The Sedona Conference Journal

Volume 18 2017

The Sedona Conference TAR Case Law Primer

The Sedona Conference

Recommended Citation:

The Sedona Conference TAR Case Law Primer, 18 Sedona Conf. J. 1 (2017).

For this and additional publications see: https://thesedonaconference.org/publications

The Sedona Conference Journal® (ISSN 1530-4981) is published on an annual or semi-annual basis, containing selections from the preceding year's conferences and Working Group efforts. The Journal is available on a complimentary basis to courthouses and public law libraries and by annual subscription to others ($95; $45 for conference participants and Working Group members). Send us an email (info@sedonaconference.org) or call (1-602-258-4910) to order or for further information. Check our website for further information about our conferences, Working Groups, and publications: www.thesedonaconference.org.

Comments (strongly encouraged) and requests to reproduce all or portions of this issue should be directed to:
The Sedona Conference,
301 East Bethany Home Road, Suite C-297, Phoenix, AZ 85012 or
info@sedonaconference.org or call 1-602-258-4910.

The Sedona Conference Journal® designed by MargoBDesignLLC at www.margobdesign.com or mbraman@sedona.net.

Cite items in this volume to "18 Sedona Conf. J. _____ (2017)."

THE SEDONA CONFERENCE TAR CASE LAW PRIMER*

A Project of The Sedona Conference Working Group on Electronic Document Retention and Production (WG1)

Author:	The Sedona Conference
Editor-in-Chief:	Maura R. Grossman
Team Leaders:	Lea Malani Bays
	Sandra Rampersaud
Senior Contributing Editor:	Gareth Evans
Drafting Team Members:	Abigail Dodd
	Maureen O'Neill
	J. Michael Showalter
Steering Committee Liaisons:	Maura R. Grossman
	Joseph P. Guglielmo
	John J. Rosenthal
Judicial Observer:	Hon. Andrew J. Peck

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group 1. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.

We thank all of our Working Group Series Sustaining and Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our

sponsors, just click on the "Sponsors" navigation bar on the homepage of our website.

PREFACE

Welcome to the Final Version of The Sedona Conference *TAR Case Law Primer*, a project of The Sedona Conference Working Group on Electronic Document Retention and Production (WG1). The Sedona Conference is a 501(c)(3) research and educational institute that exists to allow leading jurists, lawyers, experts, academics, and others, at the cutting edge of issues in the areas of antitrust law, complex litigation, and intellectual property rights, to come together in conferences and mini-think tanks called Working Groups to engage in true dialogue—not debate—in an effort to move the law forward in a reasoned and just way.

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. However, its widespread application—and the realization of its potential benefits—has been impeded by uncertainty about its acceptance by the courts as a legitimate alternative to costly, time-consuming manual review of documents in discovery. This *Primer* analyzes decisions from those courts that have been required to opine on the efficacy of TAR in a variety of circumstances and explores the evolution in the courts' thinking from 2012 through the end of 2016.

The *Primer* is the product of more than a year of development and dialogue within WG1. It was originally conceived as a chapter of a larger Commentary on the use of TAR in civil litigation, but the rapid development of the case law, the volume of court decisions, and the importance of those decisions in shaping legal practice in real time required that an exposition of the case law be made available on a faster timetable than WG1's usual dialogue and consensus-building process allowed. For

that reason, the *Primer* strives to present the case law in as neutral a fashion as possible. It avoids making recommendations regarding particular TAR methodologies, nor does it propose principles, guidelines, or best practices for TAR application, independent of those suggested by the courts themselves.

As the title suggests, the *Primer* is a starting point. The evolution in the case law is far from complete, nor is the analysis. The Sedona Conference hopes that the *Primer*, as all of the output of its Working Groups, will evolve into an authoritative statement of the law. We welcome your input on the *Primer* as we continue to receive new decisions that present novel facts, issues, and arguments. Your comments and suggestions may be sent to comments@sedonaconference.org.

I want to thank all the drafting team members for their dedication and contribution to this project, including team leaders Lea Malani Bays and Sandra Rampersaud; senior contributing editor Gareth Evans; drafting team members Abigail Dodd, Maureen O'Neill, and J. Michael Showalter; and WG1 Steering Committee Liaisons Joseph R. Guglielmo and John J. Rosenthal. Special thanks go to Hon. Andrew J. Peck, who as Judicial Observer contributed his all-important view from the bench; and to Editor-in-Chief and Steering Committee Liaison Maura R. Grossman, without whose determination, hard work, and willingness to devote countless hours, this publication would not have been possible.

Kenneth J. Withers
Deputy Executive Director
The Sedona Conference
January 2017

TABLE OF CONTENTS

I. INTRODUCTION

The jurisprudence regarding technology-assisted review (TAR)[1] is not yet well developed, and the case law reflects a number of inconsistencies and unresolved issues. This *Primer* represents our best efforts to synthesize and summarize the current state of the law (and the open questions), in a neutral fashion, as of the end of 2016. It does not reflect an exhaustive compendium of all TAR issues that may have come before the courts, nor does it cover TAR protocols that parties have negotiated and the courts have so ordered.

As discussed in Section II, below, *Da Silva Moore v. Publicis Groupe*, decided in 2012, was the first published opinion recognizing TAR as an "acceptable way to search for relevant ESI in appropriate cases."[2] Since then, as discussed in Section III, a number of courts have encouraged the use of TAR, or commented on its availability to reduce cost and burden. Some parties have stipulated to the use of TAR without disputes requiring court intervention. And some requesting parties have used TAR to review large volumes of documents produced by responding parties (or third parties). As discussed in Section IV,

---

1. Technology-Assisted Review, or TAR, is a "process for prioritizing or coding a collection of Electronically Stored Information using a computerized system that harnesses human judgments of subject matter expert(s) on a smaller set of documents and then extrapolates those judgments to the remaining documents in the collection." The Sedona Conference, *The Sedona Conference Glossary: E-Discovery and Digital Information Management*, Fourth Edition, 15 SEDONA CONF. J. 305 (2014) (definition adopted from Maura R. Grossman & Gordon V. Cormack, *The Grossman-Cormack Glossary of Technology-Assisted Review with Foreword by John M. Facciola, U.S. Magistrate Judge*, 7 FED. CTS. L. REV. 1, 32 (2013)). The terms "predictive coding" and "computer-assisted review" are often used interchangeably with TAR, to describe this process. This *Primer* will use the term "TAR," unless quoting a case that uses another term.

2. Da Silva Moore v. Publicis Groupe, 287 F.R.D. 182, 183 (S.D.N.Y. 2012).

several cases reflect the parties' use of TAR without otherwise addressing its use.

As discussed in Section V, a number of decisions have addressed substantive disputes regarding the use of TAR. These issues include, among others, whether the use of TAR can be compelled by motion (Section V.A.); whether a responding party can switch to TAR after commencing search and review with another methodology (Section V.B.); whether TAR may be preceded by keyword or other culling methods (Section V.C.); whether a party using TAR can be required to share with opposing counsel coding decisions rendered on the seed, training, or validation sets (including providing access to irrelevant documents in those sets) (Section V.D.); and whether court approval is necessary before using TAR (Section V.E.). Many, if not all, of these issues remain unresolved.

As discussed in Section V.F., courts have addressed a variety of other issues, such as recall thresholds (Section V.F.1.); post-production challenges to the use of TAR (Section V.F.2.); retraining the TAR tool for subsequent document requests (Section V.F.3.); and manual review following TAR (Section V.F.4.). As discussed in Section V.F.5., some government agencies have accepted the use of TAR as a search methodology for the production of documents in response to regulatory investigations. The Federal Trade Commission, for example, issued an update in August 2015 to its Model Second Request for merger antitrust investigations, which now asks parties using TAR to provide certain information at the end of the process.[3] Similarly, counsel

3. Fed. Trade Comm'n, *Model Request for Additional Information and Documentary Material (Second Request),* at 15–16 (revised Aug. 2015), https://www.ftc.gov/system/files/attachments/merger-review/guide3.pdf.

for the Antitrust Division of the Department of Justice has provided guidance regarding TAR protocols that should be negotiated at the outset in response to Division investigations.[4]

As discussed in Section VI, courts in Ireland, England, and Australia have approved the use of TAR.

Finally, as discussed in Section VII, there has been an evolution in thinking about TAR in the years since *Da Silva Moore*. There appears to be growing comfort within the legal community with the reliability of TAR, as reflected in *Rio Tinto PLC v. Vale S.A.*, decided in early 2015.[5] In *Rio Tinto*, which carries the subtitle "*Da Silva Moore* Revisited," and which was decided by the same judge as *Da Silva Moore*, the court wrote that TAR can no longer be considered an "unproven technology," and that, "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."[6] Moreover, TAR technologies are evolving in ways that may impact concerns about the composition of seed or training sets. For example, the court wrote in *Rio Tinto* that with TAR tools using continuous active learning, seed sets may have relatively little impact on results and, as a practical matter, there may be no discrete training sets to share.[7]

---

4. U.S. Dep't of Justice, *Request for Additional Information and Documentary Material (Model Second Request*), at 13 (June 2015), https://www.justice.gov/atr/request-additional-information-and-documentary-material-issued-weebyewe-corporation.

5. Rio Tinto PLC v. Vale S.A., 306 F.R.D. 125 (S.D.N.Y. 2015).

6. *Id.* at 127.

7. *Id.* at 128 (citing Gordon V. Cormack & Maura R. Grossman, *Evaluation of Machine Learning Protocols for Technology-Assisted Review in Electronic Discovery*, in Proceedings of the 37th Int'l ACM SIGIR Conf. on Research & Dev. in Info. Retrieval (SIGIR '14), at 153–62 (ACM New York, N.Y. 2014), http://dx.doi.org/10.1145/2600428.2609601; 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, 298 (2014) ("Disclosure of the seed or training set offers false comfort to the requesting party . . . .")).

II. THE BEGINNING: *DA SILVA MOORE*

As noted above, *Da Silva Moore v. Publicis Groupe*, decided in 2012, reflects the first published opinion recognizing TAR as an "acceptable way to search for relevant ESI in appropriate cases."[8] Before *Da Silva Moore* was decided, TAR had been available for some time, but it was not being widely used in practice. The court observed that many attorneys knowledgeable about TAR and its potential benefits were reluctant to use it because no court had yet approved its use. "While anecdotally it appears that some lawyers are using predictive coding technology, it also appears that many lawyers (and their clients) are waiting for a judicial decision approving of computer-assisted review."[9]

The court in *Da Silva Moore* approved a party-negotiated TAR protocol, which had set forth the manner of selection and review of the seed and training sets, and addressed those aspects of the protocol about which the parties disagreed.[10] According to the court, its approval of TAR meant that "[c]ounsel no longer have to worry about being the 'first' or 'guinea pig' for judicial acceptance of [TAR]."[11] The court added that, "[w]hat the Bar should take away from this Opinion is that [TAR] 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."[12] The court stated, however, "[t]hat does not mean computer-assisted review must be used in all cases, or

8. 287 F.R.D. 182, 183 (S.D.N.Y. 2012).

9. *Id.* at 182–83 (quoting Andrew Peck, *Search, Forward*, L. TECH. NEWS, Oct. 11, 2011, at 25).

10. *See id*. at 182–83, 190–93.

11. *See id*. at 193.

12. *Id*.

that the exact ESI protocol approved here will be appropriate in all future cases that utilize computer-assisted review."[13]

A. Advantages of TAR

The court described a number of the advantages of TAR over linear manual (i.e., human) review. It observed that exhaustive manual review is "simply too expensive," where millions of documents are involved, and cited a study demonstrating substantial savings for TAR—on average, a fifty-fold savings in the number of documents requiring review.[14] Additionally, the court stated that, "while some lawyers still consider manual review the 'gold standard,' that is a myth," and cited studies showing that TAR "'can (and does) yield more accurate results than exhaustive manual review, with much lower effort.'"[15]

B. Emphasis on Process

The court in *Da Silva Moore* suggested that "the best approach" when a party wishes to use TAR is to "follow the Sedona Cooperation Proclamation model" and "[a]dvise opposing counsel that you plan to use [TAR] and seek agreement."[16] If the parties are unable to reach agreement, then the court

---

13. *See id*.

14. *Id*. at 190 (citing Maura R. Grossman & Gordon V. Cormack, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, 17 RICH. J.L. & TECH. 43 (2011)).

15. *See id*. at 190 (quoting Maura R. Grossman & Gordon V. Cormack, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, 17 RICH. J.L. & TECH. 43, 48 (2011)); *see also id*. (citing Herbert L. Roitblat, Anne Kershaw & Patrick Oot, *Document Categorization in Legal Electronic Discovery: Computer Classification v. Manual Review*, 61 J. AM. SOC'Y FOR INFO. SCI. & TECH. 70, 79 (2010)).

16. *Id*. at 184 (quoting Andrew Peck, *Search, Forward*, L. TECH. NEWS, Oct. 11, 2011, at 29).

stated that parties should "consider whether to either abandon [TAR] for that case or go to the court for advance approval."[17]

With respect to court approval, the court stated that it "recognizes that [TAR] is not a magic, Staples-Easy-Button, solution appropriate for all cases."[18] While the technology should be used where appropriate, courts should consider the particular protocol that is proposed. "[I]t is not a case of machine replacing humans: it is the process used and the interaction of man and machine that the courts need to examine."[19] The court emphasized that in doing so, perfection is not required of TAR. "While this Court recognizes that [TAR] is not perfect, the Federal Rules of Civil Procedure do not require perfection."[20]

C. The Dispute and the Court's Decision

Although the parties in *Da Silva Moore* agreed in principle to the defendant's use of TAR, they disagreed about aspects of the protocol that the defendant would follow—in particular, whether training would consist solely of seven "iterative rounds," and whether the quality-control process would be adequate. Plaintiffs expressed concerns about whether the protocol would work.[21] The parties agreed to some aspects of the protocol, including the composition of the seed set and that the defendant would share the training and quality-control sets (except for privileged documents).[22]

The court observed that plaintiffs' concerns about the reliability of the TAR process were premature until the process was

17. *Id*.

18. *Id.* at 189.

19. *Id*.

20. *Id.* at 192.

21. *Id.* at 187–88.

22. *Id.* at 191–92 (citing FED. R. CIV. P. 1 & 26(b)(2)(C)).

underway or complete. It accepted defendant's proposal for seven iterative rounds of training, with the caveat that additional rounds might be required if the parties did not agree that the predictive model was "stabilized" after seven rounds.[23]

The court concluded that defendant's use of TAR was appropriate, considering the following factors: (1) the parties' agreement to use TAR (even though they disagreed on certain aspects of its implementation), (2) "the vast amount of ESI to be reviewed (over three million documents)," (3) "the superiority of [TAR] to the available alternatives (i.e., linear manual review or keyword searches)," (4) "the need for cost effectiveness and proportionality" under Federal Rule of Civil Procedure 26(b)(2)(C), and (5) "the transparent process proposed by [defendant]."[24]

The court added that defendant's "transparency in its proposed ESI search protocol made it easier for the Court to approve the use of [TAR]" because "such transparency allows the opposing counsel (and the Court) to be more comfortable with [TAR], reducing fears about the so-called 'black box' of the technology," and addressing concerns about "garbage in, garbage out" in training the tool. While the court encouraged parties to provide such transparency in future cases, it also indicated that it is not necessarily required for the use of TAR.[25]

23. *Id.* at 187.

24. *Id.* at 191–92.

25. *Id.* at 192.

III. SINCE *DA SILVA MOORE*, MANY OTHER COURTS HAVE ENCOURAGED THE USE OF TAR

After *Da Silva Moore* recognized TAR as an acceptable search methodology, many other courts have encouraged its use, or commented on its availability to potentially reduce cost and burden. However, most of these cases have not involved substantive discussions or approval of its use in the particular case.

For example, shortly after *Da Silva Moore* was decided, the court in *National Day Laborer Organizing Network v. U.S. Immigration & Customs Enforcement Agency*[26] wrote:

> [P]arties can (and frequently should) rely on latent semantic indexing, statistical probability models, and machine learning tools to find responsive documents. Through iterative learning, these methods (known as 'computer-assisted' or 'predictive' coding) allow humans to teach computers what documents are and are not responsive to a particular FOIA or discovery request and they can significantly increase the effectiveness and efficiency of searches.[27]

Similarly, in *In re Domestic Drywall Antitrust Litigation*,[28] the court referred to the availability of TAR for searching large volumes of documents produced by the opposing party. And in *Malone v. Kantner Ingredients, Inc.*,[29] the court noted that, "[p]redictive coding is now promoted (and gaining acceptance) as not

26. 877 F. Supp. 2d 87, 109 (S.D.N.Y. 2012).

27. *Id*.

28. 300 F.R.D. 228, 233 (E.D. Pa. 2014).

29. Case No. 4:12-CV-3190, 2015 WL 1470334, at *3 n.7 (D. Neb. Mar. 31, 2015).

only a more efficient and cost effective method of ESI review, but a more accurate one."

The courts in *Harris v. Subcontracting Concepts, LLC*,[30] and *Chevron Corporation v. Donziger*[31] commented on TAR as a means to reduce cost and burden. In *Harris*, the court rejected a burden argument on the grounds that "[w]ith the advent of software, predictive coding, spreadsheets and similar advances, the time and cost to produce large reams of documents can be dramatically reduced."[32] Similarly, in *Chevron*, the court pointed to the availability of TAR in rejecting a burden argument, observing that "predictive coding is an automated method that credible sources say has been demonstrated to result in more accurate searches at a fraction of the cost of human reviewers."[33]

Courts have encouraged the parties to consider the use of TAR in a number of other cases. In *FDIC v. Bowden*,[34] the court ordered the parties to "consider the use of predictive coding." In *Deutsche Bank National Trust Co. v. Decision One Mortgage Co., LLC*,[35] the court stated that, "if the parties agree that predictive coding would be appropriate in this case, they are encouraged to use that tool."

Some courts have gone beyond encouragement and have ordered parties to consider using TAR. In *Aurora Cooperative Elevator Co. v. Aventine Renewable Energy*,[36] the court ordered the parties to "consult with a computer forensic expert to create

30. Case No. 1:12-MC-82, 2013 WL 951336, at *5 (S.D.N.Y. Mar. 11, 2013).

31. Case No. 11-Civ.-0691, 2013 WL 1087236, at *32 n.255 (S.D.N.Y. Mar. 15, 2013).

32. *Harris*, 2013 WL 951336, at *5.

33. *Chevron*, 2013 WL 1087236, at *32 n.255.

34. No. 4:13-cv-245, 2014 WL 2548137, at *13 (S.D. Ga. June 6, 2014).

35. No. 13 L 5823, 2014 WL 764707, at *1 (Ill. Cir. Ct. Jan. 28, 2014).

36. No. 4:12-civ-230, slip op. at 1–2 (D. Neb. Mar. 10, 2014).

search protocols, including predictive coding as needed, for a computerized review of the parties' electronic records." Similarly, in *Johnson v. Ford Motor Co.*,[37] the court ordered the parties to "involve their IT experts and to consider other methods of searching such as predictive coding."[38]

37. No. 3:13-cv-06529, 2015 WL 4137707 (S.D. W. Va. July 8, 2015).

38. *Id.* at *11. *See also* Section V.A.2., *infra*.

IV. ADDITIONAL CASES REFLECTING THE PARTIES' USE OF TAR

Several cases reflect the parties' use of TAR, without otherwise addressing its use. Some cases have reflected that counsel for plaintiffs have used TAR in analyzing and reviewing documents they had received in document productions from defendants or third parties. In *New Mexico State Investment Council v. Bland*,[39] for example, the court, in approving settlements, noted that, "[i]n reviewing documents, [plaintiff's counsel] implemented various advanced machine learning tools such as predictive coding, concept grouping, near-duplication detection and e-mail threading."[40] The court further stated that, "[t]hese tools . . . enabled the reviewers on the document analysis teams to work more efficiently with the documents and identify potentially relevant information with greater accuracy than the standard linear review."[41] Additionally, in approving a settlement and an award of attorney's fees in *Arnett v. Bank of America*,[42] the court noted that plaintiff's counsel reviewed the more than 1.1 million documents produced in the case using "search terms, predictive coding, and manual review methods."[43]

In *Gabriel Technologies Corporation v. Qualcomm Inc.*,[44] the court awarded more than $2.8 million in fees incurred for the use of "computer assisted, algorithm-driven document review" for almost 12 million documents. The court awarded the defendant attorney's fees and TAR-related costs under federal patent law and for misappropriation claims under California's Uniform Trade Secrets Act based on its finding that the plaintiff

39. No. D-101-cv-2011-01434, 2014 WL 772860 (D.N.M. Feb. 12, 2014).

40. *Id.* at *6.

41. *Id*.

42. No. 3:11-cv-1372, 2014 WL 4672458 (D. Or. Sept. 18, 2014).

43. *Id.* at *9.

44. Case No. 09-cv-1992, 2013 WL 410103, at *10 (S.D. Cal. Feb. 1, 2013).

acted in bad faith by bringing "objectively baseless claims." The court further found that the defendant's use of TAR was "reasonable under the circumstances" of the case.[45]

45. *Id*.

V. DISPUTED ISSUES REGARDING TAR

A number of decisions have addressed various disputed issues regarding the use of TAR. Many or all of these issues remain open, either because of a lack of consensus among the decisions, an absence of in-depth analysis in the decisions, the fact-specific nature of certain decisions, or the paucity of decisions addressing an issue.

A. Requiring the Use of TAR

Several cases have involved attempts to require a responding party to use TAR, either at the behest of the requesting party or at the behest of the court.

1. Motion by The Requesting Party

In *Kleen Products LLC v. Packaging Corporation of America*,[46] a consolidated antitrust action alleging that defendants conspired to fix prices in the containerboard industry, plaintiffs sought to require defendants to use "content-based advanced analytics"—a form of TAR—rather than (according to plaintiffs) the "antiquated Boolean search of [defendants'] self-selected custodians' ESI and certain central files." Defendants already had used a keyword-based search to produce documents, at a cost of more than $1 million.[47] Defendants objected to plaintiffs' pro-

46. Case No. 10-cv-5711, 2012 WL 4498465 (N.D. Ill. Sept. 28, 2012).

47. Pls.' Statement of Position with Respect to Disputed Items for Dec. 15, 2011 Status Conference at 4–5 & n.6, Kleen Prods. LLC v. Packaging Corp. of Am., Case No. 1:10-cv-05711 (N.D. Ill. Dec. 13, 2011).

posal, arguing that it would require them to "jettison their previous work product and adopt [a] new, untested document gathering and production protocol."[48]

The dispute in *Kleen* led to two days of evidentiary hearings, during which plaintiffs' consultants testified regarding the efficacy of their proposed TAR protocol, and defendants' consultants testified regarding the discovery protocol already in place, including the development, testing, revision, and validation of defendants' search terms.[49] The court ultimately declined to require defendants to adopt one technology over another; instead the court ordered the parties to meet and confer regarding modifications to the existing search methodology.[50] That defendants had already substantially completed their review and plaintiffs were seeking to have them start over using a TAR methodology likely factored significantly in this outcome. The court also cited Principle 6 of *The Sedona Principles*, which states, "[r]esponding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information."[51]

The parties ultimately reached a stipulation by which plaintiffs withdrew their demand that defendants apply TAR for the first corpus of documents, but reserved the right to raise objec-

48. *See* Defs.' Statement of Position with Respect to Disputed Items for Dec. 15, 2011 Status Conference at 4–16, *Kleen Prods.*, Case No. 1:10-cv-05711 (N.D. Ill. Dec. 13, 2011).

49. *See* Evidentiary Hr'g Tr., *Kleen Prods.*, Case No. 1:10-cv-05711 (Feb. 21, 2012); Evidentiary Hr'g Tr., *Kleen Prods.*, Case No. 1:10-cv-05711 (Mar. 28, 2012).

50. Evidentiary Hr'g Tr. at 297–300, *Kleen Prods.*, Case No. 1:10-cv-05711 (Mar. 28, 2012).

51. *Id.* at 297–98.

tions to defendants' search methodology—including the completeness of defendants' productions—and to propose alternative methodologies for subsequent requests for production.[52]

Similar to *Kleen*, in *In re Bridgepoint Education, Inc. Securities Litigation*,[53] the court denied plaintiffs' request to require the defendants to use TAR on custodians' documents that defendants had previously searched using traditional search terms.[54]

In *Hyles v. New York City*,[55] the court concluded that defendant New York City could not be compelled to use TAR against its will, even though it agreed with the plaintiff that, "in general, TAR is cheaper, more efficient and superior to keyword searching."[56] In contrast with prior cases, where the producing party had already expended significant effort and expense on document review and production,[57] in *Hyles* the producing party had not yet initiated its review, thus raising the issue of whether, on the requesting party's motion at the outset of discovery, a court can order a responding party to use TAR. The court declared

52. Stipulation & Order Relating to ESI Search, *Kleen Prods.*, Case No. 1:10-cv-05711 (Aug. 21, 2012).

53. No. 12-cv-1737, 2014 WL 3867495 (S.D. Cal. Aug. 6, 2014).

54. *Id.* at *4. Based on a review of the cases to date, the court in *Rio Tinto* observed, in dicta, that "where the requesting party has sought to force the producing party to use TAR, the courts have refused." Rio Tinto PLC v. Vale S.A., 306 F.R.D. 125, 127 n.1 (S.D.N.Y. 2015).

55. 10 Civ. 3119, 2016 WL 4077114 (S.D.N.Y. Aug. 1, 2016).

56. *Id*. at *2.

57. The court stated that in prior cases "where the requesting party has sought to force the producing party to use TAR, the courts have refused." *Id*. The court noted, however, that in those cases, the responding party had already "spent over $1 million using keyword search (in *Kleen Products*) or keyword culling followed by TAR (in *Biomet*)." *Id.*

that "[t]he short answer is a decisive 'NO.'"[58] The court suggested that there "may come a time when TAR is so widely used that it might be unreasonable for a party to decline to use TAR," but "[w]e are not there yet."[59]

As in *Kleen Products*, the *Hyles* court reasoned that, "[u]nder Sedona Principle 6, the City as the responding party is best situated to decide how to search for and produce ESI responsive to Hyles' document requests."[60] Although the City might have to redo its search if the plaintiff later demonstrates deficiencies in the City's production, the court nevertheless reasoned "that is not a basis for Court intervention at this stage of the case."[61] The court concluded that "it is not up to the Court, or the requesting party (Hyles), to force the City as the responding party to use TAR when it prefers to use keyword searching. While Hyles may well be correct that production using keywords may not be as complete as it would if TAR were used, the standard is not perfection, or using the 'best' tool, but whether the search results are reasonable and proportional."[62]

Similarly, in *In re Viagra Products Liability Litigation,*[63] the court denied the requesting party's motion to require that the responding party use TAR, and to allow the requesting party's representatives to be involved in the process. The responding

58. *Id.* at *1 (emphasis in original).

59. *Id*. at *3.

60. *Id.*

61. *Id*.

62. *Id.* (internal citations omitted).

63. *In re* Viagra (Sildenafil Citrate) Prods. Liab. Litig., Case No. 16-md-02691-RS (SK), slip. op. at 1–3 (N.D. Cal. Oct. 14, 2016).

party instead planned to employ an iterative search-term process, which it would test and validate through sampling.[64] Relying upon the reasoning of *Hyles*, the court held that it was not up to the court or the requesting party to force the responding party to use TAR when it preferred to use search terms.[65] The court concluded that "[e]ven if predictive coding were a more efficient and better method, which [the responding party] disputes, it is not clear on what basis the Court could compel [the responding party] to use a particular [search method], especially in the absence of any evidence that [the responding party's] preferred method would produce, or has produced, insufficient discovery responses."[66] The court therefore denied the motion, without prejudice to revisiting the issue if the requesting party later contended that the production was deficient.[67]

2. Suggested by the Court

In two cases, a court proposed the use of TAR, which was ultimately adopted by one or more of the parties.

In *EORHB, Inc. v. HOA Holdings LLC*,[68] Vice Chancellor Laster of the Delaware Chancery Court *sua sponte* ordered the parties to use TAR or, alternatively, to show cause why TAR should not be used. The defendant ultimately elected to use TAR. The plaintiff, however, was not required to do so after informing the court that, because of the low volume of documents

64. *Id.*, slip op. at 1.

65. *Id.*, slip op. at 2.

66. *Id.*, slip op. at 2–3.

67. *Id.*, slip op. at 3.

68. Civil Action No. 7409-VCL (Del. Ch. Oct. 15, 2012) (Hr'g Tr. at 66–67).

it expected to review and produce, the cost of using TAR likely would outweigh any practical benefits.[69]

In *Independent Living Center v. City of Los Angeles*,[70] the court ordered (on consent) the use of TAR to search more than two million documents after "little or no discovery was completed" before the discovery cutoff, and the parties had ongoing disputes after "months of haggling" over search terms that yielded large numbers of documents for review.[71]

B. "Switching Horses Midstream": Contradictory Decisions

Two cases—*Progressive Casualty Insurance Company v. Delaney*[72] and *Bridgestone Americas, Inc. v. International Business Machines Corp.*[73]—have reached differing conclusions on whether a responding party may switch to TAR in the middle of discovery after having previously agreed to use search terms and manual review. The differing outcomes appear to result from the unique facts of each case.

In *Progressive*, the court denied the plaintiff's request to use TAR. Factors the court cited included: the plaintiff sought to use TAR extremely late in the discovery period; it had not yet produced a single document; it had previously agreed in the parties' ESI protocol to use search terms and manual review; it was not willing to reveal its coding decisions and irrelevant documents in the seed and training sets; and it made the decision to switch to TAR unilaterally, without informing defendants or the

69. *See* EORHB, Inc. v. HOA Holdings LLC, 2013 WL 1960621 (Del. Ch. May 6, 2013).

70. No. 2:12-cv-00551, slip op. (C.D. Cal. June 26, 2014).

71. *Id.*, slip op. at 1–2.

72. Case No. 2:11-cv-00678, 2014 WL 3563467 (D. Nev. July 18, 2014).

73. Case No. 3:13-1196, 2014 WL 4923014 (M.D. Tenn. July 22, 2014).

court.[74] According to the court, the parties had "spent months narrowing search terms," at the plaintiff's insistence, to reduce its burden.[75] The narrowed search terms that the parties agreed on yielded 565,000 "hit" documents out of a total population of 1.8 million. Although the plaintiff had initially represented that it would begin production in September 2013 and complete it by the end of October 2013, it advised the requesting party on December 20, 2013, that the process of reviewing the documents retrieved by the search terms was unworkable.[76]

As an alternative to manual review, the plaintiff proposed to apply TAR to the 565,000 documents that "hit" on the search terms, and estimated that plaintiff's TAR process would result in a recall of 70–80% (i.e., that it would find 70–80% of the total number of relevant documents in the collection). Plaintiff would then manually review the documents identified by TAR for production.[77]

The court in *Progressive* rejected plaintiff's proposal, on the grounds that it had previously agreed to manually review the search-term hits and it was too late to change course. The court indicated, however, that it likely would have approved the use of TAR had it been proposed earlier in the case. "Had the parties worked with their e-discovery consultants and agreed at the onset of this case to a predictive coding-based ESI protocol, the court would not hesitate to approve a transparent, mutually agreed upon ESI protocol. However, this is not what happened."[78]

---

74. *Progressive*, 2014 WL 3563467, at *8.

75. *Id.* at *5.

76. *Id.* at *4, *5.

77. *See id*.

78. *Id.* at *9.

In *Bridgestone*, however, the court permitted the plaintiff to change its search and review methodology to TAR mid-stream, based on plaintiff's determination that it would be a much more efficient process, despite defendant's objections that the request was an "unwarranted change in the original case management order," and that it would be unfair to allow the use of TAR "after an initial screening has been done with search terms."[79] "In the final analysis," the court stated, "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." The court added that, "[i]n this case, we are talking about millions of documents to be reviewed with costs likewise in the millions. There is no single, simple, correct solution possible under these circumstances."[80]

The court in *Bridgestone* also wrote that "[t]he Magistrate Judge believes that he is, to some extent, allowing Plaintiff to switch horses in midstream. Consequently, openness and transparency in what Plaintiff is doing will be of critical importance." The plaintiff advised the court that it had agreed to "provide [to defendant] the seed documents they are initially using to set up predictive coding."[81]

C. Using Search-Term Culling Before TAR

Several cases have addressed the appropriateness of using search terms to cull the document population before applying TAR.

---

79. *See* Bridgestone Ams., Inc. v. Int'l Bus. Machines Corp., Case No. 3:13-1196, 2014 WL 4923014, at *1 (M.D. Tenn. July 22, 2014).

80. *Id*.

81. *Id*.

In *In re Biomet M2A Magnum Hip Implant Products Liability Litigation*,[82] the court denied plaintiffs' motion to require the defendant to redo their search and review process using TAR on the entire document population that it had collected. The defendant had used keywords to cull the collected document set from 19.5 million documents and attachments down to 3.9 million. After having further de-duplicated the documents, it used TAR on this smaller data set, identifying almost 2 million documents for production.

Plaintiffs argued that keyword search is less accurate than TAR and that defendant's efforts were tainted by using keyword search before TAR. The court rejected plaintiffs' arguments on the basis of proportionality, holding that the defendant's methodology satisfied the standard set forth in Federal Rules 26 and 34, namely, that its efforts must be "reasonable."

The court in *Biomet* reasoned as follows:

> It might well be that predictive coding, instead of a keyword search . . . would unearth additional relevant documents. But it would cost Biomet a million, or millions, of dollars to test the [plaintiffs'] theory that predictive coding would produce a significantly greater number of relevant documents. Even 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, I can't find that the likely benefits of the discovery proposed by [plaintiffs] equals or outweighs its

82. Case No. 3:12-MD-2391, 2013 WL 1729682 & 2013 WL 6405156 (N.D. Ind. Apr. 18 & 21, 2013).

> additional burden on, and additional expense to, Biomet.[83]

In *Progressive Casualty Insurance Company v. Delaney*,[84] in denying plaintiff's request late in the process to switch from search terms and manual review to TAR, the court criticized plaintiff's plan to apply TAR not to the entire document population, but only to documents hitting the search terms. According to the court, such a process would be inconsistent with the "best practices" guide of its TAR vendor.[85]

In *Rio Tinto PLC v. Vale S.A.*,[86] the court permitted the use of keyword culling before TAR because it was included in the parties' stipulated protocol. "The Court itself felt bound by the parties' protocol, such as to allow keyword culling before running TAR, even though such pre-culling should not occur in a perfect world." But the court also noted that "the standard for TAR is not perfection," nor "best practices," "but rather what is reasonable and proportional under the circumstances."[87]

Finally, in *Bridgestone Americas, Inc. v. International Business Machines Corp.*,[88] the court permitted plaintiff to undertake a hybrid approach, using TAR on documents initially identified through the use of search terms (but which still resulted in more than two million documents requiring review). The court expressly recognized that using predictive coding "is a judgment call."[89]

---

83. *In re Biomet*, 2013 WL 1729682, at *3.

84. Case No. 2:11-cv-00678, 2014 WL 3563467 (D. Nev. July 18, 2014).

85. *Id.* at *11.

86. Case No. 14 Civ. 3042, 2015 WL 4367250, at *1 (S.D.N.Y. July 15, 2015).

87. *See id*.

88. Case No. 3:13-1196, 2014 WL 4923014 (M.D. Tenn. July 22, 2014).

89. *Id*.

D. Disclosure of the Seed, Training, or Validation Sets

Disclosure of seed, training, or validation sets—including irrelevant documents and the responding party's coding decisions—has become one of the most contentious issues related to the use of TAR. The case law reflects a range of outcomes on the issue: courts encouraging—but not requiring—disclosure; responding parties voluntarily making disclosure; parties agreeing not to require disclosure; courts not requiring disclosure; one court requiring disclosure; and one court citing non-disclosure as a factor in its denial of a motion seeking approval to use TAR.[90]

1. Courts Encouraging Disclosure

Some courts have encouraged—but not required—disclosure of seed, training, or validation sets. For example, in *Da Silva Moore*, the defendant had voluntarily agreed to provide plaintiffs' counsel with both the documents in the seed and training sets and counsel's coding of those documents.[91] The court stated that, "[w]hile not all experienced ESI counsel believe it necessary to be as transparent as MSL was willing to be, such transparency allows the opposing counsel (and the Court) to be more comfortable with computer-assisted review."[92] The court further stated that it "highly recommends that counsel in future cases be willing to at least discuss, if not agree to, such transparency in the [TAR] process."[93]

---

90. *See, e.g.*, Rio Tinto PLC v. Vale S.A., 306 F.R.D. 125, 128 (S.D.N.Y. 2015) ("[W]here the parties do not agree to transparency, the decisions are split and the debate in the discovery literature is robust.").

91. Da Silva Moore v. Publicis Groupe, 287 F.R.D. 182, 192 (S.D.N.Y. 2012).

92. *Id*.

93. *Id*.

Similarly, in *Bridgestone*, the court advised that because it was allowing a change to the discovery approach midstream, the "Magistrate judge expects full openness in this matter."[94] In *Federal Housing Finance Agency v. JPMorgan Chase & Co.*, the court appeared to encourage disclosure of the training sets by (1) stating that for the TAR process to work, "I think it needs transparency and cooperation of counsel"; and (2) confirming that the responding party would be voluntarily providing access to the training sets.[95] In *Biomet*, while the court expressly held that it could not require such disclosure under the Federal Rules of Civil Procedure, it nevertheless encouraged the responding party to "re-think its refusal" in the "cooperative spirit" encouraged by *The Sedona Conference Cooperation Proclamation*.[96]

Additionally, in *Rio Tinto*, the court expressed its preference for disclosure, but recognized that there are alternative means of evaluating the effectiveness of the TAR process.[97] Although the parties stipulated to share such documents in their TAR protocol, which the court approved, the court chose to expand upon its order by providing guidance to litigants regarding the use of TAR. In so doing, the court observed that sharing training sets—including the irrelevant documents in the training set and counsel's coding decisions on them—is not necessary to ensure appropriate training of the TAR model. The court stated:

---

94. Bridgestone Ams., Inc. v. Int'l Bus. Machines Corp., Case No. 3:13-1196, 2014 WL 4923014, at *1 (M.D. Tenn. July 22, 2014).

95. Fed. Hous. Fin. Agency v. JPMorgan Chase & Co., No. 1:11-cv-06188-DLC (S.D.N.Y. July 24, 2012) (transcript at 9, 14).

96. *In re* Biomet M2A Magnum Hip Implant Prods. Liab. Litig., No. 3:12-MD-2391, 2013 WL 6405156, at *2 (N.D. Ind. Aug. 21, 2013).

97. Rio Tinto PLC v. Vale S.A., 306 F.R.D. 125, 128–29 (S.D.N.Y. 2015).

> [W]hile I generally believe in cooperation, requesting parties can insure [sic] that training and review was done appropriately by other means, 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.[98]

Additionally, the court cited studies showing that the contents of the "seed set" are much less significant with tools using "continuous active learning," in which the learning algorithm is continually retrained as reviewers review documents the algorithm identifies as potentially relevant (or potentially not relevant).[99]

2. Responding Parties Disclosing Voluntarily

In some cases, the responding party voluntarily agreed to disclose either a sample (or more) from the seed, training, or validation sets, or agreed to allow the opposing party to have some role in training the software.

In *Da Silva Moore,* for example, the responding party agreed to disclose the non-privileged documents in the seed set.[100] In *Bridgestone*, the plaintiff offered to share the seed documents.[101]

---

98. *See id*. (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, 298 (2014)).

99. *See id.* at 127 (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, 298 (2014) ("Disclosure of the seed or training set offers false comfort to the requesting party . . . .")) (ellipsis in original).

100. Da Silva Moore v. Publicis Groupe, 287 F.R.D. 182, 192 (S.D.N.Y. 2012).

101. Bridgestone Ams., Inc. v. Int'l Bus. Machines Corp., Case No. 3:13-1196, 2014 WL 4923014, at *11 (M.D. Tenn. July 22, 2014).

In *Federal Housing Finance Agency v. JPMorgan Chase & Co.*, the court approved defendant JP Morgan Chase's request to use TAR following its agreement to allow access to the relevant and irrelevant documents, other than privileged documents, in the seed set.[102] And in *Dynamo Holdings II* the responding party agreed to allow the requesting party to code the documents used to train the TAR algorithm.[103]

3. Courts Not Requiring Disclosure

In *Biomet*, the court denied plaintiffs' request for access to the training sets and to participate in training the TAR software.[104] Plaintiffs sought to impose a protocol for TAR similar to the one used in *In re Actos (Pioglitazone) Products Liability Litigation,*[105] in which each side nominated three experts to review the training sets and conduct quality control following TAR. The *Biomet* court rejected plaintiffs' request, observing that Federal Rule of Civil Procedure 26(b)(1) only makes relevant, non-privileged information discoverable, commenting that, "I'm puzzled as to the authority behind [the plaintiffs'] request."[106] The court also stated that although Sedona Conference principles and local discovery rules encourage parties to cooperate in discovery,

---

102. Fed. Hous. Fin. Agency v. JPMorgan Chase & Co., Case No. 1:11-cv-06188-DLC (S.D.N.Y. July 24, 2012) (transcript at 14–15, 24); *see also id.* at 8–9 (commenting that the reliability of TAR depends upon the process employed, particularly with respect to training the model using seed sets). *See also* Fed. Hous. Fin. Agency v. HSBC North America Holdings Inc., No. 1:11-cv-06188-DLC, 2014 WL 584300, at *3 (S.D.N.Y. Feb. 14, 2014) (same case).

103. Dynamo Holdings Ltd. P'ship v. Comm'r of Internal Revenue, No. 2685-11, slip op. at 6–7 (T.C. Jul. 7, 2016) (hereinafter *Dynamo Holdings II*).

104. *In re* Biomet M2a Magnum Hip Implant Prods. Liab. Litig., Case No. 3:12-MD-2391, 2013 WL 1729682 & 2013 WL 6405156 (N.D. Ind. Apr. 18 & Aug. 21, 2013).

105. MDL No. 6:11-md-2299, 2012 WL 7861249 (W.D. La. July 27, 2012).

106. *In re Biomet*, 2013 WL 6405156, at *1–2.

such cooperation does not require "counsel from both sides to sit in adjoining seats while rummaging through millions of files that haven't been reviewed for confidentiality or privilege."[107]

4. Case Requiring Disclosure

In *Independent Living Center v. City of Los Angeles*, the court ordered (on consent) the use of TAR to search more than two million documents after "little or no discovery was completed" before the discovery cutoff, and the parties had ongoing disputes after "months of haggling" over search terms that yielded large numbers of documents for review.[108] Although the defendant was initially concerned about the costs of using TAR, it agreed to do so when the court stated that it would only be required to produce the top 10,000 documents identified by the TAR tool. At the defendant's request, and to avoid subsequent disputes, the court also ordered that the plaintiff "be involved in and play an active role" in the training process, including making "relevance determinations" in the training documents.[109] The court held that the defendant was not necessarily required to engage in a quality-assurance process as part of the TAR protocol; however, if the plaintiff insisted upon such a process, then plaintiff would be required to pay for 50% of its costs.[110]

107. *See id*. at *2.

108. No. 2:12-cv-00551, slip op. at 1–2 (C.D. Cal. June 26, 2014).

109. *Id*.

110. *Id.*, slip op. at 2–3.

5. Non-Disclosure as a Factor in Denying the Use of TAR

One court has cited non-disclosure as a factor in denying a party's request to use TAR. In *Progressive*,[111] the court criticized plaintiff's unwillingness in its proposed TAR protocol to share with opposing counsel the irrelevant documents used to train the TAR tool. The court stated that "[i]n the handful of cases that have approved technology assisted review of ESI, the courts have required [sic] the producing party to provide the requesting party with full disclosure about the technology used, the process, and the methodology, including the documents used to 'train' the computer."[112]

E. Advance Court Approval for the Use of TAR

In *Dynamo Holdings I*,[113] the tax court addressed whether the court's advance approval was necessary for a party to use TAR. The court commented that the petitioner's request for advance court approval to use TAR (if the respondent's motion to compel was granted) was "somewhat unusual."[114]

---

111. Progressive Casualty Insurance Co. v. Delaney, Case No. 2:11-cv-00678, 2014 WL 3563467 (D. Nev. July 18, 2014).

112. *Id.* at *10 (citing Da Silva Moore v. Publicis Groupe, 287 F.R.D. 182 (S.D.N.Y. Feb. 24, 2012), and *In re* Actos (Pioglitazone) Prods. Liab. Litig., MDL No. 6:11-md-2299, 2012 WL 7861249 (W.D. La. July 27, 2012)). In both of those cases, however, the parties seeking to use TAR had *voluntarily stipulated* to allow access to the irrelevant training documents—the courts had not required it. *See Da Silva Moore*, 287 F.R.D. at 192 (noting that the responding party agreed to produce irrelevant documents in the seed or training sets); *In re Actos*, 2012 WL 7861249, at *4–5 (parties agreed to jointly review and code the documents used to train the predictive coding model).

113. Dynamo Holdings Ltd. P'ship v. Comm'r of Internal Revenue, 143 T.C. No. 9, 2014 WL 4636526 (Sept. 17, 2014) (hereinafter *Dynamo Holdings I*).

114. *Id.* at *3.

The court stated that "although it is a proper role of the court to supervise the discovery process and intervene when it is abused by the parties, the court is not normally in the business of dictating the process that they should use when responding to discovery."[115] "If our focus were on paper discovery," the court continued, "we would not (for example) be dictating to a party the manner in which it should review documents for responsiveness or privilege, such as whether that review should be done by a paralegal, a junior attorney, or a senior attorney."[116]

While stating that if the respondent believes "the ultimate discovery response is incomplete" then it could file a motion to compel "at that time," the court nevertheless took up the issue of whether TAR would be allowed because the court had "not previously addressed the issue of computer-assisted review tools."[117]

> Where, as here, petitioners reasonably request to use predictive coding to conserve time and expense, and represent to the Court that they will retain electronic discovery experts to meet with respondent's counsel or his experts to conduct a search acceptable to respondent, we see no reason

115. *Id*.

116. *Id*.; *cf*. The Sedona Conference, *The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production*, Principle 6 (2nd ed. 2007), *available at* https://thesedonaconference.org/download-pub/81 ("Responding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information.").

117. *Dynamo Holdings I*, 2014 WL 4636526, at *3.

> petitioners should not be allowed to use predictive coding to respond to respondent's discovery request.[118]

F. Miscellaneous Issues

A number of other issues have also arisen in cases discussing TAR. These have included what an acceptable measure of completeness might be; whether a party using TAR must respond to subsequent rounds of document requests that require it to retrain the TAR tool; and whether the party using TAR can manually review documents that TAR has identified as likely responsive before producing them.

1. Recall Thresholds

Few courts have addressed the issue of what the results of a "reasonable" TAR effort should be. Most of the cases that have addressed this issue have focused on recall, a measure of the proportion (or percent) of the responsive documents in the document population that have been correctly identified by the TAR tool or end-to-end review process.

The court in *Global Aerospace Inc. v. Landow Aviation, L.P.*,[119] approved, over the plaintiffs' objections, defendants' proposed TAR protocol targeting at least 75% recall. The case involved a multi-party action arising from the collapse of three hangars at Dulles Jet Center. Defendants moved for a protective order approving the use of TAR to review approximately 250 gigabytes of ESI, which they estimated to equate to more than two million

118. *Id.* at *4.

119. Case No. 61040 (Va. Cir. Ct. Apr. 23, 2012).

documents.[120] Defendants asserted that, "[a]t average cost and rates of review and effectiveness, linear first-pass review would take 20,000 man hours, cost two million dollars, and locate only sixty percent of the potentially relevant documents."[121] By contrast, TAR would—according to defendants—locate "upwards of seventy-five percent of the potentially relevant documents," at a fraction of the cost and in a fraction of the time of a traditional linear review.[122] Defendants proposed a TAR protocol that would ensure recall—i.e., the fraction of relevant documents that are identified by the TAR tool—of at least 75%, and would give opposing counsel access to documents reviewed in the training, stabilization, and validation processes (with the exception of privileged and sensitive irrelevant documents).[123]

Plaintiffs opposed the motion, arguing that defendants' estimate of the potential review population was overstated because they "copied every file from every computer" without any "attempt to separate the files pertaining to the Dulles Jet Center from the files pertaining to [defendants'] many other business and personal ventures" and, thus, traditional linear review of the files generated by the potential custodians "simply is not an unmanageable task."[124] The court overruled plaintiffs' objections and granted defendants' request, but made its order with-

120. *See* Defs.' Mem. in Support of Motion for Protective Order Approving the Use of Predictive Coding, Global Aerospace Inc. v. Landow Aviation, L.P., Case Nos. 61040, 2012 WL 1419842 (Va. Cir. Ct. Apr. 9, 2012).

121. *Id*.

122. *Id*.

123. *Id*.

124. *See* Opp. of Pls.: M.I.C. Indus., et al., to the Landow Defs.' Motion for Protective Order Regarding Elec. Documents and "Predictive Coding," Global Aerospace Inc. v. Landow Aviation, L.P., Case Nos. 61040, 2012 WL 1419842 (Va. Cir. Ct. Apr. 9, 2012).

out prejudice to any party raising an issue as to the completeness or contents of defendants' document production or the continued use of TAR.[125]

Similarly, in *Independent Living Center v. City of Los Angeles*,[126] the court anticipated that quality assurance would establish a recall rate of 75%, and stated that if the percentage was lower than 75%, then it would have to be brought to the court's attention.

2. Post-Production Challenge

In *Dynamo Holdings II*, the tax court addressed a post-production challenge to the sufficiency of a TAR process.[127] The Commissioner of Internal Revenue argued that the responding party's production using TAR was missing a substantial number of documents found through the use of search terms.[128] The requesting party sought to have the court order the responding party to start over with a manual review to remedy the alleged gaps in the production. The court noted that the parties had worked together to develop a TAR protocol, including how to select and review the seed and training sets.

The requesting party (i.e., the Commissioner) had coded the documents used to train the TAR algorithm and, given the option of different recall and associated precision rates, had selected a recall rate of 95%. The court assumed that the TAR process was flawed, but stated that "the question remains whether

125. *See* Order Approving the Use of Predictive Coding For Discovery, *Global Aerospace*, Consol. Case. No. CL 61040 (Va. Cir. Ct. Apr. 23, 2012).

126. No. 2:12-cv-00551, slip op. at 3 (C.D. Cal. June 26, 2014).

127. Dynamo Holdings Ltd. P'ship v. Comm'r of Internal Revenue, No. 2685-11 (T.C. Jul. 7, 2016).

128. *Id.*, slip op. at 6.

any relief should be afforded."[129] It decided that the responding party had made a "reasonable inquiry" using TAR by producing documents "that the algorithm determined [were] responsive."[130]

The court reasoned that the requesting party's motion was "predicated on two myths,"[131] i.e., that manual human review "constitutes the gold standard" and that the rules require a "perfect response."[132] Specifically, in response to discovery requests, Tax Court Rule 70(f)—which is analogous to Federal Rule of Civil Procedure 26(g)—"requires the attorney to certify, to the best of their knowledge formed after a 'reasonable inquiry,' that the response is consistent with our Rules, not made for an improper purpose, and not unreasonable or unduly burdensome given the needs of the case." The court stated that "when the responding party is signing the response to a discovery demand, he is not certifying that he turned over everything, he is certifying that he made a reasonable inquiry and to the best of his knowledge, his response is complete."[133]

The court concluded that "there is no question that petitioners satisfied our Rules when they responded using predictive coding."[134]

---

129. *Id.*, slip op. at 7.

130. *Id.*, slip op. at 9.

131. *Id.*, slip op. at 7.

132. *Id.*, slip op. at 7–8.

133. *Id.*, slip op. at 8.

134. *Id.*, slip op. at 9.

3. Retraining the TAR Tool for Subsequent Document Requests

At least one case has dealt with the issue of whether the responding party can be required to respond to additional document requests after it has already used TAR to respond to a prior round of requests. In *Smilovits v. First Solar*,[135] the court held that defendants' use of TAR in response to plaintiffs' first round of document requests did not confine plaintiffs' document discovery to the first round of requests. The court also noted that defendants had not explained why the search for additional documents required the use of TAR, nor had they provided any concrete information about the costs to "retrain" the TAR tool to deal with subsequent requests.[136]

4. Manual Review Following TAR

In *Chen-Oster v. Goldman, Sachs & Co.*,[137] plaintiffs sought to compel Goldman Sachs to produce all documents hitting on agreed-upon search terms without further review. The court observed that with TAR—the court considered the use of search terms to be a form of TAR—parties can agree to produce documents without human review, but the parties had not done so in this case. The court stated that because Goldman Sachs had not agreed to produce the documents without further human review—and the court had not ordered it—Goldman Sachs was not precluded from reviewing the documents before production.[138]

135. No. 2:12-cv-00555, slip op. at 1–2 (D. Ariz. Nov. 20, 2014).

136. *Id*.

137. Case No. 10 Civ. 6950, 2014 WL 716521 (S.D.N.Y. Feb. 18, 2014).

138. *See id*. at *1.

Similarly, in *Good v. American Water Works*,[139] the defendants proposed a privilege review using both TAR and human review, along with a Federal Rule of Evidence 502(d) claw-back order. Plaintiffs argued that to ensure expedited production, and because of the protection afforded by the 502(d) order, defendants should not be permitted to manually review the documents. The court approved defendants' proposed protocol, finding that "their desired approach is a reasonable one."[140] The court stated that it was approving the protocol "with the expectation that the defendants will marshal the resources necessary to assure that the delay occasioned by manual review" would be "minimized," and the production would be accomplished quickly.[141] The court also stated that if "undue delay" threatened to jeopardize compliance with the discovery schedule, plaintiffs could file a motion requesting that the court reconsider ordering defendants to use plaintiffs' requested approach.[142]

5. Use of TAR in Government Investigations

Some government agencies have accepted the use of TAR for search and review in connection with document productions in regulatory investigations. In August 2015, the Federal Trade Commission issued an update to its Model Second Request for merger antitrust investigations that includes specifications related to the use of TAR in response to Second Requests (requiring that the responding party disclose the specified information

139. Case No. 2:14–01374, 2014 WL 5486827, at *2–3 (S.D.W. Va. 2014).

140. *Id.* at *3.

141. *Id.* at *4.

142. *Id*.

at the end of the process).[143] In particular, the responding party must:

> [b](i) describe the collection methodology, including: (a) how the software was utilized to identify responsive documents; (b) the process the Company utilized to identify and validate the seed set documents subject to manual review; (c) the total number of documents reviewed manually; (d) the total number of documents determined nonresponsive without manual review; (e) the process the Company used to determine and validate the accuracy of the automatic determinations of responsiveness and non-responsiveness; (f) how the Company handled exceptions ('uncategorized documents'); and (g) if the Company's documents include foreign language documents, whether reviewed manually or by some technology-assisted method; and [b](ii) provide all statistical analyses utilized or generated by the Company or its agents related to the precision, recall, accuracy, validation, or quality of its document production in response to this Request; and [c] identify the Person(s) able to testify on behalf of the Company about information known or reasonably available to the organization, relating to its response to this Specification.[144]

Similarly, counsel for the Antitrust Division of the Department of Justice has provided guidance regarding TAR protocols

---

143. Fed. Trade Comm'n, *Model Request for Additional Information and Documentary Material (Second Request)*, at 15–16 (revised Aug. 2015), https://www.ftc.gov/system/files/attachments/merger-review/guide3.pdf.

144. *Id.* at 16.

in response to Division investigations, which should be addressed with the DOJ prior to embarking on a TAR-based review.[145] Notably, the Definitions and Instructions section of the DOJ's Model Second Request states the following:

> Before the company or its agent uses software or technology to identify or eliminate potentially responsive documents and information produced in response to this Request, including but not limited to search terms, predictive coding or similar technology, near-deduplication, deduplication, and email threading, the company must provide a detailed description of the method(s) used to conduct all or any part of the search.[146]

145. U.S. Dep't of Justice, *Request for Additional Information and Documentary Material (Model Second Request)*, at 13 (June 2015), https://www.justice.gov/atr/request-additional-information-and-documentary-material-issued-weebyewe-corporation.

146. *Id*.

VI. INTERNATIONAL ADOPTION OF TAR

The use of TAR has been accepted in several foreign jurisdictions.

In Ireland, the Irish High Court in *Irish Bank Resolution Corp. v. Quinn* granted a responding party's motion to use TAR over the objection of the party requesting the production of documents, a ruling upheld by the Irish Court of Appeal.[147]

In England, the English High Court in *David Brown v. BCA Trading* approved the use of TAR over the objection of the requesting party.[148] And in *Pyrrho Investments Ltd. v. MWB Property Ltd.* the parties jointly sought and obtained the approval of the English High Court to use TAR.[149]

In Australia, the Federal Court of Australia in *Money Max v. QBE Insurance Group* issued the first decision of an Australian court addressing the use of TAR. The court ordered the responding party to provide several categories of information about its TAR process and for the parties to meet and confer about any disputes regarding the process.[150] Soon thereafter, in *McConnell Dowell v. Santam Ltd,* the Supreme Court of Victoria issued another opinion approving the use of TAR, which the

147. Irish Bank Resol. Corp. v. Quinn, [2015] IEHC 175 (H. Ct.) (Ir.), upheld by the Irish Court of Appeal (*see Court of Appeal Approves use of TAR for Discovery*, McCann Fitzgerald (2016), http://www.mccannfitzgerald.com/McfgFiles/knowledge/6802-Court%20of%20Appeal%20Approves%20Use%20of%20Tar%20For%20Discovery.pdf).

148. David Brown v. BCA Trading Ltd., [2016] EWHC (Ch) 1464 (Eng.).

149. Pyrrho Inv. Ltd. v. MWB Prop. Ltd., [2016] EWHC (Ch) 256 (Eng.).

150. Money Max Int'l Pty Ltd. (Tr.) v. QBE Ins. Grp. Ltd. [2016] FCAFC 148 at 3-4 (Austl.).

parties had agreed to use and a special discovery master had recommended to the court.[151]

151. McConnell Dowell Constructors (Aust) Pty Ltd. v Santam Ltd. & Others (No 1) [2016] VSC 734 (Austl.).

VII. EVOLVING VIEWS OF TAR

There appears to be some evolution in thinking about TAR since *Da Silva Moore* was decided in 2012. For example, there seems to be an increased comfort level within the legal community with the reliability of TAR, most clearly reflected in the *Rio Tinto PLC v. Vale S.A.*[152] decision in early 2015.

In *Rio Tinto*, the court's discussion reflects that TAR technologies are evolving in ways that may impact some of the issues that have, to date, been controversial in the use of TAR, for example, some requesting parties' concerns about the composition of seed and training sets and the demand for their disclosure.

The *Rio Tinto* court also noted that recent studies have shown that with TAR tools employing continuous active learning, the seed set may have little or no impact, and that as a practical matter, there may be no discrete training sets to share.[153]

The court in *Dynamo Holdings I*[154] expressed similar views to those expressed in *Rio Tinto*; the court rejected the respondent's assertion that predictive coding is an "unproved technology," noting that "the understanding of e-discovery and electronic media has advanced significantly in the last few years, thus making predictive coding more acceptable in the technology industry than it may have previously been."[155]

---

152. 306 F.R.D. 125 (S.D.N.Y. 2015).

153. *Id.* at 128 (citing Gordon V. Cormack & Maura R. Grossman, *Evaluation of Machine Learning Protocols for Technology-Assisted Review in Electronic Discovery*, in Proceedings of the 37th Int'l ACM SIGIR Conf. on Research & Dev. in Info Retrieval, at 153–62 (ACM New York, N.Y. 2014), http://dx.doi.org/10.1145/2600428.2609601).

154. Dynamo Holdings Ltd. P'ship v. Comm'r of Internal Revenue, 143 T.C. No. 9, 2014 WL 4636526 (Sept. 17, 2014).

155. *Id.* at *5.

The *Dynamo Holdings I* court added that "[i]n fact, we understand that the technology industry now considers predictive coding to be widely accepted for limiting e-discovery to relevant documents and effecting discovery of ESI without an undue burden."[156]

Whether these evolving views of TAR will translate into widespread adoption in practice remains to be seen. But, further changes in technology are likely to continue to shape and impact the evolution of TAR case law.

156. *Id*.

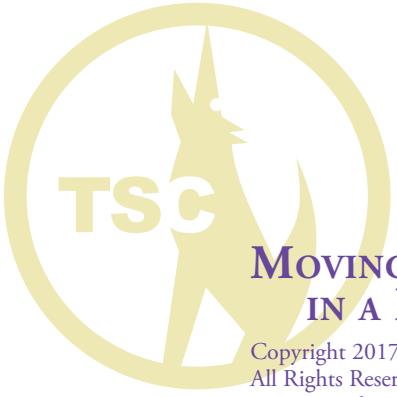
VIII. CONCLUSION

While the case law reflects a broad consensus that TAR is an acceptable search and review methodology, certain issues regarding the details of its use remain unresolved. The general principles set forth in the cases discussed in this *Primer* should provide useful guidance to courts and parties seeking to use TAR to achieve the goals of Federal Rule 1 (the just, speedy, and inexpensive resolution of legal proceedings) and Rule 26(b)(1) (proportionality).[157] The Bench and Bar should continue to actively monitor research and case law developments in this area.

157. *See* FED. R. CIV. P. 26(b)(1), Advisory Committee Note to 2015 Amendment ("Computer-based methods of searching such information continue to develop, particularly for cases involving large volumes of electronically stored information. Courts and parties should be willing to consider the opportunities for reducing the burden or expense of discovery as reliable means of searching electronically stored information become available.").

IX. INDEX OF CASES

Page(s)



Moving the Law Forward in a Reasoned & Just Way