Reality Bites: Why TAR's Promises Have Yet to be Fulfilled

William P. Butterfield, Conor R. Crowley & Jeannine Kenney[1]

I. Introduction

The complexity and costs associated with discovery of electronically stored information ("ESI") have attracted increasing attention in recent years as the volume of ESI that organizations maintain, and thus may be required to preserve and produce in litigation, has ballooned.[2] At the same time, the effectiveness of more traditional tools used to cull non-responsive from responsive ESI— both manual "eyes-on" review and the use of keyword searching—has been credibly questioned by scientific studies.

It is now well recognized, at least among information retrieval scientists, that human, manual review is not only entirely impractical in the age of ESI, but is also far from the "gold standard" of review against which other search tools should be judged: inconsistency and errors in human review lead to omission of responsive documents and inclusion of non-responsive ones compared to other methodologies.[3] And the potential problems with keyword searching using Boolean operators, the accepted methodology to which most attorneys have turned given the volume of ESI and impracticality of manual review, are well-established. Even productions resulting from carefully designed search terms developed through a fully cooperative and

---

[1] William P. Butterfield is a partner at Hausfeld LLP and is Vice Chair of the Sedona Conference Working Group on Electronic Document Retention and Production, where he served as editor-in-chief of the Case for Cooperation (2009), and co-editor of The Sedona Conference® Commentary On Preservation, Identification and Management of Sources of Information that are Not Reasonably Accessible (2008). He teaches a course in electronic discovery at American University, Washington College of Law, and serves on the advisory board of Georgetown University Law Center's Advanced E-Discovery Institute.

Conor R. Crowley is the founder of Crowley Law Office and is Chair of The Sedona Conference Working Group on Electronic Document Retention and Production, where he served as Editor-in-Chief of The Sedona Conference Commentary on Proportionality in E-Discovery (2013), and as Senior Editor of The Sedona Principles (Second Edition) (2007) and The Sedona Conference Commentary on Legal Holds (2010). He sits on the Advisory Board of Georgetown University Law Center's Advanced E-Discovery Institute and Bloomberg BNA's Digital Discovery and E-Evidence.

Jeannine M. Kenney is an associate at Hausfeld LLP, with a focus on e-discovery matters.

[2] *See, e.g.*, John Gantz & David Reinsel, *Extracting Value from Chaos*, IDC (June 2011) *available at*, http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf (discussing growth in the generation of ESI); *The Data Deluge*, The Economist, Feb. 25, 2010 *available at* http://www.economist.com/node/15579717?story_id=15579717 (noting information created increased by nearly 10-fold between 2005 and 2010).

[3] *See generally* Maura R. Grossman & Gordan V. Cormack, *Technology-Assisted Review in E-Discovery Can be More Effective and More Efficient Than Exhaustive Manual Review*, Richmond J. L. & Tech. Vol. XVII, No. 3, Article 11 (2011); *see also* Nicholas M. Pace & Laura Zakaras, RAND Inst. for Civil Justice, *Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery* 55-58 (2012), *available at* http://www.rand.org/content/dam/rand/pubs/monographs/2012/RAND_MG1208.pdf. .

iterative process may suffer from both poor precision and poor recall,[4] increasing the costs of pre-production review by the producing party and omitting too many relevant documents, thereby denying the requesting party documents that may be critical to the case. Courts increasingly recognize these concerns.[5]

In this context, Technology Assisted Review ("TAR") appears to offer a way out of the morass of traditional, but often problematic, legal search methodologies. Appropriately applied, it can offer a more accurate and cost-effective means of identifying and producing responsive ESI, at least for larger cases, compared to more traditional tools.[6]

Why then, given the advantages of this search technology, has it not been more widely adopted by parties in appropriate cases? To a certain extent, that reluctance stems from fear or lack of understanding of TAR and a comfort with old patterns.[7] Others, however, suggest that the lack of judicial decisions formally endorsing TAR has fostered concern that litigants using it are vulnerable to post-hoc production challenges that may result in costly do-overs.[8] Some suggest that without a judicial imprimatur, few litigants will rely on TAR as the sole means to identify and produce responsive ESI, although the legal basis for defending the use of TAR has been clearly outlined.[9] Finally, some potential TAR users are concerned that the cost of extensive negotiations with the opposing party and ancillary litigation over the application of a

---

[4] *See, e.g.*, The Sedona Conference, *Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery*, 8 Sedona Conf. J. 189, 201-02 (2007), *available at* http://www.thesedonaconference.org/; Maura R. Grossman & Terry Sweeney, *What Lawyers Need to Know About Search Tools: The Alternatives to Keyword Searching Include Linguistic and Mathematical Models for Concept Searching*, Nat. L. J. (Aug. 23, 2010).

[5] See, e.g., Victor Stanley, Inc. v. Creative Pipe, Inc., 250 F.R.D. 251, 256-57 (D. Md. 2008) (discussing the risks and difficulties of keyword searching); Nat'l Day Laborer Org. Network v. U.S. Immig. & Customs Enforcement Agency, 877 F. Supp. 2d 87, 109 n.112 (S.D.N.Y. 2012) ("There is increasingly strong evidence that '[k]eyword search[ing] is not nearly as effective at identifying relevant information as many lawyers would like to believe.'") (quoting Grossman & Sweeney, supra n.3) (alteration in original)).

[6] *See, e.g.*, John M. Facciola, M.J., *Foreword* to *The Grossman-Cormack Glossary of Technology-Assisted Review*, 7 Fed. Ct. L. Rev. 1, 4 (2013) ("It is now indubitable that technology-assisted review is an appreciably better and more accurate means of searching a set of data.").

Estimates of cost savings have varied widely, but all the data suggest the savings are significant. *See, e.g.*, Ann Kershaw & Joe Howie, *Judges' Guide to Cost Effective E-Discovery* 13-14 (45 percent reduction in review costs);. Pace & Zakaras, *supra* n.2, at 66-69 (estimates of reported cost savings ranged from 0 percent up to nearly 80 percent).

[7] *See* Pace & Zakaras, *supra* n.2, at 72-76, 80.

[8] *See id.* at Pace & Zakaras, *supra* n. 2, at 72-73, 77-80; Henry J. Kelston, et al., *Technology Assisted Review From the Plaintiffs' Side*, Law Technology News, Dec. 10, 2012, *available at* http://www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=1202580755574&TechnologyAssisted_Review_From_the_Plaintiffs_Side&slreturn=20130316204602

[9] *See, e.g.*, Conor R. Crowley, *Defending the Use of Analytical Software in Civil Discovery*, Digital Discovery & e-Evidence, 10 DDEE 16, Sept. 16, 2010.

TAR protocol in a given case will outweigh the cost-savings otherwise achieved by the tool.[10] Thus TAR has been used chiefly by producing parties as a means of prioritizing and expediting the review and production of documents identified by negotiated search terms or for case analysis,[11] not as a substitute for keyword searching.

Information retrieval scientists and some e-discovery practitioners now herald recent U.S. court decisions endorsing TAR as a sign that technology assisted review will replace less effective legal search tools.[12] Words like "landmark," "watershed," and "sea-change" are not infrequently invoked to describe the significance of these decisions. Unquestionably, these court decisions provide reason for optimism. But in two leading cases, courts endorsed TAR in the context of transparency and cooperation among the parties that is infrequently seen in the litigation trenches. ESI practitioners should not expect, without a dramatic change in the mindset of litigators, these decisions to alter attorneys' resistance to disclosure of information retrieval methodologies and processes. More commonly, even when traditional search term methodologies are used, requesting parties face a methodological black box, denied even basic statistical testing results that provide assurances about the outcome of the process, if not the methods employed. And other recent court decisions suggest that, if parties do not voluntarily agree to a transparent TAR process, there are limits on the degree to which courts will require advance disclosures about the producing party's methodologies. At least one such decision has suggested that some courts may place the cost burdens of TAR on requesting parties if they demand that producing parties use it.

---

[10] *See* Jeff Fowler, Waiving Work Product with Predictive Coding, Sept. 7, 2012, Podcast, ESI Bytes, *available at* http://www.esibytes.com/waiving-work-product-with-predictive-coding/ (noting that the cost of negotiating and litigating a TAR protocol before it is implemented can eliminate cost advantages of TAR); *cf.* Pace & Zakaras, *supra* n.2, at 73 (noting concerns that "costs required to completely eliminate any shortfalls in recall or precision would be out of proportion to any benefit of the additional information"); *id.* at 71 & n.3 (noting that disclosure of TAR may lead to litigation that may negate cost advantages).

[11] Kelston, et al., *supra* n.7. Others suggest TAR may be used by producing parties without disclosure. *See* Pace & Zakaras, *supra* n.2, at 71 & n.3 ("We have been told that computer-categorization techniques have, in fact, been employed in document reviews but that the use was not disclosed to opposing parties or the court.")

[12] *See, e.g.*, Correy E. Stephenson, *After breakthrough year, predictive coding catching on*, New England In-House, Feb. 28, 2013, *available at* http://newenglandinhouse.com/2013/03/05/after-breakthrough-year-predictive-coding-catching-on/; Ritter Academy, *Why Court Approval of Computer-based Searches Changes Everything* (commenting on Moore v. Publicis Groupe SA, 2012 U.S. Dist. LEXIS 58742 (S.D.N.Y. Apr. 25, 2012) approving the use of predictive coding, and asserting that "the art of discovery is dying. The science of discovery is roaring with the momentum of a shuttle launching itself away from the gravitational pull of the earth."), *available at* http://www.ritteracademy.com/blog/why-court-approval-computer-based-searches-changes-everything; Patricia Antezana, *E-Discovery: Predictive coding decisions may signal change in identifying relevant evidence*, Inside Counsel, June 12, 2012, *available at* http://www.insidecounsel.com/2012/06/12/e-discovery-predictive-coding-decisions-may-signal ("The decisions by [courts] approving the use of predictive coding have set the stage for significant change in the legal landscape of court-accepted protocols for identifying relevant and responsive ESI.").

This paper posits that, despite growing judicial acceptance of TAR, enthusiasm over the likelihood that TAR will be widely adopted by parties should be tempered with a dose of litigation reality. While information scientists may advocate cooperation and transparency as the way to achieve the best possible results from the use of TAR, parties in litigation may have tactical, strategic or ethical concerns that limit their willingness to fully cooperate in a transparent manner. Additionally, parties in litigation are not required to achieve the best possible results from TAR, but rather to only certify that they have undertaken a reasonable investigation to identify and produce non-privileged, responsive information. Finally, parties in litigation have been motivated to consider the use of TAR not only because of its promise of greater levels of precision and recall than other review methods, but also because of its promise of significantly reduced review costs. This difference in motivations and objectives may help to explain why such a promising tool has not been as widely adopted as scientists would expect.

We begin with a discussion of recent U.S. court decisions generating considerable enthusiasm about the advent of TAR as a mainstream legal search tool, focusing on what they do and do not portend for increased transparency. We then briefly review the doctrines and concerns held by producing parties that impede and discourage the transparency we believe could facilitate wider adoption of TAR. We conclude by offering a checklist of disclosures we believe parties should consider to engender effective and cooperative use of TAR and thus its widespread acceptance, and propose ways in which the competing concerns regarding transparency may be balanced.

II. Judicial Acceptance of Technology Assisted Review

Over the years, as judicial skepticism over the reliability of keyword searching has grown, courts have inched toward acceptance of advanced analytics as a viable information retrieval technology. In 2007, when directing parties to meet and confer to develop a search protocol, Federal District Court Judge Shira Scheindlin noted "recent scholarship that argues that concept searching, as opposed to keyword searching, is more efficient and more likely to produce the most comprehensive results."[13] Five years later, Judge Scheindlin, after discussing the shortcomings of keyword searches, concluded that "parties can (and frequently should) rely on latent semantic indexing, statistical probability models, and machine learning tools."[14] Federal District Court Judge Paul Grimm has likewise pointed to "a growing body of literature that highlights the risks [of] conducting an unreliable or inadequate keyword search or relying exclusively on such searches," and suggested more advanced tools like concept searching may be more appropriate methodologies.[15] And notably, the Southern District of New York's Pilot Project for Case Management Techniques for Complex Civil Cases requires participants in the project to make a joint ESI submission regarding disclosure of ESI search techniques to be used, and identifies advanced analytical tools, including "machine learning" as appropriate

[13] *Disability Rights Council of Greater Wash. v. Wash. Metro. Transit Auth*., 242 F.R.D. 139, 148 (D.D.C. 2007).

[14] *Nat'l Day Laborer Org. Network*, 877 F. Supp. 2d at 109.

[15] Victor Stanley, 250 F.R.D. at 256-57. When *Victor Stanley* was decided Judge Grimm served as a Magistrate Judge. He now serves as a District Court Judge for the U.S. District Court of Maryland.

approaches.[16] Finally, Kroll Ontrack reported that, in 2012, nine percent of cases addressing e-discovery issues discussed TAR or predictive coding.[17]

No court had endorsed use of TAR as a defensible production methodology in a specific case until Magistrate Judge Peck broke the ice in *Da Silva Moore v. Publicis Groupe*, formally "recogniz[ing] that computer assisted review is an acceptable way to search for relevant ESI in appropriate cases."[18] At least four subsequent court decisions in the past year further highlight the growing acceptance of TAR by the court or the parties, or both. More importantly, these decisions suggest a shift in focus away from the acceptability of TAR as a tool and toward the process used to employ it and the results obtained. Much ink has been spilled about *Da Silva Moore* and these subsequent decisions, and thus we limit our discussion here to only those details necessary to distill key issues relevant to our thesis.

A. Global Aerospace, Kleen Products & *Biomet*: Court Deference to Producing Parties

Several widely discussed court decisions, which came on the heels of *Da Silva Moore*, have attracted the attention of the information retrieval community. First, in *Global Aerospace, Inc. v. Landow Aviation, LP*,[19] a state court permitted defendants, over plaintiffs' objections, to use a TAR methodology that did not involve seed sets or keywords to search a corpus of some 1.3 million potentially responsive, searchable documents. Importantly, the court did so without prejudice to plaintiffs' right to object, *post-hoc*, to the results obtained and the completeness of the production. During the hearing, the court noted that the producing party ordinarily has the right to select its search and review methodology, subject to the propounding party's right to object if it finds responsive documents have not been produced. Importantly, the court did not specifically endorse the technology or conclude it was better suited than other legal search tools; it merely let defendants proceed as they chose. According to subsequent reports, following application of TAR, defendants presented results showing the methodology produced a precision rate of 80 percent, and a recall rate of 81 percent,[20] to which the plaintiffs remaining in the case apparently did not object or respond.[21]

---

[16] Standing Order, *In re: Pilot Project Re: Case Mgmt. Techniques for Complex Civil Cases in the S. Dist. of N.Y.*, No. 11-mc-0388 (ECF No. 1) at 21 (S.D.N.Y. Nov. 11, 2011).

[17] Julie Beck, *Study Analyzes E-discovery Trends in 2012*, Inside Counsel, Dec. 31, 2012, *available at* http://www.insidecounsel.com/2012/12/31/study-analyzes-e-discovery-trends-in-2012.

[18] No. 11-civ-1279, U.S. Dist. LEXIS 23350 at *1 (S.D.N.Y. Feb. 24, 2012) *adopted and obj. overruled, sub nom. Moore v. Publicis Groupe SA*, 2012 U.S. Dist. LEXIS 58742 (S.D.N.Y. Apr. 25, 2012) (Carter, J.).

[19] No. CL 61040, 2012 VA. Cir. LEXIS 50 (Va. Cir. Ct., Apr. 23, 2012).

[20] Joe Palazzola, *How a Computer Did the Work of Many Lawyers*, Wall St. J. Law Blog, Jan. 17, 2013, *available at* http://blogs.wsj.com/law/2013/01/17/how-a-computer-did-the-work-of-many-lawyers/?mod=WSJBlog.

[21] Evan Koblentz, *Predictive Coding Completed in* 'Global Aerospace' *Case*, L. Tech. News, Jan. 16, 2013, *available at* http://www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=1202584738091&Predictive_Coding_Completed_in_Global_Aerospace_Case&slreturn=20130321180628.

Second, in an interesting twist, in *Kleen Products LLC v. Packaging Corporation of America*,[22] it was plaintiffs, not defendants, who argued that TAR was necessary to ensure that responsive documents were captured and produced, asserting that keyword searching used by defendants was not sufficiently reliable compared to TAR. Defendants balked, at least in part, because they had already applied a keyword search methodology along with iterative testing and validation procedures, to respond to the first set of production requests. Plaintiffs, on the other hand, were requesting a "do-over." In the end, the court conducted two days of evidentiary proceedings to resolve the dispute and required the parties to meet and confer to identify whether refining searches or other steps could provide plaintiffs with assurances that they were receiving a high percentage of responsive documents. Five months and many meet and confers and court conferences later, plaintiffs agreed to withdraw their demand that TAR be applied to existing production requests, but the parties agreed to meet and confer on the appropriate search methodology for any subsequent requests, leaving the door open to using TAR for later productions. The Court, in recounting this resolution, noted that "responding parties are best situated to evaluate the procedures, methodologies, and techniques appropriate for preserving and production their own [ESI]."[23]

More recently, in *In re Biomet M2A Magnum Hip Implant Products Liability Litigation*, the court permitted the producing party to use a hybrid methodology, using search terms to cull a 19.5 million document set to 2.5 million documents, to which TAR was then applied to identify responsive documents.[24] Plaintiffs objected not to the use of TAR *per se*, but to Biomet's unilateral decision to apply search terms to cull the corpus prior to application of TAR—a process they challenged as too exclusionary—without meeting and conferring with plaintiffs. Plaintiffs asked the TAR process be reapplied without the use of search terms and demanded involvement in the computer training process. In response, Biomet offered to produce all 2.5 million documents resulting from the keyword search (responsive or not)—a proposal that did not address plaintiffs' concerns that responsive documents were improperly culled from the original set of 19.5 million documents.

The court's rejection of plaintiffs' proposed do-over was based on both its assessment of the reliability statistics disclosed in the case (which it concluded satisfied Biomet's production obligations) and consideration of the seven-figure cost involved in rerunning the predictive coding process. The court noted that:

> Statistical sampling tests of a random sample [of the unselected documents] projected, with a 99 percent confidence rate, that between .55 and 1.33 percent of the unselected documents would be responsive and (with the same confidence level) that between 1.37 and 2.47 percent of the original 19.5 million documents

[22] No. 10-cv-571, 2012 U.S. Dist. LEXIS 139632 (N.D. Ill. Sept. 28, 2012).

[23] *Id.* at *18 (citing Principle 6 of the Sedona Best Practices Recommendations & Principles for Addressing Electronic Document Production of The Sedona Conference, The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery, 8 Sedona Conf. J. 189, 193 (Fall 2007)).

[24] Order Regarding Discovery of ESI (ECF No.396), No. 12-md-2391 (N.D. Ind. Apr. 18. 2013).

> were responsive. In comparison, [Defendant's] keyword/deduplication approach had identified 16 percent of the original 19.5 million.[25]

Despite plaintiffs' argument that Biomet chose to gamble on the defensibility of its methodology when it unilaterally deployed it, the court concluded that if plaintiffs wanted a do-over, they would have to pay for it.[26]

These cases are important not only because of their outcome, but also because of the deference the courts paid to the producing parties' choice of search methodology and procedures.[27] That deference may lend greater force to objections by producing parties that a mandate for transparency in TAR deviates from the principle espoused by Principle 6 of the Sedona Best Practices Recommendations & Principles for Addressing Electronic Document Production—that responding parties are best situated to evaluate the appropriate means of producing ESI.[28] Thus, as discussed *infra* Part II, litigants are often reluctant to reveal specifics about their ESI itself and the tools used to search it; and they are even less willing to involve the requesting party in their search process. Yet that lack of transparency and cooperation regarding the nature of the ESI and the search methodologies to be applied may impede the parties' ability to reach agreement on the use of TAR. Consider the following observation by counsel for Plaintiffs in *Global Aerospace* explaining their reluctance to grant what they describe as Defendants' request for an "advance waiver" on the acceptability of predictive coding absent disclosures:

> [C]ounsel for the . . . plaintiffs were reluctant to provide any such waiver without first obtaining information showing that the proposed use of predictive coding was the best method for locating responsive documents.
>
> We never received this comfort; many of our questions about the data went unanswered. The [defendants'] counsel refused to state how many emails they had accumulated, how many emails dated from the key time period, how many emails they had collected from the principal officer of the [defendant] entities, or how any of the documents were foldered or organized. The plaintiffs therefore did not consent to the use of predictive coding. The [defendants] then filed a motion for a protective order.
>
> . . . .

---

[25] *Id.* at 2.

[26] *Id.* at 7.

[27] In contrast to this deference, Delaware Chancery Court Judge Travis Laster, in addition to ordering parties to select a joint e-discovery vendor to host and review documents, ordered the parties to show cause why predictive coding should not be used in the case although, apparently, neither party had proposed its use. *See* Tr. 66:12-67:13, *EORHB, Inc. v. HOA Holdings, LLC*, C.A. No. 7409-VCL (Del. Ch. Oct. 19, 2012).

[28] *See* David J. Kessler, *Search Terms Are More Than Mere Words*, New York L. J., Mar. 21, 2011, *available at* http://www.fulbright.com/images/publications/20110321SearchTermsAreMoreThanMereWords.pdf .

> As the failed negotiations in *Global Aerospace* demonstrate, it is difficult to gain the comfort necessary to bless a particular search protocol without receiving extensive information about the documents.[29]

Would plaintiffs have provided that waiver with the appropriate disclosures, avoiding the risk that Global might have to rerun its production?

Did Biomet take an ill-advised gamble by making unilateral decisions on its hybrid search methodology given growing court recognition of the deficiencies (including under-inclusiveness) of search terms, particularly for such a large production? Its statistics on reliability, though accepted by the court, do not seem to support a finding that a reasonable search was conducted. In commentary on the decision, information retrieval experts Maura Grossman and Gordon Cormack noted that these statistics raise concerns about the reliability of the hybrid approach. They conclude that recall for the keyword search itself was just 50 percent; and assuming that predictive coding resulted in about 70 percent recall in the set to which the methodology was applied, the hybrid approach likely identified only 35 percent of the relevant documents from among the 19.5 million.[30] Given those statistics, although Biomet may not be required to re-run its predictive coding, it may not avoid costly post-production challenges to the adequacy of its production.

B. Da Silva Moore & *Actos:* Transparency & Cooperation Lead to Acceptance

The approaches taken by the parties in *Da Silva Moore* and *In re Actos Products Liability Litigation* ("*Actos*") strongly suggest that, despite the deference to the producing parties' choices of review tools and procedures noted in the cases discussed *supra*, transparency and cooperation are more likely to promote both judicial and party acceptance of TAR. Both cases predicated their approval of TAR on significant transparency—to the court and the requesting party—and on significant involvement of the requesting party in the TAR process.

Importantly, in *Da Silva Moore*, Magistrate Judge Peck sent an early signal that he viewed transparency with respect to TAR as a necessary predicate to his endorsement of its use, advising defendants that if they used that methodology they would be required to turn over their seed set to plaintiffs, including documents tagged as non-responsive in the training rounds, to enable them to determine whether the computer was being appropriately trained.[31] After having agreed to a detailed predictive coding protocol, that cooperation broke down as defendants and plaintiffs ultimately engaged in a heated dispute over how the success of TAR would be measured.

---

[29] *Id.*

[30] *In re: Biomet – Doing the Math on Court Approved Multimodal Review*, IT-Lex Technology Law Blog, April 22, 2013, *available at* http://it-lex.org/in-re-biomet-doing-the-math-on-court-approved-multimodal-review/.

[31] Da Silva Moore v. Publicis Groupe, 2012 U.S. Dist. LEXIS 23350, at *7 (S.D.N.Y. Feb. 24, 2012).

What transpired up to that point, before the process unraveled, is, to us, far more important than the unraveling.[32] Notably, the parties had entered into an extensive TAR protocol replete with plaintiff input into the process, with the parties agreeing: (1) on data sources from which documents could be collected; (2) on the confidence level and interval for sampling, and the sample size for the initial seed set; (3) that plaintiffs would receive all documents in the seed set; (4) that the documents hit by targeted search terms used to populate a portion of the seed set would be disclosed to plaintiffs along with the defendants coding of those documents; (5) that plaintiffs would provide additional search terms to further populate the seed set; (6) that all documents reviewed in the seed set would be turned over to plaintiffs, whether designated as responsive or not; and (7) that after the seventh round, defendants would review a random sample of documents the algorithm deemed to be non-responsive in order to assess recall, and disclose those documents to plaintiffs. The court concluded that resolution of subsequent disputes over the reliability and success of the methodology was premature. Notably, Judge Peck discounted concerns about relevance standards given that the level of transparency would permit plaintiffs to raise concerns with the court as they arose. However, proceeding with the use of TAR without reaching agreement on such a fundamental issue could result in costly, prolonged motion practice that could undercut the savings gained from the use of TAR. Furthermore, requiring the production of non-responsive documents is facially inconsistent with the language of Rule 26(b) that requires, with limited exception, the production of only those documents "relevant to any party's claim or defense."[33]

Transparency was also a critical feature in *Actos*, where the court entered a case management order outlining a "search methodology proof of concept" that provided a set of detailed, step-by-step requirements governing how a specific, identified TAR product would be applied to a subset of a collection to test whether the tool was appropriate for the full ESI collection. [34] The Order included not only disclosure requirements, but obligations for substantial cooperation: (1) the parties agreed to the custodians from whom documents would be collected for the sample collection population; (2) the court set the size of the random "control set;" (3) defendants' experts reviewed documents for privilege first, providing a privilege log; (4) the parties each nominated three experts to review the control set; (5) following review of the control set, using an active-learning approach, the experts coded random sample training sets, permitting the experts to work collaboratively to determine responsiveness, subject to meet and confers when disagreements about coding designations arose; (6) the court required sufficient training rounds until the system stabilized; and (7) the parties are required to meet and confer to agree upon the relevance score above which documents were produced, and below which they were withheld. Notably, the clear difficulty of repeatedly obtaining agreement on relevance

---

[32] For a discussion of why the subsequent disputes over the predictive coding process in *Da Silva Moore* may have been much ado about nothing, *see* Herbert L. Roitblat, *Da Silva Moore Plaintiffs Slash and Burn Their Way Though eDiscovery*, Information Discovery Blog, Mar. 25, 2013, *available at* http://orcatec.blogspot.com/2012/03/da-silva-moore-plaintiffs-slash-and.html.

[33] Fed. R. Civ. P. 26(b).

[34] Case Management Order: Protocol Relating to the Production of Electronically Stored Information ("ESI") (ECF No. 1539) at 6-16, *In re Actos (Pioglitazone)* Prods. Liab. Litig., No. 11-md-2299 (W.D. La. July 27, 2012).

from six experts, may result in a longer and more contentious discovery process than the parties envisioned.

The levels of transparency in *Da Silva Moore* and *Actos* arguably provided the requesting party with the information and involvement in the document retrieval process necessary to assess whether the process and methodology was sufficiently reliable and thus defensible. Notably, in both cases, the requesting party enjoyed full disclosure about the technology itself, and had full access not only to the sample sets themselves (with the exception of privileged documents) but also to the coding of those sets, permitting them to determine the adequacy of the critical control point of TAR—the human teaching of the tool.

Why are transparency and cooperation deemed so critical to adoption of TAR? *Biomet* notwithstanding, requesting parties and courts increasingly understand that not all TAR tools are alike or equally effective and that none are more reliable than traditional legal search methods if they are not used and applied correctly, with a clearly defined process and sound statistical methodology supporting them. Too much can go wrong: a seed set can be improperly selected, improper judgmental sampling can bias the learning process, coding of the seed set may be inaccurate, the learning process may be terminated before it has stabilized, inappropriate relevance thresholds may be adopted, and so forth.

Transparency and cooperation at each stage of a TAR process are thus useful to assure the requesting party that the tool as applied is effective. Without those assurances, merely invoking the magic words "predictive coding"—despite its promise of reliability and efficiency—will not guarantee that either requesting parties or courts will endorse the process.

Much has been written about the steps required to achieve a defensible e-discovery process—a process that reflects the producing party's compliance with its discovery obligations to undertake reasonable efforts to identify and produce responsive documents.[35] One way to reduce the susceptibility of the chosen review methodology and process to *post-hoc* challenges is for the parties to reach advance agreement not only on the methodology but on how it will be applied.[36]

III. Objections to Transparency: Privilege Doctrines & Litigation Realities

Nonetheless, producing parties are often reluctant to disclose even the most basic information about their search methodologies. The aversion to disclosure of TAR processes is often founded on assertions that (1) revealing methodological decisions reveals protected work-product; (2) discovery about discovery is beyond the scope of Rule 26 of the Federal Rules of Civil Procedure; (3) revealing documents that are non-responsive exposes the producing party to

[35] *See, e.g.,* Hon. Craig B. Schaffer, M.J., "*Defensible" By What Standard?*, 13 Sedona Conf. J. 217 (Fall 2012); Mike Hamilton, *"Defensible E-Discovery": What does the phrase really mean?,* E-Discovery Beat, June 28, 2012, *available at http://www.exterro.com/e-discovery-beat/2012/06/28/%E2%80%9Cdefensible-e-discovery%E2%80%9D-what-does-the-phrase-really-mean/*

[36] *Id.* ("The most 'defensible' search methodology is one [that] has been jointly adopted by the parties and endorsed by the court.").

unnecessary litigation risk; and (4) producing parties are bound by the Federal Rules requiring that they conduct a reasonable search for responsive documents, and thus no greater assurances are required or warranted.

The work-product doctrine protects from discovery documents or tangible things that express the mental impressions of an attorney about a particular matter.[37] An attorney is thus protected from disclosing his or her written thoughts about relevant vs. irrelevant facts, legal theories and strategies.[38] Likewise, the attorney-client privilege protects from discovery communications between an attorney and his or her client if made for the purpose of providing or obtaining legal advice.[39]

TAR advocates should expect that assertions of work-product that have plagued traditional keyword search negotiations will, if asserted as to TAR methodologies, impede adoption and acceptance of technology assisted review. Even under a keyword search strategy, some defense attorneys assert that the search terms themselves and related strategies for searching are work product because they "reflect counsel's strategy for the case" and thus courts cannot force, and propounding parties cannot insist on, their disclosure.[40] Additionally, it is not uncommon for a producing party to refuse to disclose their precision and recall statistics on grounds of work product.

In the context of TAR, counsel raise similar concerns that work product may be impermissibly disclosed if producing parties are required, for example, to reveal the seed set, particularly if judgmental sampling is used, because it may reveal thought processes about the case. If search terms are used to identify documents used to populate the seed set, as with traditional use of keyword searching, attorneys express concern that revealing those terms discloses their assessments about the case. Some suggest that even disclosure of a sources used to generate random seed set reveals the attorney's strategy on sources to be searched and

---

[37] *Hickman v. Taylor*, 329 U.S. 495, 509 (1947); Paul F. Rothstein & Susan W. Crump, FEDERAL TESTIMONIAL PRIVILEGES §11:1 (2nd ed. 2007-08). *See also* Fed. R. Civ. P. 26(b)(3).

[38] *Hickman*, 329 U.S. at 510-11.

[39] FEDERAL TESTIMONIAL PRIVILEGES, *supra* n.34, at §2:10

[40] *See, e.g.*,Kessler, *supra* n. 27 ("What should not become common is the forced disclosure of search terms by courts. Not only are search terms not within the bounds of discovery, but are windows into how counsel are considering and evaluating their case. . . . [T]he search terms a party applies in discovery are generally created with the advice of counsel. Thus they represent the work product and mental impressions of the party's counsel."); *Graff v. Haverhill N. Coke Co.*, No. 1:09-cv-670, 2010 U.S. Dist. LEXIS 142925, at *17 (S.D. Ohio Dec. 22, 2010) (noting "Defendants claim that the search terms and methodology used to obtain all relevant documents from the company's electronic database of documents is protected by the work-product doctrine.").

Some courts have concluded, however, that details about search methodology are not work product at all, since they merely go to the "underlying facts of what documents are responsive to . . . document requests and [do] not delve into the thought processes of . . . counsel." *Romero v. Allstate Ins. Co.*, 271 F.R.D. 96, 110 (E.D. Pa. 2010). Rather than reveal an attorney's assessment of what is relevant or even important to the case, disclosures regarding search methodologies merely reveal assessments of what documents are responsive to production requests—a purely factual question. *Id.* at 110 n.9

custodians likely to possess responsive documents.[41] Producing parties' aversion to disclosing reliability statistics for search terms will apply equally to determinations of relevance thresholds and F_1 statistics[42] in the context of TAR. Finally, if, as in *Da Silva Moore* and *Actos*, the producing party must not only reveal the seed set, but also the coding decisions of reviewing attorneys, expect claims that doing so requires disclosure of attorneys' work product—their judgments as to relevance.

More serious concerns arise with respect to disclosure of documents in the seed set coded as non-responsive by producing attorneys. Doing so reveals to the requesting parties documents that are not subject to any production obligation. That might be beyond the scope of discovery obligations, which require production of relevant information.[43] Moreover, producing parties that are the target of repeat litigation by the same set of plaintiffs' attorneys may well be reluctant to turn over documents entirely unrelated to the matter at hand, such as in prescription drug mass torts. [44]

Finally, some contend that courts and requesting parties should need no greater assurance as to the appropriateness of TAR methodologies than the producing party's obligation to conduct a reasonable search under the federal rules and their certification that they have done so,[45] arguing courts and requesting parties have long had to rely for assurance simply on this certification as to the reasonableness of the investigation conducted to identify responsive documents. Further, they assert that the producing party has a strong incentive to implement an accurate and effective methodology, given the risk that opposing counsel will discover through depositions or other means responsive documents that were withheld. All of this, though, may be insufficient to allay the concerns of requesting parties who may acknowledge the theoretical superiority of TAR, but understand that given its complexity, even attorneys acting in good faith can improperly implement an otherwise effective search tool, excluding from production large amounts of responsive documents.

But calls for transparency and cooperation are not without support. In 2008, The Sedona Conference® issued a *Cooperation Proclamation*, which "urges parties to work in a cooperative rather than an adversarial manner to resolve discovery issues in order to stem the monetary costs

---

[41] *See* Fowler, *supra* n.9.

[42] The F_1 statistic is "a measure of the effectiveness of a search or review effort, which accounts for the tradeoff between Recall and Precision. In order to achieve a high F_1 score, a search or review effort must achieve both high Recall and high Precision." Maura R. Grossman & Gordon V. Cormack, *The Grossman-Cormack Glossary of Technology-Assisted Review*, 7 Fed. Ct. L. Rev. 1, 16 (2013).

[43] *See* Fed. R. Civ. P. 26(b)(1) ("Unless otherwise limited by court order, the scope of discovery is as follows: Parties may obtain discovery regarding any nonprivileged matter that is relevant to any party's claim or defense—including the existence, description, nature, custody, condition, and location of any documents or other tangible things and the identity and location of person who know of any discoverable matter. . . .").

[44] *See* Fowler, *supra* n.9.

[45] *See* Fed. R. Civ. P. 26(g).

of discovery disputes."[46] Numerous courts have now cited to the Proclamation, some noting that "the best solution in the entire area of electronic discovery is cooperation among counsel."[47]

Even if parties are inclined to provide the transparency and cooperation envisioned in *Da Silva Moore* and *Actos,* the cost of doing so might outweigh any savings achieved by employing TAR. In *Da Silva Moore,* even though the parties agreed on the use of TAR, the process of reaching agreement on the specifics of the protocol to be employed, required days of hearings and testimony from competing experts. In *Actos,* where the parties agreed on the use of TAR and a joint protocol for its use, the costs associated with having six experts review the control set and training sets, and the costs resulting from the need to resolve their disagreements whether by negotiation or by seeking a judicial ruling, could significantly reduce the savings that the parties expected to gain from employing TAR.

Regardless of the relative merits of the arguments against transparency and cooperation in designing and implementing TAR, there is reason for concern that mandating the disclosures and the level of cooperation present in *Da Silva Moore* and *Actos* will have a chilling effect, steering producing parties back to less robust methodologies where less must be revealed.[48] As a consequence, those advocating for TAR as a viable and preferable legal search tool in appropriate cases face a difficult task: they must find a way to balance the conflicting concerns of producing parties resisting disclosure with requesting parties demanding it.

IV. Transparency and Cooperation

For all the reasons discussed, enthusiasm over the promise of TAR should be tempered by a strong dose of litigation reality; its widespread adoption likely depends as much on how and whether the parties agree to cooperate and what the producing party discloses as on the technology itself and judicial acceptance of the idea of predictive coding. We offer here a checklist of the information that the parties should consider disclosing to facilitate a cooperative and defensible information retrieval process using TAR. To information retrieval scientists, the importance of these disclosures may be clear. However, as discussed above, in the world of litigation where parties have traditionally not been required to disclose such information and have reasonable concerns about the appropriateness of such disclosures, the disclosures are controversial.

The Tool & the Provider: Not all TAR tools are the same. Knowing what tool from which vendor is proposed aids understanding of the strengths and weaknesses of the tool. Does it use feature engineering, such as latent semantic analysis, to prepare the feature set for

---

[46] The Sedona Conference, *The Case for Cooperation*, 10 SEDONA CONF. J. 339, 348-49 (2009).

[47] *SEC v. Collins & Aikman Corp.*, 256 F.R.D. 403, 406-07 (S.D.N.Y. 2009); *see also, e.g., Kleen Prods.*, 2012 U.S. Dist. LEXIS 139632, at *6; Cartel Asset Mgmt. v. Ocwen Fin. Corp., Civil Action No. 01-cv-01644-REB-CBS, 2010 U.S. Dist. LEXIS 17857, at *40 (D. Colo. Feb. 8, 2010); *DeGeer v. Gillis*, 755 F. Supp. 2d 909, 918 (N.D. Ill. 2010); *William A. Gross Constr. Assocs. v. Am. Mfrs. Mut. Ins. Co.*, 256 F.R.D. 134, 136 (S.D.N.Y. 2009)*; Capitol Records, Inc., v. MP3tunes, LLC*, 261 F.R.D. 44, 47 (S.D.N.Y. 2009).

[48] *See* Fowler, *supra* n.9.

classification? What classifier design/machine learning does the tool use? Nearest neighbor? Language modeling? Bayesian?[49] What methodologies are used to categorize documents? Supervised learning? Active learning? Both? Some experts suggest that any tool should include an active learning component,[50] a process in which the system identifies documents most likely to be misclassified.

The more information disclosed about the tool, the more comfortable a requesting party can be in the appropriateness of that tool for the task. Importantly, disclosure of these facts should provide producing parties with no discomfort.

The Nature of the Documents and the Size of the Corpus: What types of documents will be included in the corpus from which samples will be drawn? Email and text documents? Paper scans with unsearchable text? Image files? Spreadsheets with limited text? If the underlying tool relies on semantic analysis, these disclosures inform the appropriateness of the tool's application.

And what is the overall volume? These disclosures aid in understanding whether certain documents are appropriate for machine learning or should be excluded and searched using an alternative methodology.

Culling the Corpus: Will culling techniques be used? If so, which and how? Will date, file type, domain or custodian filters be applied? Disclosure of date filtering should not be controversial as the relevant date range is one likely to be resolved through negotiations over the production requests themselves. Similarly, disclosures of associated custodians and sources of ESI are routine in modern litigation, and often mandated at the outset of litigation in a preservation order. Disclosing that certain data stores will be included or excluded should not therefore trouble the producing party, and indeed, physical and human sources of discoverable documents are among the type of discoverable information under Rule 26(b)(1).

If keywords are used to cull the corpus, what are the search terms? As with traditional keyword searching, producing parties may object to revealing search terms, but if used to cull documents, their disclosure will be important not only to gain the requesting party's agreement, but also to defend the methodology if it is later challenged. Concerns about revealing work product or confidential information in disclosing search terms can be addressed through protective orders. One such order, in the context of keyword searching, provided that such

---

[49] For a helpful discussion on the different features of search protocols (feature construction, feature transformation, classifier design, and methods), as well as types of legal search algorithms, *see* Ralph Losey, The Many Types of Legal Search Software in the CAR Market Today, E-discovery Team Blog, Mar. 3, 2013, *available at* http://e-discoveryteam.com/2013/03/03/the-many-types-of-legal-search-software-in-the-car-market-today/ (crediting Doug Oard for the description of the layers of search algorithms, and Herb Roitblat for the description of search algorithms). Losey advises that good TAR design requires all four features of search protocols must work well together.

[50] *See* Losey, *supra* n.48 ("If software does not have [active learning] built in as a planned method of use, then, in my opinion, it is not *bona fide* predictive coding, machine learning, type of CAR (TAR).")

disclosure does not constitute waiver of work-product protection or attorney client privilege, the terms will be treated as highly confidential information, and if necessary to file search terms with the court, they shall be filed under seal.[51] And under Federal Rule of Evidence 502, a federal court may order that privilege or work-product protection is not waived by disclosure connected to the litigation, and the parties' agreement on the effect of disclosure is binding.[52]

Control Sets, Sampling Techniques to Seed the Training Set: Does the system provide for a control set against which the tool's performance can be judged (*i.e.*, to assess the prevalence of responsive documents in the corpus)? If so, how large is the control set and what confidence level and interval was used to derive it?

How will documents be selected for the training set? Purely randomly machine-selected? Through the use of judgmental sampling? Or both? If judgmental sampling is used, what tool is being used to select that sample? If keywords are used to do so, what are they? As noted, disclosure of keywords can facilitate significant agreement between the parties.

What is the size of the training set, and how was that sample size established? What confidence level and interval was selected? Was the sample pulled from the entire corpus of documents to be predictively coded?

Workflow Considerations: All TAR tools use some procedure for iterative refinements after the initial training session, whether done sequentially or automatically. What approach does the tool use? How many training rounds are generally necessary to reach stability? If the tool uses active learning, how many documents in total must generally be reviewed before stability is achieved?

Coding the Training Set: Who will be coding the training sets? Experienced attorneys? Relatively junior attorneys with less litigation experience? Contract attorneys less familiar with the case? How many coders will be training the system?

With TAR,the effectiveness of the methodology is only as good as the coding used to train the system. Requesting parties will want assurances that experienced attorneys that are intimately familiar with the issues in the case are training the tool and that as few coders as possible are used to minimize the inconsistent relevance determinations that are inherent in human review.

As in *Da Silva Moore* and *Actos*, requesting parties may also want not only to know how documents were coded during the training rounds, but also to play an active role in that coding. As noted, this degree of disclosure and cooperation justifiably generates significant concern among producing parties about not only disclosure of work product but also of non-responsive documents outside of the scope of discovery.

---

[51] *See In re CV Therapeutics, Inc. Sec. Litig.*, 2006 U.S. Dist. LEXIS 38909, at *31-32 (N.D. Cal. Apr. 3, 2006).

[52] Fed. R. Evid. 502(d) & (e).

However, there are tools to minimize those concerns. First, as in *Actos*, the producing party may first cull out potentially privileged documents from the training set. Second, to address both the producing party's concern that non-responsive documents will be disclosed and work product revealed, and the requesting party's concern that documents will be improperly coded, parties can consider employing a third party neutral, schooled in the facts of the case, to either conduct the training herself, or to evaluate the accuracy of the producing party's coding.

Relevance Thresholds and Reliability Statistics: What confidence level and interval will be used to determine sample sizes and precision and recall? What level of precision and recall will be deemed acceptable. How will the responsiveness cutoff be determined? This determination is necessarily dependant on determinations of acceptable precision and recall rates (or F_1 statistics). The producing party will want high precision and possibly high recall (depending on its risk aversion); the requesting party will want both. Reaching agreement on the appropriate balance inures to the benefit of the producing party should the production later be challenged.

V. Conclusion

Although information retrieval scientists have developed wonderful and exciting tools that can provide improved accuracy while reducing the cost of document review, practical considerations specific to litigation have resulted in less widespread adoption than would have been expected in the abstract. The processes employed to maximize the efficacy of these tools must be adapted for use by lawyers and their clients, who face limits on the degree of transparency and must decide how much machine training is enough to achieve defensibility rather than perfection. As lawyers and judges become better educated about the processes needed to employ TAR effectively, agreement about the specifics of TAR protocols should become easier to achieve. This should reduce the costs associated with reaching agreement on TAR protocols and encourage wider adoption of TAR in civil discovery.