

Developing the Defensible TAR Approach

From e-books to FaceTime chats, technology has rapidly inserted itself into corners of everyday life where it never existed before, which adds data volumes and creates complexity throughout the discovery process. Data is doubling every two years, estimates technology research firm IDC. Traditional human review of potentially responsive documents has proven unable to keep up with exploding amounts of data in a cost-effective, reliable, and consistent way. This has led to the evolving use of TAR.

For attorneys with years of experience, the slow but inexorable move to TAR has led to many questions and uncertainties. While some attorneys may insist that TAR is faster, cheaper and more reliable than human review, for others it can be difficult to digest and accept, particularly when rules and precedent are still relatively new.

However, as judges begin to clarify and even encourage the use of TAR, clients are beginning to demand it. The ripples from decisions such as *Da Silva Moore v. Publicis Group*,¹ which supported the use of TAR as part of the review process, continue to increase and spread.

Attorneys who have shied away from TAR and other new technologies will have to understand and use them in order to competently represent their clients, as required by the American Bar Association amended Model Rule 1.1, “To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology, engage in continuing study

and education and comply with all continuing legal education requirements to which the lawyer is subject.”²

Proposed amendments to the Federal Rules of Civil Procedure around proportionality, including FRCP 26, could essentially require legal teams to utilize TAR in certain situations. The proposed amendments would build proportionality requirements into Rule 26(b)(1)’s generally applicable limits on the scope of permissible discovery, which must be “proportional to the needs of the case considering

the amount in controversy,

the importance of the issues at stake in the action,

the parties’ resources,

the importance of discovery in resolving the issues, and

whether the burden or expense of the proposed discovery outweighs its likely benefit.”³

¹*Da Silva Moore v. Publicis Group* (S.D.N.Y. Feb. 24, 2012)

²http://www.americanbar.org/content/dam/aba/administrative/ethics_2020/20120808_house_action_compilation_redline_105a-f.authcheckdam.pdf

³Preliminary Draft of Proposed Amendments to the Federal Rules of Bankruptcy and Civil Procedure, page 265. <http://www.uscourts.gov/uscourts/rules/preliminary-draft-proposedamendments.pdf>

In the near future, attorneys will not only need to learn to accept TAR; they should understand both the technology and the processes involved in creating a defensible discovery process using these new technologies. The uses of TAR vary—it is sometimes simply used as part of an early case assessment or to prioritize documents all of which will eventually be reviewed. In fact, many legal teams are using TAR in those contexts, even in smaller and less complex matters. But TAR can also be used to cull down the document set for a portion to be set aside from further human review. In these cases the ability to defend the process is paramount.

UNDERSTANDING TAR

Even as it gains acceptance, the use of TAR seems complicated and uncertain, but a basic understanding is easy to gain. While specific TAR workflows may differ depending on the technology used, generally, senior reviewers identify potentially responsive content (“exemplar responsive documents”) within a subset of documents. Using that information, TAR software then applies those decisions across a much larger data set. For certain use cases such as culling in which defensibility may be challenged, the process may be iterative and undergo several rounds by applying statistical sampling and quality control techniques that improve precision as it progresses.

As part of the process, the system ranks documents and assigns probability scores to indicate how likely it is that the document is responsive. If the legal team is using TAR simply for prioritization, documents with higher priority can be reviewed and analyzed first. If TAR is part of a culling process, the lowest ranking documents, with appropriate sampling, might be set aside for review altogether. This allows for review that is more consistent and far faster than review done solely by human reviewers.

However, highly accelerated review is only a good thing if the initial parameters are identified correctly. The process must be rigorously controlled and audited, and TAR should not be considered “push-button” technology (even if a do-it-yourself software solution is used), where attorneys and their team can sit back and then let the computer do all the work.

KEY ATTRIBUTES OF A DEFENSIBLE TAR

Along with an appropriate understanding of the technology and uses, attorneys must know what is involved in creating a replicable, defensible TAR process. Attributes of a thorough approach should include:

Building a team with TAR expertise

When it comes to TAR, lawyers do not need to become software or statistics experts. In cases where TAR is used for culling prior to production—as opposed to prioritizing for review—legal teams may want to hire those who are well versed in these areas, so lawyers can concentrate on subject matter and case strategy. The experts will also be the ones defending the process if required to attend a Rule 26(f) meet and confer, appear before a judge or submit an affidavit.

Measuring quality of results

To assure ongoing QC control and the final quality assurances, statistically valid measures of precision and recall must be generated at each stage.

Ensuring statistical validity

At the core of successful TAR lies statistical validity. Legal teams must be able to show that they calculated a set of defensible metrics for their particular use case through extensive testing, and that they have measured what they intended to measure in a repeatable way.

Utilizing an adaptive approach

Though TAR is often used to prioritize documents for review or for QC purposes for incoming productions, when it is used for document reduction, or culling, it may be necessary to conduct multiple cycles of review so the software can “learn” more each time. The legal team should analyze results after each round, correcting and refining at every phase.

Combining methodologies

“TAR can be a powerful tool, but it is only one of the tools that counsel should rely on. Other approaches such as keyword searches, conceptual analysis, and random and judgmental sampling should be used in conjunction with TAR to leverage the strengths and offset the weaknesses of many techniques,” said Pam Davis, Director, Legal - Discovery & Litigation Support at Google Inc.

Reproducibility

In order to be defensible, the team needs to be able to replicate the processes and steps they used and achieve the same results with the same raw set of documents. This requires honed and consistent approaches and documentation.

Creating an audit trail

The TAR process must be transparent across the entire corpus, as well as around individual decisions. In order to verify how inputs were chosen, decisions were made and results achieved, the legal team must create a standardized methodology and approach to reporting and documentation that can stand up to the scrutiny of opposing counsel and the courts. (Again, this may only be a concern when the TAR use case necessitates defensibility.)

Cooperating with opposing counsel

As the Federal Rules of Civil Procedure require and as The Sedona Conference has encouraged, parties involved in litigation should cooperate with each other to clarify potential protocols for all aspects of the discovery process. The earlier this conversation is held, the more effective it will be for certain types of use cases, such as culling.

CREATING THE TAR PROCESS

Once attorneys have identified the key attributes of a defensible TAR-based review, the next step is to create the processes around them. While every matter is different and should be customized accordingly, a repeatable, defensible process should be built on a foundation that includes:

1) Defined roles and responsibilities

A team approach will be necessary, and the team should include several key players.

Subject matter experts will be needed at different phases, including the people most knowledgeable about the matter.

Project managers and others must serve as “traffic cops” to ensure that the review process proceeds smoothly and on time, and that nothing falls through the cracks.

Large, complex matters or ones with high exposure may require consultants, including statisticians, linguists and/or software experts who can build the best processes and break down complicated

topics in ways that are easy to understand. These experts can often aid defensibility and also reduce costs. The legal team should consider at the outset whether experts are likely to be required for a particular matter, and identify them early for quick engagement.

2) Managing the project

“The TAR-based review should be built around processes that are applicable to any major project. However, with TAR it is particularly important to identify the process and outline the project before starting. Everyone involved must be clear on the strategy and goals from the beginning,” said Julie Brown, Litigation Technology Manager of Vorys, Sater, Seymour and Pease LLP.

The project manager must ensure that all standards are identified and metrics are assigned from the outset. PMs can also fill the role of facilitating conversations to determine statistical goals and update those as necessary based on continuous measurement throughout the TAR cycles.

Even if the technology is new, the old adage of “garbage in, garbage out” still applies. The team must be able to demonstrate that the right documents were included in the review, the results that were produced were validated and the entire process is documented.

3) Creating inputs and measuring outputs

A TAR review begins with a small sample where expert human reviewers who are most knowledgeable about the case assess every document for responsiveness. Often, multiple people will review the same documents. Then, QC is conducted and the exemplar responsive documents are used to train the TAR software to apply similar decisions to categorize the entire data set.

Once they are processed, documents are assigned a score based on the software’s determination of the likelihood that they are potentially responsive. Documents above a certain range might be automatically designated for human review. Documents that score below a certain percentage are generally deemed non-responsive and, depending on the case strategy, may not be reviewed by humans, except for sampling to test and make sure that responsive documents are not slipping through.

Scores in the mid-range often represent documents that do not provide enough information to determine if it is likely responsive or not. The best approach is to review a sample of these documents again in order to create a supplemental training set, and then run these documents through the system again. Another option is to immediately route those documents for manual review.

It is important to understand that not all documents are appropriate to run through TAR, such as image documents with no text and poorly OCR scanned (paper) documents. These document types must be identified and will require a different workflow.

4) Continuous feedback and refinement

Approaches and changes should be constantly fine-tuned, and each new batch of results should be carefully evaluated. The team may decide to widen or narrow the range of documents that must be input into the system again. The key is to stop to consider and document reasons behind the decisions made.

Decisions that should be constantly considered include:

- Composition of the seed set;
- Adjusting the sample size in order to achieve the desired level of statistical certainty; and
- Knowing when to stop. This threshold becomes easier to recognize if the team already set goals from the outset. Carefully monitoring precision and recall will help to indicate when it's time to finish with TAR.

This stage represents a critical aspect of developing a TAR approach that will withstand questions from the courts and opposing counsel. At some point, the team must consider efficiency as well. The results should be compared back to the statistical model every time. A technology-based approach requires starting with expected results based on a statistical model and then evaluating actual outcomes every step of the way.

5) QC strategies

When considering QC strategies, it may be helpful to think of checking the quality of a batch of light bulbs. Before starting the manufacturing process, engineers have already determined a “defect rate:” the expected number of bulbs that are defective. Once manufacturing is complete, a sample of the

bulbs is tested to make sure that the number of defects is near—or below—expectation. If it is not, the entire batch is called into question. That same mathematical principle underpins the final acceptance sampling in TAR to test for recall.

6) Reporting and Documentation

In order to prove defensibility, decisions must be documented from beginning to end. A good TAR system will include reporting functionalities to help track progress and allow the team to make sense of how far it has come and where it needs to catch up. While the type of reporting can be tool- and vendor-specific, whatever tool is used should be able to document all decisions and at what point decisions were made.

The key to a defensible TAR process is not to jump in, and then make decisions as information emerges. Standards and processes must be developed up front, rather than on the fly. A key part involves keeping track of statistical thresholds, which documents were used for training and what the overturn rates were. Most importantly, attorneys must be able to demonstrate from the beginning what the expected outcomes would be and what those outcomes ended up being. As with any other type of discovery process, this should be based on multiple factors, including risk profile and other considerations.

Interested in learning more?

Contact us at info@elitediscovery.com



400 N. Saint Paul, Suite 1300, Dallas, TX 75201
866.896.2626 | www.elitediscovery.com