

# The COMPUTER & INTERNET *Lawyer*

Volume 35 ▲ Number 2 ▲ FEBRUARY 2018

Ronald L. Johnston, Arnold & Porter, LLP, Editor-in-Chief

## Court Guideposts for the Path to Technology Assisted Review Adoption

By James A. Sherer, David Choi, and Csilla Boga-Lofaro

Technology Assisted Review (TAR), also known as Computer Assisted Review (CAR), uses software to analyze, search, and categorize documents that are relevant for the purposes of discovery in those cases or investigations (matters) that require the review and production of materials in electronic or “E-discovery” form. TAR and CAR often are used interchangeably with the term “Predictive Coding”—though, to be more precise, TAR and CAR encompass many applications, of which the most commonly discussed is predictive coding. TAR was introduced primarily as a means of reducing costs associated with discovery efforts, and predictive coding is a type of TAR in which an algorithm leverages decisions made by an attorney reviewer regarding

a subset of documents to make predictions regarding the relevancy of each document in the data set.<sup>1</sup> This article employs the generally preferred term “TAR” to refer to the application of predictive coding to prepare document productions in the discovery phase of matters.<sup>2</sup>

In 2012, US Magistrate Judge Andrew J. Peck issued the first judicial opinion endorsing the use of TAR in *Da Silva Moore, et al. v. Publicis Groupe*.<sup>3</sup> Three years later, Judge Peck noted in *Rio Tinto PLC v. Vale S.A.*<sup>4</sup> that case law had developed to a point where the choice to use TAR had become “black letter law.” Despite that pronouncement (as well as those other, similar court opinions Judge Peck referenced), we assert that most judges, litigators, and clients are still on the journey to fully embracing TAR as their preferred discovery process—and have decidedly not yet reached the destination. In support of that journey, and in aid to those practitioners who are considering the use of TAR and are seeking court guidance on an approach, this article details what courts have held regarding TAR as a holistic approach as well as TAR-process specifics addressed by the courts.

While the oft-cited *Sedona Principles* posit that parties responding to discovery requests are best

---

**James A. Sherer** is a partner and **David Choi** and **Csilla Boga-Lofaro** are associates at Baker & Hostetler LLP in New York, NY. This article was submitted on October 2, 2017 as part of the proceedings of the 2017 Georgetown Advanced eDiscovery Institute. The views expressed herein are solely those of the authors, should not be attributed to their places of employment, colleagues, or clients, and do not constitute solicitation or the provision of legal advice.

situated to evaluate procedures, methodologies, and technologies appropriate for preserving and producing electronically stored information (ESI),<sup>5</sup> judges have still taken opportunities to either opine on or address a number of specific considerations parties may have when applying TAR in response to requests. We have categorized those procedures and considerations involved in the TAR process and addressed by the courts into the following:

- Proportionality and Party Objections;
- Cooperation, Transparency, and the Provision of Seed Sets;
- Recall and Precision; and
- Audit Practices or Validation.

These considerations do not provide a fool-proof method to deploy TAR that all parties and courts would accept or agree to. But the examination of these categories should help practitioners think critically about proposed TAR implementations, and will aid in the provision of discovery generally.

## Proportionality and Party Objections

Considerations of proportionality follow from the benefit-versus-burden test provided in Federal Rule of Civil Procedure (FRCP) 26(b)(1), where “[a]ll discovery is subject to the limitations imposed by Rule 26(b)(2)(C).” Specifically, “the Rule 26 proportionality test allows [a c]ourt to ‘limit discovery if it determines that the burden of the discovery outweighs its benefit.’”<sup>6</sup> Because TAR functions to leverage computers in lieu of manual review and attorney billable-hour time and lowers the cost of review accordingly, proportionality considerations generally support the application of TAR in “large-data” cases. In addition to being less expensive, courts considering TAR also have agreed that its results can be more effective than manual review as well. In *Moore v. Publicis Groupe*,<sup>7</sup> the court opined specifically on the favorability of TAR and the “myth” that some lawyers cling to in believing manual review to be the “gold standard.” Studies debunk that myth and “clearly show that computerized searches are at least as accurate, if not more so, than manual review.”<sup>8</sup>

In *Moore*, the court permitted defendant to use TAR over plaintiffs’ objection because of the vast amount of ESI to be reviewed (over three million documents), the superiority of computer-assisted review to the available alternatives (*i.e.*, linear manual review or

keyword searches), and the need for cost effectiveness and proportionality under Rule 26(b)(2)(C). The court noted that no lawyer “using any search method could honestly certify that its production is ‘complete,’”<sup>9</sup> but that TAR was a proven option that “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.”<sup>10</sup>

---

**Because TAR functions to leverage computers in lieu of manual review and attorney billable-hour time and lowers the cost of review accordingly, proportionality considerations generally support the application of TAR in “large-data” cases.**

---

In addition to considering TAR an appropriate measure to factor into proportionality considerations, courts have consistently approved the use of TAR over opposing party’s objections. In *Bridgestone Americas, Inc. v. International Business Machines Corp.*,<sup>11</sup> the court approved plaintiff’s request to use predictive coding to review over two million documents, despite defendant’s objections that (a) switching to TAR was an unwarranted change in the original case management order; and (b) that it would be unfair to use predictive coding after completing an initial screening with search terms. The court opined that ultimately, “the use of predictive coding is a judgment call, hopefully keeping in mind the exhortation of Rule 26 that discovery be tailored by the court to be as efficient and cost-effective as possible.”<sup>12</sup> This case involved millions of documents to be reviewed with costs likewise in the millions. In allowing plaintiff to switch to TAR, the court noted that discovery, especially when dealing with a large volume of ESI, should be a flexible and transparent process, and that principles of efficiency and proportionality may justify “switch[ing] horses in midstream.”<sup>13</sup>

While courts are willing to approve the use of TAR over an opposing party’s objection, courts are nevertheless reluctant to force a party to use TAR against its choice, even if proportionality considerations would support the application of TAR. For example, in *Hyles v. New York City*,<sup>14</sup> while Judge Peck noted that, “in general, TAR is cheaper, more efficient and superior to keyword searching,” he ultimately declined to compel defendants to use TAR, finding that that TAR has not

yet displaced other tools to the point where alternate approaches are unreasonable. The court nevertheless felt compelled to note that “for most cases today, TAR is the best and most efficient search tool.”<sup>15</sup> Similarly, the court in *In re Viagra (Sildenafil Citrate) Products Liability Litig.*,<sup>16</sup> found *Hyles* “well-reasoned.” The *Viagra* court acknowledged that “it may well be correct that production using keywords may not be as complete as it would be if TAR were used. But the standard is not perfection, or using the ‘best’ tool, but whether the search results are reasonable and proportional.”<sup>17</sup> Ultimately, there was no basis upon which the court could compel defendant to use TAR, despite proportionality considerations warranting it.<sup>18</sup>

Another case that involved a party “switch[ing] horses in midstream” from using keyword searches to TAR highlights the importance of proportionality considerations in the application of TAR, and illustrates that such considerations will not always err in favor of TAR under all circumstances involving large amounts of data. The court in *In re Biomet M2a Magnum Hip Implant Products Liability Litig.*<sup>19</sup> held that defendant’s initial use of “keyword culling” against the population of ESI before switching to the use of TAR was reasonable under proportionality principles of FRCP 26(b). The court did not compel defendant to restart from the beginning and use TAR on the entire ESI population, despite plaintiffs’ arguments to the contrary.

The court acknowledged that “[i]t might well be that predictive coding, instead of a keyword search, at [an earlier stage] of the process would unearth additional relevant documents.” But the court also acknowledged it would cost defendant “a million, or millions, of dollars to test” the plaintiff’s theory “that predictive coding would produce a significantly greater number of relevant documents.” Instead, proportionality considerations ruled the day. The court concluded that the potential expense of restarting the process to apply TAR was too great “[e]ven in light of the needs of the hundreds of plaintiffs in this case, the very large amount in controversy, the parties’ resources, the importance of the issues at stake, and the importance of this discovery in resolving the issues.”<sup>20</sup>

### **Cooperation, Transparency, and the Provision of Seed Sets**

Cooperation and transparency—principles that are important in conducting discovery in every litigation—are all the more important when TAR is involved. The primary objectives when applying TAR are efficiency and cost-savings. But getting mired in disputes and motion practice over the application

and execution of TAR can negate those benefits. It is easy for disputes to arise in this context because, as courts have acknowledged, “the [TAR] process is complicated.”<sup>21</sup>

---

**The primary objectives when applying TAR are efficiency and cost-savings. But getting mired in disputes and motion practice over the application and execution of TAR can negate those benefits.**

---

US Magistrate Judge Peggy A. Leen reminded the bar of how important cooperation and transparency are when using TAR. In one case over which Judge Leen presided, a party sought leave to deviate from the agreed upon ESI protocol, and asked the court for permission to use TAR rather than keyword searches. Judge Leen acknowledged that TAR may be an improvement over traditional methods of review and production “as a far more accurate means of producing responsive ESI in discovery.” But even where the court noted the potential supremacy of TAR, it could not approve the plaintiff’s request to use it when the parties already had demonstrated a lack of willingness to engage in the type of cooperation and transparency that is needed to establish an acceptable TAR protocol. Specifically, the court stated that those “cases which have approved technology assisted review of ESI have required an unprecedented degree of transparency and cooperation among counsel in the review and production of ESI responsive to discovery requests;” but found that discovery in the present case had been “contentious” and had forced the court to already resolve “many, many discovery disputes.” Thus, even where the court realized the applicability (and even stated a preference for) TAR, the court recognized that the parties—and *not* the court—would need to undertake cooperative measures to properly implement TAR. The court simply did not trust that TAR could be efficiently applied under such contentious circumstances and ruled accordingly.

One TAR issue that remains open is how transparent and cooperative the parties need to be with respect to the “seed” sets (also known as training sets).<sup>22</sup> Judge Peck explored this issue—without having decided on it—in *Rio Tinto*,<sup>23</sup> and noted a then-current split in decisions among courts.

The least contentious scenarios are where parties agree to work together in establishing the seed set, or otherwise agree to share the seed set with each other.

For example, the defendant in *Da Silva Moore* volunteered to turn over to plaintiffs “[all non-privileged] documents that are reviewed as a function of the seed set, whether . . . ultimately coded relevant or irrelevant.”<sup>24</sup> The *Da Silva Moore* court then “highly recommend[ed] that counsel in future cases be willing to at least discuss, if not agree to, such transparency in the computer-assisted review process.”<sup>25</sup> In *In re Actos (Pioglitazone) Products Liability Litig.*,<sup>26</sup> the parties’ protocol had “experts” from each side simultaneously review and code the seed set. In *Bridgestone*,<sup>27</sup> the court made clear that “full openness in the matter” was “of critical importance,” where plaintiff agreed to provide the responsive and non-responsive seed set documents to defendant.

The parties in *West Penn Allegheny Health Systems, Inc. v. UPMC*<sup>28</sup> agreed to a protocol in the event of a dispute as to whether a document was appropriately part of a seed set. Upon good cause, a party would be required to “identify within their responsive documents those documents that were part of the initial seed set.” The parties further agreed that the requesting party could contribute to the producing party its own coded documents as a “supplemental seed set used to provide additional training to the [producing party’s] system.”<sup>29</sup> More recently, *In re Bair Hugger Forced Air Warming Products Liability Litig.*<sup>30</sup> involved parties that each agreed to identify up to 250 documents to be included in their own initial seed sets; each party also was required to review and code for inclusion in their seed set “no less than 1,500 randomly selected documents.”

In his *Rio Tinto* decision, Judge Peck noted that the only case in which a court mandated a party share its seed set was in *Federal Housing Financial Agency v. HSBC North American Holdings Inc.*<sup>31</sup> In a bench decision on July 24, 2012, Judge Cote required that defendant give plaintiff full access to the seed set’s non-privileged responsive and non-responsive documents. Nevertheless, when forced to make a decision on this issue, most published opinions show that courts tend to be reluctant to order parties to share their seed sets. The *Biomet*<sup>32</sup> court conceded that there was no authority to provide the basis for requiring defendant to share seed set documents with plaintiffs’ counsel, but encouraged defendant to reconsider its “troubling position,” and chastised defendant’s “unexplained lack of cooperation.” More recently, the court in *Aurora Coop. Elevator Co. v. Aventine Renewable Energy-Aurora W, LLC*<sup>33</sup> took a similar view, and conceded the rules of discovery do not technically require the disclosure of seed sets because Rule 26(b)(1) does not authorize ordering a party to disclose irrelevant information (which a seed set should include). While the *Aurora* court pointed out that that position is supported “by the language, if not the spirit, of the civil

discovery rules,” the judge “encourage[d] defendants to reconsider their position and work cooperatively [so as to] allay the risk of having to repeat the process if the plaintiff later challenges, and the courts agrees, that the defendant’s unilaterally created computer review training was faulty or unreliable.”<sup>34</sup>

---

**Case law generally indicates that courts encourage, but are reluctant to mandate, the sharing of seed sets; and some practitioners have gone so far as to argue that aspects of the seed set selecting process could constitute work product.**

---

Case law therefore generally indicates that courts encourage, but are reluctant to mandate, the sharing of seed sets; and some practitioners have gone so far as to argue that aspects of the seed set selecting process could constitute work product.<sup>35</sup> Generally, cooperation would make the issue significantly easier on the parties, but if agreement cannot be reached regarding the provision of seed sets, then—as noted by Judge Peck in the *Rio Tinto* decision—there are other means by which parties may insure that TAR was appropriately implemented, “such as statistical estimation of recall at the conclusion of the review as well as by whether there are gaps in the production, and quality control review of samples from the documents categorized as non-responsive.”<sup>36</sup>

## Recall and Precision

Recall and Precision, essential concepts to gauge the TAR process, are a measure of completeness and selectivity correlated in such a way that when recall increases, precision often decreases. Ultimately, the recall rate and precision rate measure the effectiveness of the information retrieval provided by the operation of the TAR process, and help the producing and receiving parties understand more about the documents produced and the universe of materials from which the documents were produced.

The recall rate is the percentage of the relevant documents in the collection that are found by the review. In other words, the higher the recall, the fewer relevant documents that are left behind, and a recall rate of 100 percent would mean that a reviewer retrieved every relevant document from a collection (a recall rate of 80 percent would mean that 20 percent of the relevant documents were left behind and

not included in the production). The precision rate is the percentage of documents selected by the TAR process that actually are relevant, with the balance of the documents being irrelevant. In other words, the higher the precision, the fewer “false positives” there are. Therefore, a precision rate of 70 percent would mean that 30 percent of the documents pulled would be irrelevant.

When parties agree to the use of TAR, court decisions and supporting literature indicate that the parties also should consider the targets for recall and precision in tandem, understanding their interrelation: as the target rate for one increases, a higher target for the other becomes more difficult to achieve. This tradeoff is logical—a broad search that misses few relevant documents usually will capture a lot of irrelevant documents, while a narrower search that minimizes “false positives” will be more likely to miss some relevant documents.

In seeking leave of court to use TAR, the defendant in *Global Aerospace Inc. v. Landow Aviation, L.P.*,<sup>37</sup> argued that among the benefits of TAR were improvements in both greater recall and precision rates when compared to those seen when human review is used. In particular, the defendant cited studies that, when taken together, determined that linear review misses, on average, 40 percent of the relevant documents (*i.e.*, 60 percent recall rate), and the documents pulled by human reviewers are nearly 70 percent irrelevant (*i.e.*, 30 percent precision rate).

Based on those studies, the producing party proposed, and the court endorsed, an acceptable recall rate of 75 percent. This meant that “predictive coding will conclude once the sampling program establishes that at least 75 percent of the relevant documents have been retrieved from defendant’s ESI available to [defendant] for discovery purposes (*e.g.*, privilege review, responsiveness review, etc.)” The parties did not specify a target precision rate.

The parties in *Bair Hugger*<sup>38</sup> adopted a protocol which stated that “adequate training of the CAR system shall be reached when both Recall rate is at or exceeds 80 percent.” The parties further agreed that if a “producing party is unable to meet the 80 percent threshold for Recall with reasonable effort, the Parties shall meet and confer on other steps that can reasonably be taken to meet that threshold and/or modification of the threshold if necessary. If the Parties cannot agree on other steps and/or modification of the threshold, the Parties shall bring the issue to the court for resolution.”<sup>39</sup>

Setting a high recall rate, even one close to 100 percent, does not necessarily mean that a court will

demand perfection. The parties in *Dynamo Holdings Ltd. Partnership v. Comm’r of Internal Revenue*<sup>40</sup> agreed to set the recall rate at 95 percent, which the court noted meant the model chosen “would return more non-responsive documents.” Indeed, this ultimately resulted in a precision rate of 3 percent. But despite setting such a high recall rate, respondents argued that the recall rate was in fact too low (or petitioner’s TAR method was otherwise flawed) based on the absence in petitioner’s production of some relevant documents found through Boolean searches. In response, the court rejected respondent’s assertions regarding the TAR process and noted that FRCP 26(g) requires only a “reasonable inquiry,” not perfection.<sup>41</sup> The court also noted that petitioner was transparent in the discovery process and had worked cooperatively with respondent, and therefore fulfilled its obligations.<sup>42</sup>

### Audit Practices

When parties agree to the use of TAR, they traditionally agree on an audit process or method to evaluate results to assess if TAR achieved the parties’ goals. The audit process typically involves the random selection of documents from both relevant (responsive) and irrelevant (not responsive) document sets that are then reviewed for accuracy. There is no single audit process endorsed by the courts, and parties in different cases have therefore agreed to different audit processes and protocols.

---

**When parties agree to the use of TAR, they traditionally agree on an audit process or method to evaluate results to assess if TAR achieved the parties’ goals.**

---

The defendant in *Landow Aviation*<sup>43</sup> proposed (a) taking random samples from both the relevant and irrelevant document sets; (b) reviewing both sample sets to examine the relevance of each document to determine the number of relevant documents in both the relevant and irrelevant sets; and (c) after logging and removing the privileged and irrelevant sensitive documents, making them available to opposing counsel. The size of the sample sets would be determined at the conclusion of the TAR process once the distribution of relevant documents became apparent. The defendant also noted an “escape valve” of sorts for the process that would involve a court determination, including language in the proposal ultimately accepted by the court that, “if

counsel disagree[d] with the position taken by Landow, the court may determine the relevance of any document or the propriety of withholding any particular document.”<sup>44</sup>

The parties in *Rio Tinto* proposed a validation protocol that would take place cooperatively between the parties *before* the producing party produced its documents.<sup>45</sup> After an initial exchange of documents, parties agreed to a post-production audit protocol under which each party could request the preparation of a 2,400 document sample from the producing party’s production; the requesting party also would have the right to specify how many documents from a given source, an “Audit Category,”<sup>46</sup> should be in the audit sample. The producing party would then have that sample coded by human reviewers who had no prior knowledge of the Audit Categories; how the documents were previously coded by TAR; or whether they were previously produced. The producing party would share the non-privileged documents from the audit sample and the human coding results. Upon the requesting party’s own review of the audit sample, the requesting party could raise any disputes regarding how the documents in the audit sample were coded.

The parties in *Bair Hugger*<sup>47</sup> took the process one step further, agreeing to provide an extremely detailed documentation for each “run” of the predictive coding process. As an audit protocol, the parties essentially agreed to total transparency and detailed documentation of and throughout the TAR process. Each producing party was to prepare an Excel spreadsheet with columns detailing:

- The TAR run date & time;
- A list of control numbers for all documents the system identified as responsive, along with a copy of those documents except that no copy need be provided for those documents withheld for privilege;
- A list of control numbers for all documents the system identified as non-responsive, along with a copy of those documents the system incorrectly identified as non-responsive and which were moved from the non-responsive set to the responsive set, except that no copy need be provided for those documents withheld for privilege;
- A count of the number of privileged documents correctly identified as responsive but withheld for privilege;
- A count of the number of privileged documents incorrectly identified as responsive but withheld for privilege;
- The metadata for all documents... depending on the privilege status of the document; and
- Four additional columns in the spreadsheet indicating:
  - Whether the document was marked by the CAR system as responsive or non-responsive;
  - Whether, after manual review, the document was changed from the non-responsive set to the responsive or vice-versa;
  - If a copy of the document is not produced by reason of privilege, the reason for withholding the document on the basis of privilege; and
  - For any document for which a copy is not being produced due to non-responsiveness or privilege, information sufficient to allow the receiving party to determine why the document was considered non-responsive and/or privileged;<sup>48</sup>

## Conclusion

By and large, courts are supportive of parties using or even considering the use of TAR in litigation, and have recognized the benefits TAR can provide when facing proportionality considerations under Rule 26(b)(1). But courts generally have been reluctant to mandate TAR’s use, and have stayed away from directing the technical aspects of using TAR (such as whether parties must provide seed sets; the parties’ audit practices; and establishing target recall vs. precision rates). The parties have instead been directed or at least entrusted to come up with an agreed upon protocol, and the cases discussed above exhibit just some of the varying ways parties have approached each issue.

While reluctant, courts have sometimes been forced to consider the appropriateness of TAR as an option in certain cases given proportionality considerations. When faced with this difficult situation, courts have still shied away from compelling parties to use TAR—even in cases that involve large amounts of data—and instead have restated the generally accepted principle that each party is best situated to determine how to prepare its own production. Understandably, courts have been even more explicit in stating that they “[do not]

endorse any vendor<sup>49</sup> or specific technologies for TAR. As Judge Peck noted in *Hyles*,<sup>50</sup> TAR has not yet displaced other tools to the point where using something else is unreasonable.

---

**For those parties considering the implementation of TAR as a discovery mechanism, the published case law does provide guidance for a sound protocol and approach, and indicates how courts will consider the specifics of a party's ultimate determinations and resulting productions.**

---

Given the ever-increasing acceptance and popularity of TAR, we may one day reach the point where TAR is the default form of review and a required step of the process of production. Over time, courts may be more inclined to establish standards with respect to the technical factors involved in preparing a sufficient production using TAR. In the meantime, for those parties considering the implementation of TAR as a discovery mechanism, the published case law does provide guidance for a sound protocol and approach, and indicates how courts will consider the specifics of a party's ultimate determinations and resulting productions.

## Notes

1. See Maura R. Grossman & Gordon v. Cormack, "The Grossman-Cormack Glossary of Technology-Assisted Review," 7 *Fed. Cts. L. Rev.* 1 (2013).
2. See *Rio Tinto PLC v. Vale S.A.*, 306 F.R.D. 125, 126 (S.D.N.Y. 2015) ("technology assisted review, or TAR, now seems to be the preferred term of art").
3. *Da Silva Moore, et al. v. Publicis Groupe*, No. 11-Civ.-1279 (ALC) (AJP), 2012 WL 607412 (S.D.N.Y. Feb. 24, 2012).
4. *Rio Tinto PLC v. Vale S.A.*, 306 F.R.D. 125, 127 (S.D.N.Y. 2015) ("In the three years since *Da Silva Moore*, the case law has developed to the point that it is now black letter law that where the producing party wants to utilize TAR for document review, courts will permit it.>").
5. "The Sedona Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production," *The Sedona Conference* (2017 Public Comment Version), <https://thesedonaconference.org/publication/The%20Sedona%20Principles>. ("Because the responding party generally is best situated to evaluate, select, and implement the procedures, methodologies, and technologies appropriate to meet its preservation and discovery obligations, there should be no preemptive restraint placed on a responding party that chooses to proceed on its own with determining how best to fulfill its preservation and discovery obligations. Thus, as a general matter, neither a requesting party nor the court should prescribe or detail the steps that a responding party must take to meet its discovery obligations, and there should be no discovery on discovery, absent an agreement between the parties, or specific, tangible, evidence-based indicia (versus general allegations of deficiencies or mere 'speculation') of a material failure by the responding party to meet its obligations.").
6. *Kleen Products LLC v. Packaging Corp. of Am.*, No. 10-Civ.-5711, 2012 WL 4498465, at \*9 (N.D. Ill. Sept. 28, 2012)  *citing In re IKB Deutsche Industriebank AG*, No. 09-Civ.-7582, 2010 WL 1526070, at \*5 (N.D. Ill. Apr. 8, 2010).
7. *Moore v. Publicis Groupe*, 287 F.R.D. 182, 190 (S.D.N.Y. 2012), adopted sub nom. *Moore v. Publicis Groupe SA*, No. 11-Civ.-1279 (ALC) (AJP), 2012 WL 1446534 (S.D.N.Y. Apr. 26, 2012).
8. *Id.*
9. *Id.* at 188.
10. *Id.* at 193.
11. *Bridgestone Americas, Inc. v. Int. Bus. Machs. Corp.*, No. 13-Civ.-1196, 2014 WL 4923014 (M.D. Tenn. July 22, 2014).
12. *Id.* at \*1.
13. *Id.*
14. *Hyles v. New York City*, No. 10-Civ.-3119 (AJP), 2016 WL 4077114 (S.D.N.Y. Aug. 1, 2016).
15. *Id.* at \*3.
16. *In re Viagra (Sildenafil Citrate) Prods. Liab. Litig.*, No. 16-MD-02691 (RS) (SK), 2016 WL 7336411, at \*2 (N.D. Cal. Oct. 14, 2016).
17. *Id.*
18. *Id.* See also *Burnett v. Ford Motor Co.*, No. 13-Civ.-14207, 2015 WL 4137847, at \*11 (S.D.W. Va. July 8, 2015) (ordering parties "to consider" the use of TAR, but not outright ordering the use of TAR).
19. *In re Biomet M2a Magnum Hip Implant Prods. Liab. Litig.*, No. 12-MD-2391, 2013 WL 1729682, at \*3 (N.D. Ind. Apr. 18, 2013).
20. *Id.* at \*3.
21. *Rabin v. PricewaterhouseCoopers LLP*, No. 16-Civ.-02276, 2017 US Dist. LEXIS 125404 (N.D. Cal. Aug. 08, 2017) (despite acknowledging that TAR is a "complicated process," the efficiencies of TAR justified adopting an "aggressive" discovery timeline).
22. A seed set is a sample of documents from the universe of documents that is reviewed by individuals knowledgeable about the subject matter. The determinations made on the seed set comprise the primary reference data to teach the predictive coding machine how to recognize patterns of relevance in the larger document set.
23. *Rio Tinto*, 306 F.R.D. at 128–29.

24. *Da Silva Moore*, 287 F.R.D. 182 at 187.
25. *Id.* at 187.
26. *In re Actos (Pioglitazone) Prods. Liab. Litig.*, No. 11-MD-2299, 2012 WL 7861249, at \*4 (W.D. La. July 27, 2012).
27. *Bridgestone*, 2014 WL 4923014, at \*1.
28. *West Penn Allegheny Health Sys., Inc. v. UPMC*, No. 09-Civ.-0480 (JFC), 2013 WL 12134102, at \*3 (W.D. Pa. Feb. 28, 2013).
29. *Id.*
30. *In re Bair Hugger Forced Air Warming Prods. Liab. Litig.*, No. 15-MDL-2666 (JNE) (FLN), 2016 WL 3702959, at \*2 (D. Minn. July 8, 2016).
31. *Fed. Housing Fin. Agency v. HSBC N. Am. Holdings Inc.*, Case Nos. 11 Civ. 6189-6190, 6193, 6195, 6198, 6200-6203, 6739, & 7010 (S.D.N.Y. July 24, 2012) (DLC).
32. *Biomet*, 2013 WL 6405156, at \*2.
33. *Aurora Coop. Elevator Co. v. Aventine Renewable Energy-Aurora W., LLC*, No. 12-Civ.-230, 2015 WL 10550240, at \*2 (D. Neb. Jan. 6, 2015).
34. *Id.*
35. *See, e.g.*, John M. Facciola & Philip J. Favro, "Safeguarding the Seed Set: Why Seed Set Documents May Be Entitled to Work Product Protection," 8 *Fed. Cts. L. Rev.* 1 (2015).
36. *Rio Tinto*, 306 F.R.D. at 129 (*citing* Maura R. Grossman & Gordon v. Cormack, Comments On "The Implications of Rule 26(g) on the Use of Technology-Assisted Review," 7 *Fed. Cts. L. Rev.* 285, 301-12 (2014)).
37. Defendants' Memorandum in Support of Motion for Protective Order Approving the Use of Predictive Coding, No. CL-61040, 2012 WL 1419842 (Va. Cir. Ct. Apr. 9, 2012). The court granted defendants' motion in *Global Aerospace Inc. v. Landow Aviation, L.P.*, 2012 WL 1431215 (Va. Cir. Ct. Apr. 23, 2012).
38. *Bair Hugger*, 2016 WL 3702959, at \*2.
39. *Id.*
40. *Dynamo Holdings Ltd. P'ship v. Comm'r of Internal Rev.*, No. 2685-11, 2016 WL 4204067, at \*4 (T.C. July 13, 2016).
41. *Id.*
42. *Id.* at \*4 (Petitioner "provided the [respondent] with seed sets of documents... and [respondent] determined which documents were relevant. That selection was used to develop the predictive coding algorithm.")
43. *Landow Aviation*, 2012 WL 1419842.
44. *Id.*
45. Stip. and Order re: Revised Validation and Audit Protocols for the Use of Predictive Coding in Discovery, *Rio Tinto PLC v. Vale S.A.*, No. 14-Civ.-3042 (RMB) (AJP) (S.D.N.Y. Sept. 8, 2015), ECF No. 338.
46. *E.g.*, documents that were not keyword hits, documents that were identified as likely non-responsive by the TAR process, documents that were marked as "non-responsive" by manual reviewers, etc.
47. *Bair Hugger*, 2016 WL 3702959, at \*2.
48. *Id.*
49. *Da Silva Moore*, 287 F.R.D. at 193. In fact, the *Da Silva Moore* court made it a point to be "very careful not to mention the names of the parties' vendors in the body of th[e] Opinion." *Id.*
50. *Hyles*, 2016 WL 4077114, at \*3.