarly, Gareth Evans noted in 2015 that the rate of adoption for TAR was slower than initially predicted, in part because of fear related to the amount and nature of "transparency" that would be required, should they elect to use TAR. [13] Judge Peck noted the same in *Rio Tinto PLC v. Vale SA*: "fear of spending more in motion practice than the savings from using TAR for review," he wrote, effectively "discourages parties from using TAR" in the first place. [14] Kate Bauer wrote in 2018 that "inconsistent rulings about the amount of disclosure TAR requires have hobbled its adoption." [15]

They are all correct. If you choose attorney review when responding to document requests, there is very little ability for opposing counsel to monitor or second-guess your process.

This simple privacy stems from the days before e-discovery, where privilege, confidentiality and the doctrine of attorney work product allowed lawyers to work alone in their offices, without opposing counsel demanding to see how documents were being categorized in the conference room. Of course, attorneys had ethical and rules-based obligations to produce responsive documents and, if there was any indication of a problem, there were mechanisms for courts to require additional assurances.

You do not enjoy the same privacy when using TAR for your responsiveness review. The twin requirements of cooperation and transparency become heavy handed quite quickly. Opposing counsel will have at least some ability to know what you are doing, how you are doing it, and to demand information about both inputs and outputs. "Garbage in, garbage out," they say.

There will also almost certainly be disputes over what the TAR protocol should look like in the first place. Lawyers, vendors and consultants could be required to weigh in. Particularly unreasonable litigants might demand to see all documents coded by an attorney as a TAR input, whether responsive or nonresponsive, whether privileged or not, as well as all documents in any quality-control sample. Disputes before the court are common and require expensive motion practice as well as the court's time and energy.

We are not aware of any study that looks at the total time, money and judicial resources spent on preemptive negotiation over review methodology, much less one that demonstrates any sort of a benefit in terms of enhanced review quality. Perhaps there should be one — some type of critical review that looks at (A) what a party would do if left to its own devices, compared to (B) what a party is forced to do as a result of preemptive fighting over TAR protocols, and then makes an evaluation of whether there's any actual difference in the output of either method, especially in light of the time and resources required to go from (A) to (B).

Our hunch — based on representing many different responding parties using many different review methodologies — is that the thousands of dollars and potentially months of time spent on fighting over TAR protocols has very little, if any, benefit when it comes to output quality. From our vantage point, it's no wonder some litigants opt for the traditional route of linear attorney review.

Mr. Rabiej asserts that the big problem with TAR abstention is that it "has thwarted the development of a body of law on the meaning and limits of 'cooperation.'" [16] He correctly notes that the issue of cooperation "is rarely formally presented, and courts are denied the opportunity to develop common-law guidance." [17]

This is because smart litigators want to be as collegial and cooperative as possible and show the court that they are good actors in word and deed. Complaining about your adversary's TAR curiosity comes off as obtuse, especially where there is significant case law that suggests they can and should be involved in the process, at least to a certain extent. Thus, cooperation remains undefined.

No definition of cooperation, however, whether it comes from common law or a top-down rule, will

solve the problem of preemptive fighting over review methodology. What is to be done?

### **The Proof Is in the Pudding**

We need to focus on results. From a broad level, the only thing that really matters in discovery is that a requesting party timely receives documents to which it is entitled. A producing party is required to take reasonable steps to make that happen.

Guardrails do exist in this instance. There are rules and cases that help define the scope of discovery. The rules of professional conduct deter skullduggery. Local rules provide guidance as to an individual judge's preferences and requirements. With that predicate, why should we spend lots of time and money preemptively dictating how a producing party honors its obligations?

If a producing party can sift through documents, identify those that are responsive, and weed out those deemed nonresponsive, does it matter whether they use attorneys, TAR, or a mix of the two? No. All that matters in the end is that the process was reasonably effective.

To that end, we should stop all preemptive fighting over review methodology and instead use a simple, uniform report card that demonstrates the overall quality of the final production. In other words, let's look at the grades, not the study habits. The objective quality-control metrics we propose employing can be derived from sampling the results of large-scale document reviews.

"Recall" is the percentage of responsive documents in a set that are successfully identified as responsive (so if you can find 80 out of 100 responsive documents in a set, your recall score is 80%). "Precision" is the percentage of documents identified as responsive that are actually responsive (for example, if 100 documents in a set are identified as responsive by your process and a quality control check determines that only 80 were truly responsive, then your precision is 80%).<sup>[18]</sup>

To analyze these numbers from a statistical point of view, it's important to understand the size of the set, the sample size, the level of confidence, etc. All of these metrics are objective numbers that can be written down and shared.

This is precisely what we propose be done. Parties can use whatever review methodology they like — attorneys, TAR 1.0, TAR 2.0 (continuous active learning or continuous multimodal learning), linguistically based TAR, whatever they decide is most appropriate for the data set — and at the end of the process, they must supply standardized reporting of metrics on a predetermined form that allows their opponents and the court to independently verify the quality of the review.

The goal would be to provide enough information about sampling methodologies and methods of calculating recall and precision to allow opposing parties and the court to determine whether you have done a reasonably good job and presented correctly calculated metrics. The forms should only be exchanged at the end of the process (not for interim testing) and should be signed using a Rule 26(g) standard (or state-law equivalent). A sample form can be **found here**.

This approach puts all methodologies on the same footing in terms of the level of required disclosure, addressing Judge Peck's concern that "it is inappropriate to hold TAR to a higher standard than keywords or manual review."<sup>[19]</sup> It restores the sanctity of a lawyer's work by preventing opposing counsel from second-guessing or otherwise interfering with work that is still in process. It also cuts down significantly on preemptive fighting over search terms, TAR protocols, sample disclosures, etc.

If there is a problem with the quality of the review (as indicated by the report card or otherwise) then the parties and the court can evaluate potential causes and fixes. Critically, it enables the search process to fully embrace the concept of proportionality regardless of the methodology chosen by the producing party.

Our suggested framework requires that we trust lawyers. If a lawyer conducts a sample and reports the results on a form that she signs, we should presumptively accept those results unless and until we find cause to do otherwise.

And yes, our proposed approach is imperfect. For example, courts might fear hearing lots of whining from litigants with bad grades about the cost and time associated with going back to fix poor-quality

reviews.[20] But complaints like that would be avoided altogether if counsel in charge of the document review and production monitors the process throughout, and adjusts course as necessary if problems arise.

The issue could also be addressed by active judicial management. Courts can give litigants the option early on to: (1) engage in preemptive negotiation and get a preapproved review process that, if followed, will not be revisited later, no matter the quality of the results; or (2) use the report card framework and perhaps be required to redo the process if the results are poor. An either/or option that holds firm would incentivize litigants to bring their A-game the first time.

Finally, this framework will not prevent bad actors from acting badly. Should we just accept that risk and take a chance that lawyers will lie, cheat and steal if not forced to do their work out in the open? Yes. We firmly believe that most lawyers are trustworthy and take their duties regarding ethical representation seriously.

The risk that some lawyers may act unethically is present in every area of the law. But that risk is insufficient to impose significant and largely wasted costs and burdens on parties and judges. Absent a specific showing of past unethical conduct, courts around the country should adopt in the first instance a results-oriented framework for large-scale discovery reviews.

---

*Christine Payne is a partner at Redgrave LLP.*

*Michelle Six is a partner at Kirkland & Ellis LLP and vice chair of the firm's electronic discovery committee.*

*The authors would like to thank Leslie Gutierrez, Dan Nichols, Gareth Evans, Victoria Redgrave, Jonathan Redgrave, Logan Wiggins, Andre Welbon, Leslie Shirk, Nick Snavely, Craig Ball, Brian Shoaf, Jim Mutchnik, Mark Premo-Hopkins, Dan Raffle, Lon Troyer and Jeff Salling for their assistance, especially in the form of zealous debate.*

*The opinions expressed are those of the author(s) and do not necessarily reflect the views of the firm, its clients, or Portfolio Media Inc., or any of its or their respective affiliates. This article is for general information purposes and is not intended to be and should not be taken as legal advice.*

[1] Culling techniques such as search terms are often used to narrow a data set before review. There's a hot debate over whether you can use search terms in addition to TAR. That is the subject for a different article, however, and here we focus on a key issue for clients — how to review the data set once it is finally determined.

[2] Roitblat, Kershaw & Oot, Document Categorization in Legal Electronic Discovery: Computer Classification vs. Manual Review, 61 J. A. Soc'y for Info. Sci. & Tech. 70, 79 (2010).

[3] Roitblat, *supra* at 70.

[4] Grossman & Cormack, Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhausted Manual Review, XVII Rich. J.L. & Tech. 11, \*32 (2011).

[5] *Id.* at 44.


[6] An Unlikely Pair: Lawyers and AI, Epiq Angle (July 31, 2019) <https://www.epiqglobal.com/en-us/thinking/blog/an-unlikely-pair-lawyers-and-ai>.

[7] Thomas C. Gricks III and Robert J. Ambrogi, A Brief History of Technology Assisted Review, Law Technology Today (Nov. 17, 2015) <https://www.lawtechnologytoday.org/2015/11/history-technology-assisted-review/>.

[8] Ariel Darvish, AI and TAR: Law Firms of a New Age, Fordham Journal of Corporate & Financial Law (April 8, 2018) [https://news.law.fordham.edu/jcfl/2018/04/08/ai-and-tar-law-firms-of-a-new-age/#\\_edn14](https://news.law.fordham.edu/jcfl/2018/04/08/ai-and-tar-law-firms-of-a-new-age/#_edn14).


[9] What is Technology-Assisted Review or TAR?, Zapproved (Aug. 15, 2018) <https://zapproved.com/blog/what-is-technology-assisted-review-tar/>.

[10] 287 F.R.D. 182, 192 (S.D.N.Y. 2012).

[11] See, e.g. **Progressive Casualty Insurance Co. v. Delaney** , No. 2:11-CV-00678-LRH-PAL, 2014 WL 3563467 at \*11 (D. Nev. July 18, 2014) (rejecting a producing party's bid to use TAR because that party was "unwilling to engage in the type of cooperation and transparency that its own e-discovery consultant [had] so comprehensibly and persuasively explained is needed for a predictive coding protocol to be accepted by the court or opposing counsel as a reasonable method to search for and produce responsive ESI").

[12] Email from John K. Rabiej, Director, Center for Judicial Studies, Duke University, to John D. Bates, Chairman of the Advisory Committee on Civil Rules, Proposed Rule Amendment, Dec. 19, 2019 ("Rabiej Proposal").

[13] Gareth Evans, Predictive Coding: Can It Get A Break? Excerpted from the Gibson Dunn Mid-Year E-Discovery Update (July 23, 2015) <https://www.linkedin.com/pulse/predictive-coding-can-get-break-gareth-evans/>.

[14] **Rio Tinto PLC v. Vale SA** , No. 14 Civ. 3042 (RMB)(AJP) (S.D.N.Y. March 3, 2015).

[15] Kate Bauer, Technology Assisted Review: Overcoming the Judicial Double-Standard, Richmond Journal of Law and Technology (January 24, 2018) <https://jolt.richmond.edu/2018/01/24/technology-assisted-review-overcoming-the-judicial-double-standard/>.

[16] Rabiej Proposal (see note 10, above).

[17] Id.

[18] There are other metrics that can be derived from these two; "F1," for example, is the harmonic mean of precision and recall.

[19] Rio Tinto at 6.

[20] What counts as good grades and bad grades? The answer, in true lawyerly fashion, is it depends. Parties and the court are in the best position to gauge, based on proportionality factors. Indeed, they are already doing in matters that TAR protocols requiring that objective metrics be exchanged.

**PROTOTYPE FORM\***

<b>FINAL RESULTS REPORT: DOCUMENT REVIEW</b>								
<b>GENERAL INFORMATION</b>								
<b>Case Name:</b>					<b>Production Date:</b>			
<b>Case No. &amp; Jurisdiction:</b>					<b>Producing Party:</b>			
<b>Counsel for Plaintiffs:</b>								
<b>Counsel for Defendant(s):</b>								
<b>Judge:</b>								
<b>SAMPLE INFORMATION</b>								
<b>Name and bar number of attorney(s) conducting sample review:</b>								
	<b>Estimated Score</b>	<b>Confidence Level &amp; Margin of Error</b>	<b>Number of Documents in Sample</b>	<b>Source of Sample</b>	<b>Sampling Method</b>	<b>Documents Coded &amp; Responsive During Sample Review</b>	<b>Documents Coded Non-Responsive During Sample Review</b>	<b>Sample Reviewed at Family Level or Document Level?</b>
<b>Precision</b>								
<b>Recall</b>								
<b>POPULATION INFORMATION</b>								
<b>Number of documents identified as responsive and non-privileged:</b>				<b>Number of documents identified as non-responsive:</b>				

**FINAL RESULTS REPORT: DOCUMENT REVIEW**

**Number of documents identified as uncategorized:**

**Description of collection methodology, parameters, and filters (if any):**

**Number of files collected after de-NIST and de-duplication by file type:**

*\*This is a prototype form. Do you love it? Hate it? Think it only works for binary review protocols? Should it include a specific methodology for calculating recall? Please let us know what you would add, change, or take away. Email: [contactus@redgravellp.com](mailto:contactus@redgravellp.com)*