

NY Commercial Division Backs Technology-Assisted Review

By **Elizabeth Sacksteder and Ross Gotler** (July 29, 2018, 9:26 PM EDT)

It is generally agreed that the most expensive stage of complex commercial litigation today is document review. A 2012 RAND Corp. study found that document review consumes on average 73 percent of the total cost of document production in cases involving electronic discovery, notwithstanding such common economies as the use of vendors to do first-level document review.[1] Achieving greater efficiency in this resource-intensive stage of litigation — making review of electronically stored information cheaper, faster and more accurate — is a shared goal of litigants, their counsel, and the courts. Sophisticated litigants know that the use of technology-assisted review — of which there are many types, ranging from widely used software tools like keyword searching to more sophisticated algorithmic technologies such as predictive coding — can yield substantial cost savings, as well as streamline and accelerate document review and production. But neither the Federal Rules of Civil Procedure nor most state procedure codes, including the New York Civil Practice Law and Rules, or CPLR, expressly address whether, in what circumstances, or how a party may use technology-assisted review to fulfill its disclosure obligations.

A new rule in the Commercial Division of the Supreme Court of the State of New York aims to fill that gap. Promulgated on July 19, 2018, by Chief Administrative Judge Lawrence K. Marks as a new subdivision (f) of Rule 11-e of the Rules of Practice for the Commercial Division, and effective on Oct. 1, 2018, the new rule provides:

(f) The parties are encouraged to use the most efficient means to review documents, including electronically stored information (“ESI”), that is consistent with the parties’ disclosure obligations under Article 31 of the CPLR and proportional to the needs of the case. Such means may include technology-assisted review, including predictive coding, in appropriate cases. The parties are encouraged to confer, at the outset of discovery and as needed throughout the discovery period, about technology-assisted review mechanisms they intend to use in document review and production.

The new rule serves several purposes. First, it confirms, consistent with the Commercial Division’s goal to “conserve client resources, encourage proportionality in discovery, [and] promote efficient resolution of matters,”[2] that efficiency is not only an appropriate consideration in determining how to review ESI, but that the most efficient means to review documents is actually preferred, provided that the method



Elizabeth Sacksteder



Ross Gotler

chosen by the producing party meets that party's disclosure obligations. Having this confirmation in a formal rule should provide comfort to litigants and their counsel that the court will help them avoid needless waste of resources in reviewing the large volumes of ESI that are typical in complex commercial cases.

Second, the new rule makes clear that the use of technology-assisted review, including predictive coding and other algorithmic, machine-learning-based technologies, is not only appropriate but encouraged when it is the most efficient means for a party to meet its disclosure obligations. In this respect, the Commercial Division has placed itself in the vanguard of promoting the thoughtful use of technology and analytics to ease the burden of document review on litigants.

Third, the new rule makes clear that, as with other aspects of the discovery process, parties are expected to confer with one another about the technology-assisted mechanisms they intend to use in document review and production. This mandate aligns with the existing Commercial Division requirement that parties consult prior to preliminary or compliance conferences with the court, including about "the scope or method for searching and reviewing ESI."^[3]

The new rule was developed and recommended to the chief administrative judge by the Commercial Division Advisory Council, which was established in 2013 by then-Chief Judge Jonathan Lippman to "help fulfill the long-term strategic goals of a world-class Commercial Division in New York State."^[4] In advancing the rule, which was adopted after a public comment period, the council sought not only to provide guidance to commercial litigants in the nation's leading commercial center about the appropriate use of available tools to make document review cheaper, faster and more accurate, but also to highlight the sophistication of the justices of the Commercial Division and their desire to be on the leading edge of the smart use of technology in complex litigation.

Technology-Assisted Review

Litigants in complex commercial cases today use a wide range of technology-assisted review techniques to facilitate the review of what is often an enormous volume of ESI. In such document-intensive cases, human review of each and every collected document for responsiveness can be slower, more costly, and less accurate than the appropriate use of technology-assisted review,^[5] which relies on software to help cull irrelevant documents from a large data set, to group together similar documents so as to promote efficient review and consistency of results, or to "teach" a computer to recognize those documents that are most likely to be responsive.

The threshold challenge faced in reviewing a large volume of ESI is that most ESI is unstructured, meaning that it is not organized in any predetermined way — a common example being email, which has few predetermined data fields and typically is stored without regard to subject matter. Traditionally, review of ESI begins by collecting a large volume of unstructured ESI, frequently limited only by custodian and date range, and then running a keyword search, which uses software to identify words or phrases that are likely to be found in responsive documents, to identify the documents to be reviewed.^[6] A more sophisticated variant is "concept searching," which uses advanced technology to identify documents incorporating concepts similar to the specific search terms used.^[7]

The efficiency of the document review can be enhanced through techniques to group similar or related documents together, such as "email threading," which packages together email strings and any attachments as one chronological thread;^[8] "near-duplicate identification," which groups together similar documents based on their textual similarities (e.g., different drafts of a document);^[9] and

"clustering," which uses conceptual analytics technology to group and categorize similar documents.[10]

While these techniques can help to cull a data set and organize it for review, none of them obviates the need for human review for responsiveness. The form of technology-assisted review generally referred to as predictive coding brings practitioners closer to just that. Predictive coding uses a machine learning algorithm[11] to extrapolate human judgments about responsiveness, based on human review of a sample "seed set" or "training set" of documents, across the remaining document collection.[12] Because predictive coding requires an upfront investment of time in "teaching" the computer to recognize the characteristics of responsive documents, often iteratively, it generally is cost-effective only when dealing with a large volume of unstructured ESI, but in those circumstances it has the potential to enhance the speed, accuracy and cost-effectiveness of document review.[13] Recent advances in predictive coding technology have reduced the reliance on training sets and iterations, however, making predictive coding a viable option in many more situations.

Rationale for the New Rule

Both federal and state courts have endorsed the use of technology-assisted review. The United States District Court for the Southern District of New York, for example, has noted that "[p]redictive 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." [14] Indeed, the Delaware Chancery Court has actually required a party to use predictive coding.[15] Courts have noted in particular the utility of predictive coding for reviewing a large volume of ESI. In the Southern District of New York, for example, Magistrate Judge Andrew Peck (retired) has observed that "computer-assisted review is an available tool and should be seriously considered for use in large-data-volume cases." [16] Another federal district court has granted a plaintiff's request, over the defendant's objection, to use predictive coding to review approximately 2 million documents for responsiveness.[17] Foreign courts have likewise recognized the utility of predictive coding in reviewing large volumes of ESI.[18]

The new rule aligns the Commercial Division with these decisions in supporting the use of technology-assisted review, including predictive coding, in appropriate cases. The rule does not, however, prescribe whether or when any particular form of technology-assisted review may or should be used. Because these technologies are evolving at a rapid rate, any effort to prescribe permissible or impermissible methodologies would quickly become obsolete, and in any event the appropriateness of a given methodology can only be determined in the context of the particular case and the data set to be reviewed. Nothing in the proposed rule is intended to limit the role of the presiding justice in supervising document disclosure,[19] or to insulate the responding party's production from challenge.[20]

Need for Proportionality

Regardless of the method a party uses to review a large collection of ESI for responsiveness, the result will not be perfect. "There simply is no review tool that guarantees perfection. ... [T]here are risks inherent in any method of reviewing electronic documents." [21] Courts have recognized that the standard for a review, whether technology-assisted or entirely human, "is not perfection, or using the 'best' tool, but whether the search results are reasonable and proportional." [22] "The goal is for the review method to result in higher recall and higher precision than another review method, at a cost proportionate to the 'value' of the case." [23]

This concept of proportionality is embedded in the Commercial Division rules[24] and the CPLR[25] as well as the Federal Rules.[26] Accordingly, it should not be a legitimate objection to a party's use of

predictive coding or other technology-assisted review that the chosen method may not deliver perfect results. If the methodology chosen is reasonable in the circumstances, then it should be deemed sufficient to meet a party's disclosure obligations. For that reason, the new rule incorporates proportionality as a relevant consideration in determining the appropriateness of a document review method.

Parties Encouraged to Cooperate

Because the responding party knows best what kinds and volume of documents it has, how they are stored, and what it will cost to review them, "[r]esponding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information." [27] "Unless [the responding party's] choice is manifestly unreasonable or the requesting party demonstrates that the resulting production is deficient, the court should play no role in dictating the design of the search ..." [28] Nonetheless, disputes can be avoided if the parties confer at the outset and as needed throughout the discovery process about any technology-assisted review techniques they propose to use, and the rule expressly encourages such cooperation. Should the parties be unable to agree, however, the rule does not prevent the requesting party from challenging the producing party's chosen means as inadequate or a production as incomplete.

* * *

As society becomes ever more reliant on electronic communications of all kinds, and advanced technologies such as machine learning and artificial intelligence become ever more sophisticated, technology-assisted review will become an increasingly important tool in commercial litigators' toolkit. The Commercial Division's new Rule 11-e(f) embraces the potential of technology to make disclosure quicker, cheaper and better in the information-rich world in which we all live today.

Elizabeth M. Sacksteder is a partner at Paul Weiss Rifkind Wharton & Garrison LLP and a member of the Commercial Division Advisory Council. Ross M. Gotler is e-discovery counsel at Paul Weiss.

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

[1] Nicholas M. Page & Laura Zakaras, Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery, RAND Institute for Civil Justice, 2012, at xv-xvi, 25-27, 41, available at https://www.rand.org/content/dam/rand/pubs/monographs/2012/RAND_MG1208.pdf.

[2] 22 NYCRR § 202.70(g), Preamble; N.Y. Ct. Rules § 202.70(g) (Uniform Civil Rules for the Supreme Court and County Court).

[3] *Id.* at Rule 8(b)(ix).

[4] Commercial Division of the Supreme Court of the State of New York, Press Release, Chief Judge Names Members of Commercial Division Advisory Council, Mar. 26, 2013, available at http://ww2.nycourts.gov/sites/default/files/document/files/2018-05/PR13_05.pdf.

[5] See, e.g., Pace & Zakaras at 55-58, 61-69.

[6] See, e.g., The Sedona Conference, Commentary on Defense of Process: Principles and Guidelines for Developing and Implementing a Sound E-Discovery Process, 25 (Public Comment Version, 2016).

[7] See, e.g., Relativity, Concept Searching, available at <https://www.relativity.com/relativity/Portals/0/Documents/8.0%20Documentation%20Help%20Site/Content/Features/Analytics/Concept%20searching.htm>.

[8] See Nik Balepur, 5 Email Threading Facts That May Surprise You, The Relativity Blog (Apr. 16, 2015), available at <http://blog.kcura.com/relativity/blog/5-email-threading-facts-that-may-surprise-you>.

[9] See D4, Near-Duplicate Detection Finds Documents No One Thought Could Be Found, D4 Case Studies Blog (June 11, 2015), available at <http://d4discovery.com/discover-more/near-duplicate-detection-finds-documents-no-one-thought-could-be-found#sthash.tl5DevpH.dpbs>; see also Equivio, Choosing a Near-Duplicate Identification Solution (2012), available at <http://www.equivio.com/files/files/White%20Paper%20-%20Choosing%20A%20Near-Duplicate%20Identification%20Solution.pdf>.

[10] See CloudNine, Document Clustering for eDiscovery Review, available at <https://www.ediscovery.co/legacy/document-clustering>.

[11] See Maura R. Grossman & Gordon V. Cormack, The Grossman-Cormack Glossary of Technology-Assisted Review, with Foreword by John M. Facciola, U.S. Magistrate Judge, 2013 Fed. Cts. L. Rev. 7, 26 (2013).

[12] See *id.* at 29, 32-33. Other implementations of predictive coding use a "Continuous Active Learning" model in which the computer "learns" while humans review documents, allowing for the reclassification of documents as the software continuously evolves its "understanding."

[13] See John Tredennick et. al, TAR for Smart People: How Technology Assisted Review Works and Why IT Matters for Legal Professionals, 35-41 (2016), available at https://catalystsecure.com/pdfs/book/Catalyst_TAR_for_Smart_People.pdf.

[14] *Chevron Corp. v. Donziger*, 2013 WL 1087236, at *32 n.255 (S.D.N.Y. Mar. 15, 2013).

[15] *OSI Restaurant Partners LLC v. United Ohana LLC*, 2017 WL 396357 (Del. Ch. Ct. Jan. 27, 2017).

[16] *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182, 193 (S.D.N.Y. 2012).

[17] See *Bridgestone Ams, Inc. v. Int'l Bus. Machs. Corp.*, 2014 WL 4923014, at *1 (M.D. Tenn. 2014).

[18] See, e.g., *Irish Bank Resolution Corp. Ltd & ors v. Quinn & ors*, [2015] IEHC 175 (Ir.); *Brown v. BCA Trading Ltd.*, [2016] EWHC 1464 (Ch) (Eng.); *Pyrrho Invs. Ltd. v. MWB Prop. Ltd.*, [2016] EWHC 256 (Ch) (Eng.).

[19] See CPLR 3104(a).

[20] See CPLR 3124.

[21] Da Silva Moore v. Publicis Groupe, 2012 WL 1446534, at *3 (S.D.N.Y. Apr. 26, 2012) (affirming Magistrate Judge Peck's acceptance of predictive coding).

[22] Hyles v. N.Y. City, 2016 WL 4077114, at *3 (S.D.N.Y. Aug. 1, 2016); see Winfield v. City of New York, 2017 WL 5664852, at *11 (S.D.N.Y. Nov. 27, 2017).

[23] Da Silva Moore, 287 F.R.D. at 190.

[24] See 22 NYCRR § 202.70(g), Preamble ("The Commercial Division is mindful of the need to ... encourage proportionality in discovery[.]").

[25] See CPLR 3101(a) (limiting the scope of disclosure to "all matter *material and necessary* in the prosecution or defense of an action") (emphasis added).

[26] See Fed. R. Civ. P. 26(b)(1) (limiting the "scope of discovery" to "any nonprivileged matter that is relevant to any party's claim or defense *and proportional to the needs of the case*, considering the importance of the issues at stake in the action, the amount in controversy, the parties' relative access to relevant information, the parties' resources, the importance of discovery in resolving the issues, *and whether the burden or expense of the proposed discovery outweighs its likely benefit*") (emphases added).

[27] Hyles, 2016 WL 4077114, at *3 (quoting The Sedona Conference, The Sedona Principles: Second Edition, Best Practices Recommendations & Principles for Addressing Electronic Document Production, Principle 6, Ill. i (2d ed. 2007), available at www.thesedonaconference.org); see also Winfield, 2017 WL 5664852, at *9.

[28] Mortg. Resolution Servicing Inc. v. JPMorgan Chase Bank NA, 2017 WL 2305398, at *2 (S.D.N.Y. May 18, 2017).