

Guidelines Regarding the Use of Predictive Coding

Like any new and disruptive technology, the use of predictive coding¹ has been both contagious and controversial since its introduction to the discovery process.² On the one hand, predictive coding has found welcome recipients in clients, counsel, and the courts, all of whom are seeking to expedite the ESI search and review process.³ Lawyers and litigants have additionally gravitated toward predictive coding given its utility in identifying the key documents required to establish their claims or defenses.⁴

Nevertheless, there has been disagreement over the “what,” “when,” and “how” of predictive coding⁵ and misconceptions persist regarding what is predictive coding, when it should be used, and the process for how to successfully implement it into a discovery workflow.⁶ Moreover, the few judicial opinions on predictive coding are based on specific fact patterns that make general application for practitioners difficult.⁷

In an effort to dispel confusion over these issues and to help provide informed direction on the use of predictive coding, the Coalition of Technology Resources for Lawyers has prepared a few guidelines that it believes will aid attorneys, clients, and judges regarding the use of predictive coding. Developed so as

¹ In these Guidelines, predictive coding shall mean and refer to a process for selecting and ranking a collection of documents using a computerized system that incorporates the decisions that lawyers have made on a smaller set of documents and then applies those decisions to the remaining universe of documents. *See* Bennett Borden & Jason R. Baron, *Finding the Signal in the Noise: Information Governance, Analytics, and the Future of Legal Practice*, 20 RICH. J.L. & TECH. 7, ¶ 10 (2014); Maura R. Grossman & Gordon V. Cormack, *The Grossman-Cormack Glossary of Technology-Assisted Review*, with a Foreword by John M. Facciola, U.S. Magistrate Judge, 2013 FED. CTS. L. REV. 7, 38 (2013).

² *See, e.g.*, *In re Domestic Drywall Antitrust Litig.*, 88 Fed. R. Serv. 3d 966 (E.D. Pa. May 12, 2014) (touting the benefits of predictive coding as a new and “sophisticated” discovery search methodology); Dana Remus, *The Uncertain Promise of Predictive Coding*, 99 IOWA. L. REV. 1691, 1695 (2014) (suggesting that lawyers “proceed with deliberate care in the use and adoption of predictive-coding technologies.”).

³ *Nat’l Day Laborer Org. Network v. U.S. Immigration & Customs Enforcement Agency (NDLON)*, 877 F. Supp. 2d 87, 109 (S.D.N.Y. 2012) (“[P]arties can (and frequently should) rely on . . . machine learning tools to find responsive documents.”); Hon. Patrick J. Walsh, *Rethinking Civil Litigation in Federal District Court*, 40 LITIG. 6, 7 (2013) (urging lawyers to use “21st-century computer technology” including predictive coding to address digital age discovery issues and to stop relying on legacy discovery technologies and methods).

⁴ Walsh, *supra* note 3, at 7 (“Their goal should be to have the computer sift through the millions of documents and distill and organize the hundreds or thousands of documents that are critical to the case . . .”).

⁵ Remus, *supra* note 2, at 116-17 (observing disapprovingly that “the litigation community is uncritically embracing predictive coding as if its definition is unitary and clear, its accuracy and efficacy well-established.”).

⁶ Borden, *supra* note 1, at ¶ 17 (“[W]e bow to the reality that in a large class of cases the use of predictive coding is currently infeasible or unwarranted, especially as a matter of cost.”).

⁷ *Compare* *Progressive Cas. Ins. Co. v. Delaney*, No. 2:11-cv-00678, 2014 WL 3563467, at * 10-12 (D. Nev. July 18, 2014) ECF No. 89 (interpreting case management order provision to prevent the use of predictive coding) *with* *Bridgestone Americas, Inc. v. Int’l Bus. Mach. Corp. (Bridgestone)*, No. 3:13-cv-1196 (M.D. Tenn. July 22, 2014) ECF No. 89 (allowing the use of predictive coding despite an arguably contrary case management order).

to apply irrespective of the predictive coding technology or workflow used, these general guidelines should help practitioners through the decision-making process on the following key issues:

- What are the touchstones of a defensible predictive coding use plan?
- What are some essential aspects of a defensible predictive coding workflow?
- Should the use of predictive coding be disclosed to litigation adversaries?
- What are the benefits and drawbacks of entering into a stipulated predictive coding use protocol?

Finally, a stipulation and order regarding the use of predictive coding is provided in the appendix that can be used as a model if required by a court or if the circumstances in a particular matter lend themselves to reaching such a stipulation with a litigation adversary. The model stipulation and order is provided with the caveat that we disclaim that obtaining such a stipulation and order is necessarily a best practice.⁸ For example, whether such a process will satisfy the mandate of Federal Rule of Civil Procedure 1 that “every action and proceeding” be resolved in “a just, speedy, and inexpensive” manner⁹ will depend upon the circumstances of each case, requiring the application of legal judgment.¹⁰

I. WHAT ARE THE TOUCHSTONES OF A DEFENSIBLE PREDICTIVE CODING USE PLAN?

Counsel must be prepared to defend the reasonableness of its use of predictive coding just as it would with any other discovery search methodology.¹¹ This means that counsel’s use of predictive coding must accord with the notions of relevance, proportionality, and reasonableness, the traditional touchstones of the discovery process.¹² These points, along with the general need to establish a defensible predictive

⁸ See *infra* Part IV.

⁹ FED. R. CIV. P. 1.

¹⁰ See *In re Biomet M2a Magnum Hip Implant Products Liability Litig. (Biomet I)*, No. 3:12-MD-2391, 2013 WL 1729682 (N.D. Ind. Apr. 18, 2013) (holding that the defendant’s combined use of keyword and predictive coding search methodologies satisfied its response obligations under Rule 26 and Rule 34); Remus, *supra* note 2, at 126 (cautioning lawyers not to yield the interests of their clients in the quest for reasonable cooperation in the discovery process).

¹¹ *Bridgestone* (successfully defending the combined use of keywords and predictive coding). See also William A. Gross Const. Associates, Inc. v. Am. Mfrs. Mut. Ins. Co., 256 F.R.D. 134, 134 (S.D.N.Y. 2009) (“This Opinion should serve as a wake-up call to the Bar in this District about the need for careful thought, quality control, testing, and cooperation with opposing counsel in designing search terms or ‘keywords’ to be used to produce emails or other electronically stored information”).

¹² See FED. R. CIV. P. 26(b)(1); 26(b)(2)(3); 26(g)(1); *Victor Stanley, Inc. v. Creative Pipe, Inc.*, 269 F.R.D. 497, 523 (D. Md. 2010) (observing that “all permissible discovery must be measured against the yardstick of proportionality.”). The recently proposed amendments to Federal Rule of Civil Procedure 26(b)(1) clarify that all discovery must be viewed through the lenses of relevance and proportionality. See JUDICIAL CONFERENCE OF THE UNITED STATES, REPORT OF ADVISORY COMMITTEE ON CIVIL RULES 84 (May 2, 2014) (REPORT) (“Proportional discovery relevant to any party’s claim or defense suffices . . .”). Satisfying these discovery touchstones has become more challenging in recent years with the advent of “eDiscovery.” eDiscovery requires more attention than ever before to understanding client data, deadlines, amounts in controversy, privacy protections, and the law of each case. If unfamiliar with eDiscovery, lawyers should either educate themselves or engage competent counsel to better satisfy procedural requirements and ethical duties. See STATE BAR OF CALIFORNIA STANDING COMM. ON PROF’L RESPONSIBILITY AND CONDUCT, PROPOSED FORMAL OP. INTERIM NO. 11-0004 (2014), *available at* <http://www.calbar.ca.gov/AboutUs/PublicComment/201404.aspx> (discussing a lawyer’s ethical duties in handling of

coding process, are illustrated in the *In re Biomet M2a Magnum Hip Implant Products Liability Litigation* case.¹³

In *Biomet*, the court refused the plaintiffs' request that the defendant pharmaceutical company re-do its production of documents.¹⁴ The plaintiffs had argued that the company's production of documents was incomplete given that the company had used a combination of keyword and predictive coding search methodologies to identify responsive information.¹⁵ The company first applied keyword searches and deduplication methods to reduce the universe of potentially responsive information from 19.5 million to 2.5 million documents.¹⁶ It then searched the remaining subset of data using a predictive coding process. Relying on scholarly research and statistical reports, the plaintiffs challenged the ability of keyword searches to return an acceptable recall of responsive information.¹⁷ Because the recall of keyword searches was arguably too low and could leave out too much responsive data, the plaintiffs urged the company to re-do its production by running the predictive coding process against the original universe of 19.5 million documents.¹⁸

The court declined to order the re-production, holding instead that the company's production of documents satisfied its discovery obligations under Federal Rules of Civil Procedure 26(b) and 34(b)(2).¹⁹ Nothing in the Rules, observed the court, required the company to forego keyword searches.²⁰ Moreover, even if some marginally relevant information had been bypassed in connection with the keyword searches, re-doing the production at the anticipated seven-figure cost estimate would violate the proportionality standards set forth in Rule 26(b)(2)(C).²¹

Lastly, the court sympathized with but ultimately disregarded the plaintiffs' complaints that the company was not being cooperative.²² While touting the virtues of cooperation enshrined in the Seventh Circuit

eDiscovery); MODEL RULES OF PROF'L CONDUCT R. 1.1 cmt. 8 (2013) (competency requires an understanding of the "benefits and risks associated with relevant technology").

¹³ *Biomet I*.

¹⁴ *Id.* at *3.

¹⁵ *Id.* at *2.

¹⁶ *Id.* at *1.

¹⁷ *Id.*

¹⁸ *Id.* at *2.

¹⁹ *Id.* ("What Biomet has done complies fully with the requirements of Federal Rules of Civil Procedure 26(b) and 34(b)(2).").

²⁰ *Id.*

²¹ *Id.* at *3 ("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 the [plaintiffs] equals or outweighs its additional burden on, and additional expense to, Biomet.").

²² *Id.*

Pilot Program²³ and The Sedona Conference Cooperation Proclamation,²⁴ the court observed that 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.”²⁵ Moreover, even though the *Biomet* court later noted that the company’s lack of cooperation in certain aspects was “troubling” and could possibly lead the court to conclude the company was “hiding something,” the court took no action since it had no “authority to compel discovery of information not made discoverable by the Federal Rules.”²⁶

The *Biomet* case stands for the proposition that counsel can use predictive coding – with or without disclosing it to the other side – so long as the process is defensible.²⁷ Significantly, what is left unstated by the court’s order is an implicit understanding that the approach the defendant company adopted with respect to its predictive coding process was defensible, *i.e.*, that it reasonably identified highly relevant, responsive information and that it was proportional under the circumstances.²⁸ Were it not so and had the company not produced the key information in the case, the court undoubtedly would have required it to re-do the production despite the cost burdens.²⁹

The *Biomet* case ultimately teaches that a predictive coding process will be evaluated on the quality and nature of the responsive information disclosed to an adversary.³⁰ The extant predictive coding jurisprudence makes clear that courts do not require that counsel’s efforts meet a mythical standard of perfection.³¹ Instead, they expect the predictive coding process to result in document productions that are

²³ 7TH CIR. ELEC. DISCOVERY COMM., PRINCIPLES RELATING TO THE DISCOVERY OF ELECTRONICALLY STORED INFORMATION, at princs. 1.01-.03 (2010), available at http://www.discoverypilot.com/sites/default/files/Principles8_10.pdf.

²⁴ The Sedona Conference, *The Sedona Conference Cooperation Proclamation*, 10 SEDONA CONF. J. 331 (2009).

²⁵ *Biomet I*, at *2.

²⁶ *In re Biomet M2a Magnum Hip Implant Products Liability Litig. (Biomet II)*, No. 3:12-MD-2391, 2013 WL 6405156, at *2 (N.D. Ind. Aug. 21, 2013).

²⁷ Hon. Craig B. Shaffer, “*Defensible*” by *What Standard?*, 13 SEDONA CONF. J. 217 (2012).

²⁸ See *Biomet I*, at *2-3; *Bridgestone* (observing that “the exhortation of Rule 26 [is] that discovery be tailored by the court to be as efficient and cost-effective as possible.”). *But see* *Indep. Living Ctr. of S. Cal. v. City of L.A.*, 2:12-cv-00551-FMO-PJW, at *1 (C.D. Cal. June 13, 2014) ECF No. 371 (forbidding the defendant from using documents to train its predictive coding algorithm that it previously identified through keyword searches).

²⁹ See *Indep. Living Ctr.*, at *1.

³⁰ *Fed. Hous. Fin. Agency v. HSBC N. Am. Holdings Inc.*, 11-cv-6189, 2014 WL 584300, *2 (S.D.N.Y. Feb. 14, 2014); *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182, 191 (S.D.N.Y. 2012)

³¹ *Fed. Hous. Fin. Agency*, 2014 WL 584300 at *2 (“[While] [p]arties in litigation are required to be diligent and to act in good faith in producing documents in discovery . . . no one could or should expect perfection from this process. All that can be legitimately expected is a good faith, diligent commitment to produce all responsive documents uncovered when following the protocols to which the parties have agreed, or which a court has ordered.”); *Da Silva Moore*, 287 F.R.D. at 191 (“While this Court recognizes that computer-assisted review is not perfect, the Federal Rules of Civil Procedure do not require perfection.”).

reasonable and proportional under the circumstances, regardless of the level of cooperation between the parties.³²

II. WHAT ARE SOME ESSENTIAL ASPECTS OF A DEFENSIBLE PREDICTIVE CODING WORKFLOW?

There are various issues that must be considered in connection with a predictive coding workflow.³³ To be sure, those issues will inevitably vary depending on the quality and nature of the technology adopted by the responding party, along with the facts and circumstances of a particular matter.³⁴ For example, the time when counsel decides to apply predictive coding or other search methodologies to the universe of potentially responsive information – particularly when the search and review process begins before the Rule 26(f) conference – may impact workflow defensibility.³⁵

Irrespective of these and other issues, however, we refute any assertion that the ultimate defensibility of a predictive coding workflow is somehow dependent on the so-called “best practices” advanced by any particular eDiscovery technology provider.³⁶ Instead, the defensibility of such a workflow should be

³² See *Biomet I*, at *3; *Biomet II*, at *1-2.

³³ See Nicholas Barry, Note, *Man Versus Machine Review: The Showdown between Hordes of Discovery Lawyers and a Computer-Utilizing Predictive-Coding Technology*, 15 VAND. J. ENT. & TECH. L. 343, 354-55 (2013) (describing a predictive coding workflow and related issues associated with its implementation).

³⁴ The circumstances of a case may lend themselves to using other search methodologies either with or without predictive coding for a document production. Indeed, the nature of the data, text extraction or optical character recognition, images, and numerically based documents may very well impact whether a client should use predictive coding and/or other search methodologies in a particular workflow. See Elle Byram, *The Collision of the Courts and Predictive Coding: Defining Best Practices and Guidelines in Predictive Coding for Electronic Discovery*, 29 SANTA CLARA COMPUTER & HIGH TECH. L.J. 675, 694 (describing several issues and other variables affecting the choice of a particular discovery search methodology). These issues and others spotlight that predictive coding is not a stand-alone solution and it can be used in conjunction with search terms, manual review, and/or other search procedures. See, e.g., *NDLON* (“And beyond the use of keyword search, parties can (and frequently should) rely on latent semantic indexing, statistical probability models, and machine learning tools to find responsive documents.”); *Biomet I* (approving the combined use of keyword search terms and predictive coding); *Bridgestone* (approving the combined use of keyword search terms and predictive coding).

³⁵ Decisions on search methodologies and discovery technologies are frequently made before a complaint is filed or opposing counsel is identified. Clients who have made a significant investment in in-house tools should be able to leverage that investment to determine how to produce responsive information in a manner that reasonably satisfies their discovery obligations without having to seek permission to do so from litigation adversaries. See, e.g., *Bridgestone*, at *1-2 (holding that the responding party was free to combine predictive coding and keyword search methodologies to accomplish its production of documents despite arguably contrary language from a case management order); THE SEDONA CONFERENCE, THE SEDONA PRINCIPLES: BEST PRACTICES RECOMMENDATIONS & PRINCIPLES FOR ADDRESSING ELECTRONIC DOCUMENT PRODUCTION (*SEDONA PRINCIPLES*) 38 (Jonathan M. Redgrave et al. eds., 2d ed. 2007) (providing under Principle Six that the responding party is best situated to determine how it should search for, review, and produce its responsive documents), available at <https://thesedonaconference.org/publication/The%20Sedona%20Principles>.

³⁶ *But see* *Progressive Cas. Ins. Co. v. Delaney*, No. 2:11-cv-00678, 2014 WL 3563467, at *4, *8-11 (D. Nev. July 18, 2014) (denying the plaintiff’s request to use predictive coding to assist with its production of documents since (among other reasons) its proposed predictive coding methodology did “not comply with all of [its technology

evaluated based on whether the production satisfies the touchstones of relevance, proportionality, and reasonableness.³⁷ Three issues that are fundamental to doing so are discussed below.

First, counsel should confirm that it has accurately determined the prevalence of responsive information within the universe of documents. As an initial step in this process, counsel should be free to consider using keywords or other search methodologies in a reasonable manner to help narrow the subset of potentially responsive documents.³⁸ After doing so, counsel should then ensure that the “control set” or “sample set” of documents reflects the approximate percentage of responsive data found within that universe of documents.³⁹ Having an accurate reading of prevalence is essential to establishing overall search and production objectives for the predictive process.⁴⁰ If the prevalence evaluation is off the mark, the ultimate evaluation of the review and production quality will be difficult.⁴¹ The production could be under-inclusive, leaving potentially key, responsive information out of the production.⁴² Alternatively, the production could be over-inclusive, resulting in the production of too much marginally responsive or non-responsive data.⁴³

Next, counsel should prepare a seed or training set of relevant documents designed to elicit responsive information from the universe of documents.⁴⁴ A seed or training set of data is a relatively small subset of data that contains examples of the categories of information being sought (and sometimes includes negative examples to enhance training, depending on the algorithm).⁴⁵ The predictive coding algorithms

vendor’s] recommended best practices.”).

³⁷ *Bridgestone*, at 2 (“In the final analysis, the uses of predictive coding [are] 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.”).

³⁸ Charles Yablon & Nick Landsman-Roos, *Predictive Coding: Emerging Questions and Concerns*, 64 S.C. L. REV. 633, 638-39 (2013) (describing the cost benefits of removing clearly non-responsive documents from the universe of potentially responsive information).

³⁹ *Id.* at 640.

⁴⁰ *Id.*

⁴¹ *Id.*

⁴² *Id.*

⁴³ *Id.*

⁴⁴ See Daniel Martin Katz, *Quantitative Legal Prediction - Or - How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, 62 EMORY L.J. 909, 946 (2013) (explaining that predictive coding “approaches are inductive and typically involve the seeding of the algorithm with training (or labeled) data from which the machine infers the ‘true’ function for assigning a document to a particular group (i.e., relevant versus not relevant).”).

⁴⁵ Yablon, *supra* note 38, at 638-39, 643-44 (detailing the importance of the seed set to establishing a reasonable predictive coding workflow).

use the characteristics of the seed set to find similar documents.⁴⁶ Indeed, as United States District Judge Denise Cote opined in the *Federal Housing Finance Agency v. UBS Americas* case, “you should train your algorithm from the kinds of relevant documents that you might actually uncover in a search.”⁴⁷ Those seed documents should then be run through the predictive coding process to train the algorithm to obtain documents for the production.⁴⁸ This training process should then be repeated until it fails to yield new materially responsive results or when the algorithm has obtained sufficient information to accurately estimate the probability that any document is relevant to the matter.⁴⁹

Finally, counsel should validate the final production results from the predictive coding process through different forms of testing.⁵⁰ This will invariably entail taking statistically valid samples to ensure that the predictive coding process reached reasonable levels of recall and precision.⁵¹ This result will likely vary depending on the nature of the case, any agreement between the parties, or any court order addressing these issues.⁵²

III. SHOULD THE USE OF PREDICTIVE CODING BE DISCLOSED TO LITIGATION ADVERSARIES?

The issue of whether to disclose the use of predictive coding is significant and more than just a single inquiry. If counsel is inclined to reveal its use of predictive coding, how much information will it share? Will counsel merely divulge the fact that predictive coding will be used, enter into a stipulated use protocol, or adopt an approach that is somewhere between those positions?⁵³

⁴⁶ *Id.* (“Judgmental sampling, on the other hand, requires that attorneys with knowledge of the case select documents--already uncovered through discovery--as ‘seeds’ that they have determined are clearly fitting or not fitting a particular document category (e.g., a document is clearly relevant or not, privileged or not). That seed set of documents is fed into the software to train it for assessing relevancy.”).

⁴⁷ Transcript of Record at 114, *Fed. Hous. Fin. Agency v. UBS Americas*, 11-cv-06188 (S.D.N.Y. July 31, 2012) ECF No. 134. *But see* Karl Schieneman and Thomas Gricks, *The Implications of Rule 26(g) on the Use of Technology-Assisted Review*, 7 *FED. CTS. L. REV.* 239, 261-62 (2013) (“To the extent that the seed set reflects only the ‘most interesting’ documents, and is less than representative of the entire population of relevant documents, the use of technology-assisted review to propagate the seed decisions throughout the collection may be under-inclusive and run afoul of Rule 26(g).”).

⁴⁸ Transcript of Record at 114, *Fed. Hous. Fin. Agency v. UBS Americas*, 11-cv-06188 (S.D.N.Y. July 31, 2012) ECF No. 134; Yablon, *supra* note 38, at 639.

⁴⁹ Remus, *supra* note 2, at 111-12; Yablon, *supra* note 38, at 639. *See Biomet I* at *1 (describing the inextricably intertwined process used for seeding and training the predictive coding workflow).

⁵⁰ Byram, *supra* note 34, at 697 (“Sampling furthers the desirability of [predictive coding] by aiding in confirming a party’s results and satisfying the other parties of the contents of the production. In essence, sampling supports defensibility.”).

⁵¹ Barry, *supra* note 33, at 369-70.

⁵² *Id.*

⁵³ Schieneman, *supra* Note 47, at 254-57, 261-63 (describing some considerations surrounding disclosure and transparency with respect to predictive coding).

Disclosure is also likely a matter of first impression for many courts,⁵⁴ with outcomes influenced by a variety of factors. The history of the judge(s) overseeing the matter,⁵⁵ the nature of the client, the temperament of litigation adversaries,⁵⁶ and the particular phase of litigation in which predictive coding is to be used⁵⁷ may impact the decision-making process of both counsel and the courts on this issue.

The principal factor weighing in favor of some form of disclosure is that of greater certainty, *i.e.*, that an adversarial party will find it difficult to impugn the adequacy of the responding party's search methodology if it is aware that the responding party is using predictive coding.⁵⁸ Against this position stands the counter-argument to disclosure: potentially costly satellite litigation over an adversary's (real or perceived) dissatisfaction with the responding party's disclosed use of predictive coding.⁵⁹ That such satellite litigation may occur is borne out by the disclosures made – and the resulting protracted motion practice – in the *Da Silva Moore v. Publicis Groupe*.⁶⁰

Ultimately, the decision whether to disclose the use of predictive coding remains in the hands of the responding party and its counsel.⁶¹ This is consistent with the prevailing discovery practice reflected in case law and memorialized in other authoritative sources that the responding party is in the best and most appropriate position to determine how to produce its responsive documents and otherwise satisfy its discovery obligations imposed by the FRCP, local rules, and case law.⁶² While it may be strategically beneficial to enter into a cooperative dialogue with the requesting party, doing so may not always be

⁵⁴ See, e.g., Transcript of Record at 114, *Fed. Hous. Fin. Agency v. UBS Americas*, 11-cv-06188 (S.D.N.Y. July 31, 2012) ECF No. 134 (describing her experience with predictive coding, U.S. District Judge Denise Cote stated to counsel that “I’m learning about predictive coding as we go.”).

⁵⁵ *Id.*; Yablon, *supra* note 38, at 673 (discussing the divergence of views over “how much control courts may and should exercise over discovery practice.”).

⁵⁶ See *infra* Part IV.

⁵⁷ See *supra* Part II.

⁵⁸ Byram, *supra* note 34, at 699 (“Courts will look more favorably upon a party who discloses its key custodians and how it will [search] for the requested documents. Where a party is transparent, ‘opposing counsel and the Court are more apt to agree to your approach’”).

⁵⁹ See *infra* Part IV.

⁶⁰ See *infra* note 100.

⁶¹ *Kleen Products LLC v. Packaging Corp. of Am.*, No. 10-cv-5711, 2012 WL 4498465 (N.D. Ill. Sept. 28, 2012), *5 (“the Court observed that under Sedona Principle 6, ‘[r]esponding parties are best situated to evaluate the procedures, methodologies, and techniques appropriate for preserving and producing their own electronically stored information.’”); *Ford Motor Co. v. Edgewood Props.*, 257 F.R.D. 418, 427 (D.N.J. 2009) (“The Sedona Principles wisely state that it is, in fact, the producing party who is the best position to determine the method by which they will collect documents. The producing party responding to a document request has the best knowledge as to how documents have been preserved and maintained.”); *SEDONA PRINCIPLES*, at 38.

⁶² *Id.*

possible or advantageous.⁶³ Moreover, blanket requirements of transparency are not required by the spirit or letter of the law on this issue.⁶⁴

IV. WHAT ARE THE BENEFITS AND DRAWBACKS OF ENTERING INTO A STIPULATED PREDICTIVE CODING USE PROTOCOL?

A critical decision for counsel is whether to enter into a stipulated protocol regarding the use of predictive coding. Counsel's choice on this issue will affect the course of discovery, impact the relationship between the parties, and influence the court's perception of counsel and client.⁶⁵

It is worth emphasizing that it is the *choice* of counsel and the client to enter into such a protocol.⁶⁶ While some courts and commentators have taken the position that parties should enter into a stipulated protocol to use predictive coding,⁶⁷ neither the Rules nor case law require such a step⁶⁸ unless ordered by the court.⁶⁹ Indeed, whether counsel should take this step is entirely dependent on what is best for its client,⁷⁰ not what is convenient for opposing counsel or the court, so long as the general standard of an otherwise reasonable discovery process is met.⁷¹ While there are potential benefits to entering into such a protocol, there are also risks. The purpose of this section is to outline issues regarding the use of stipulations so counsel can make an informed decision that best represents client interests.

The Principal Benefit of a Stipulated Use Protocol is Cost Savings

⁶³ See *infra* Part IV.

⁶⁴ See, e.g., *Biomet II*; Remus, *supra* note 2, at 126-28.

⁶⁵ See, e.g., *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182, 204 (S.D.N.Y. 2012), *In re Actos (Pioglitazone) Products Liab. Litig.*, 6:11-MD-2299, 2012 WL 7861249 (W.D. La. July 27, 2012); Transcript of Record at 110-111, *Fed. Hous. Fin. Agency v. UBS Americas*, 11-cv-06188 (S.D.N.Y. July 31, 2012) ECF No. 134.

⁶⁶ *Kleen Products*, at *5, *19; *Ford Motor*, at 427; *SEDONA PRINCIPLES*, at 38.

⁶⁷ See, e.g., *Da Silva Moore*, 287 F.R.D. at 192; Byram, *supra* note 34, at 699.

⁶⁸ FED. R. CIV. P. 26(b); 34(b); *SEDONA PRINCIPLES*, at 38; *Kleen Products*, at *5, *19; *Ford Motor*, at 427.

⁶⁹ *Indep. Living Ctr. of S. Cal. v. City of L.A.*, 2:12-cv-00551-FMO-PJW, at 1 (C.D. Cal. June 13, 2014) ECF No. 371 (ordering the defendant “to use a predictive coding system for identifying the 10,000 most relevant documents in its databases and, after reviewing them for privilege, etc., produc[e] them to Plaintiffs); *Aurora Co-op. Elevator Co. v. Aventine Renewable Energy – Aurora West, LLC*, 4:12-cv-230 (D. Neb. Mar. 10, 2014) ECF No. 147 (ordering the parties to use predictive coding in connection with the first phase of discovery and, in connection with that process, to designate an expert who could work with the parties to design a computerized search to examine the parties’ electronic records).

⁷⁰ MODEL RULES OF PROF’L CONDUCT PREAMBLE & SCOPE ¶ 2 (2013) (explaining generally a lawyer’s duties to its client including, but not limited to, the requirement to “zealously assert[] the client’s position under the rules of the adversary system”).

⁷¹ See *supra* Parts I, II.

The primary objective of entering into a stipulated predictive coding protocol is generally to decrease the costs associated with pursuing discovery.⁷² Proponents of this strategy argue that costs will be reduced since discovery will – in theory – proceed in a more orderly fashion with the court and all parties cooperatively involved in the process.⁷³ Many stipulated protocols have invited opposing counsel to collaborate with and help prepare its adversary’s search methodology.⁷⁴ This may include sharing irrelevant documents with opposing counsel from the sample and seed sets, allowing counsel to assist with document coding, and involving counsel in the training and testing processes.⁷⁵ According to its proponents, such a cooperative and transparent approach will reduce satellite litigation over the process the party used to search for, review, and produce responsive information.⁷⁶ All of which will arguably make discovery less costly, more efficient, and ultimately focused on disclosing information to enable the parties to resolve matters on the merits.⁷⁷

Understanding the Drawbacks of a Stipulated Protocol

Against the backdrop of potentially lower discovery costs loom several drawbacks with stipulated use protocols. The first and most obvious risk is the potential for excessive input from and wrangling with opposing counsel and the court over the process for searching, reviewing, and producing documents.⁷⁸ The ESI search and review process has always been complex; allowing opposing counsel to participate may create tensions given the parties’ adversarial interests in the litigation.⁷⁹ Along with those tensions is the likelihood of motion practice and delays, which can offset the cost and time savings otherwise offered by predictive coding.⁸⁰

⁷² See, e.g., L. Casey Auttonberry, Comment, *Predictive Coding: Taking the Devil Out of the Details*, 74 LA. L. REV. 613, 624 (2014) (“Well-designed protocols can effectively decrease some of the costs and delays associated with e-discovery.”).

⁷³ See generally *Da Silva Moore*, 287 F.R.D. 182; *In Re Actos*, 2012 WL 7861249.

⁷⁴ *Id.* See also Judge Henry Coke Morgan, Jr., *Predictive Coding: A Trial Court Judge’s Perspective*, 26 REGENT U.L. REV. 71, 77-78 (2013) (opining that “counsel for all parties will participate in the two-step process of selecting the seed set of documents for coding . . . [which] must be transparent and acceptable to all parties.”).

⁷⁵ *Da Silva Moore*, 287 F.R.D. 182; *In Re Actos*, 2012 WL 7861249. *But see Biomet II* (“The Steering Committee wants the whole seed set Biomet used for the algorithm’s initial training. That request reaches well beyond the scope of any permissible discovery by seeking irrelevant or privileged documents used to tell the algorithm what not to find. That the Steering Committee has no right to discover irrelevant or privileged documents seems self-evident.”).

⁷⁶ Schieneman, *supra* note 47, at 261-63; Byram, *supra* note 34, at 698-699.

⁷⁷ *Id.*

⁷⁸ Transcript of Record at 110-111, *Fed. Hous. Fin. Agency v. UBS Americas*, 11-cv-06188 (S.D.N.Y. July 31, 2012) ECF No. 134; *Da Silva Moore*, 287 F.R.D. at 185-189.

⁷⁹ *Id.*

⁸⁰ This is particularly evident from the docket in *Da Silva Moore*, which reflects repeated motion practice over nominal issues relating to the parties’ predictive coding protocol. See *infra* note 100.

Another hazard with stipulated use protocols is that they may require counsel to disclose to its litigation adversary non-responsive information, particularly non-responsive documents used to train the predictive coding algorithm.⁸¹ Besides the fact that such information is outside the permissible scope of discovery,⁸² disclosing non-responsive documents may violate counsel's ethical duty of confidentiality.⁸³ Non-responsive information may contain trade secrets, sensitive financial data, or other proprietary information that should not be disclosed to a litigation adversary,⁸⁴ especially if the adversary is a business competitor.⁸⁵ Moreover, depending on the nature of the information, it could be used to amplify claims in the present lawsuit⁸⁶ or to file a new lawsuit against the producing party.⁸⁷

A third problem associated with stipulated protocols is the risk of waiving attorney work product protection by voluntarily identifying the predictive coding training set to opposing counsel.⁸⁸ A training set may reflect a lawyer's litigation strategy and perceptions of relevance,⁸⁹ which are frequently held to be opinion work product.⁹⁰ By pinpointing those specific documents for a litigation adversary, counsel

⁸¹ See Remus, *supra* note 2, at 125-26.

⁸² FED. R. CIV. P. 26(b)(1); *Biomet II*, at *1-2.

⁸³ See, e.g., MODEL RULES OF PROF'L CONDUCT R. 1.6 (2013) (delineating the general rule that "[a] lawyer shall not reveal information relating to the representation of a client," along with pertinent exceptions); CAL. BUS. AND PROF'L CODE § 6068(e) ("It is the duty of an attorney to...maintain inviolate the confidence, and at every peril to himself or herself to preserve the secrets, of his or her client"); STATE BAR OF CALIFORNIA STANDING COMM. ON PROF'L RESPONSIBILITY AND CONDUCT, PROPOSED FORMAL OP. INTERIM NO. 11-0004 (2014), *available at* <http://www.calbar.ca.gov/AboutUs/PublicComment/201404.aspx>.

⁸⁴ *Id.*

⁸⁵ *Cf.* Apple v. Samsung, 5:11-cv-01846, 2014 WL 2854994, at *4 (N.D. Cal. June 20, 2014) (imposing a \$2 million sanction on the defendant's counsel (among others) for making an unauthorized disclosure to its client of documents produced by the plaintiff that were designated "attorneys' eyes only" under the governing protective order).

⁸⁶ *Cf.* In re Google Inc., 462 F. App'x 975, 978 (Fed. Cir. 2012) (the defendant's unintended disclosure of certain documents claimed as privileged arguably strengthened the plaintiff's claims).

⁸⁷ *Cf.* Fed. Hous. Fin. Agency v. HSBC N. Am. Holdings Inc., 2014 WL 584300, at *3 (S.D.N.Y. Feb. 14, 2014) (rejecting certain defendants' attempt to use documents produced in a related action to challenge the reasonableness of the plaintiff's document production in the instant action).

⁸⁸ See *Biomet II*; Remus, *supra* note 2, at 125-26 ("requiring seed-set transparency threatens core protections for attorney work product, attorney-client privilege, and confidentiality."); Yablon, *supra* note 38, at 644 ("If . . . the seed set is made up of documents selected or coded by a producing party as relevant, production of that seed set has a much higher probability of disclosing attorney impressions of the case.").

⁸⁹ *Id.*

⁹⁰ FED. R. CIV. P. 26(b)(3)(B) ("If the court orders discovery of those materials, it must protect against disclosure of the mental impressions, conclusions, opinions, or legal theories of a party's attorney or other representative concerning the litigation."); *Hickman v. Taylor*, 329 U.S. 495, 516 (1947) (Jackson, J., concurring) ("Discovery was hardly intended to enable a learned profession to perform its functions either without wits or on wits borrowed from the adversary"); *Lockheed Martin Corp. v. L-3 Communs. Corp.*, 6:05-cv-1580-Orl-31KRS, 2007 WL 2209250, at *9 (M.D. Fla. July 29, 2007) (explaining that opinion work product – which includes "an attorney's mental impressions, conclusions, opinions, or legal theories" – is entitled to "a nearly absolute immunity and can be discovered only in very rare and extraordinary circumstances").

likely yields any work product protection that may have otherwise been associated with its identification of those documents.⁹¹

A related risk with stipulated protocols is the possibility that attorney-client privileged information could be inadvertently shared with opposing counsel.⁹² This is particularly the case where privileged communications are used to train the predictive coding process.⁹³ Unless appropriate screening measures are deployed, counsel could inadvertently divulge privileged materials to opposing counsel.⁹⁴ And while properly executed orders under Federal Rule of Evidence 502(d) may address the problem of inadvertent waiver,⁹⁵ they cannot remove the nature of the privileged information now in the mind of opposing counsel.

Finally, any deviations from the stipulated protocol that counsel may want to implement will likely require consent from opposing counsel or the court.⁹⁶ While neither may be forthcoming, implementing unilateral changes to the agreed-upon process may invite the court's displeasure and adverse consequences to counsel and the client.⁹⁷

The Jurisprudence on Stipulated Protocols

There are few reported cases that address the benefits and risks of entering into a stipulated predictive coding protocol. In particular, the extant jurisprudence illustrates the risks of those protocols. This is not surprising since “reported decisions tend to involve obstructionist conduct at the most egregious end of

⁹¹ *Simmons, Inc. v. Bombardier, Inc.*, 221 F.R.D. 4, 8 (D.D.C. 2004) (“The work-product privilege may be waived by the voluntary release of materials otherwise protected by it.”).

⁹² *See, e.g.*, Transcript of Record at 14-15, *Fed. Hous. Fin. Agency v. JPMorgan Chase & Co., Inc.*, 11-cv-05201 (S.D.N.Y. July 24, 2012) ECF No. 128.

⁹³ *Id.* (addressing the issue of privilege protection in connection with sharing seed set documents with the plaintiff's counsel, counsel for one of the defendants requested the court “to make some provision, your Honor, which I think would be necessary, to deal with the problem that privileged documents could be part of the seed set, and when you don't have plaintiff's counsel in the picture, you don't have to worry about maintaining the privilege, but once you do have them in the picture, you'd have to take a step or two to make sure that that wasn't going to be an issue.”).

⁹⁴ *Id.*

⁹⁵ FED. R. EVID. 502(d); *see also* FED. R. EVID. 502 (advisory committee notes) (discussing the rule's framework for addressing the problems associated with the inadvertent production of ESI); John M. Barkett, *Evidence Rule 502: The Solution to the Privilege-Protection Puzzle in the Digital Era*, 81 *FORDHAM L. REV.* 1589, 1619-20 (2013) (discussing the importance of Federal Rule of Evidence 502(d) in reducing the costs and burdens associated with attorney-client privilege reviews in discovery).

⁹⁶ *Compare Progressive Cas. Ins. Co. v. Delaney*, No. 2:11-cv-00678, 2014 WL 3563467, at *10-12 (D. Nev. July 18, 2014) (interpreting a case management order provision to prevent the unilateral adoption of predictive coding) *with Bridgestone* (allowing the use of predictive coding despite an arguably contrary case management order).

⁹⁷ *Progressive*, 2014 WL 3563467, at *10-12. *See also Biomet II* (warning that “[a]n unexplained lack of cooperation in discovery can lead a court to question why the uncooperative party is hiding something. . .”).

the spectrum.”⁹⁸ The following three cases are instructive on these issues and are analyzed in the following paragraphs.

Da Silva Moore v. Publicis Groupe

In *Da Silva Moore*, the parties agreed in principal to use predictive coding, but only reached a stipulated protocol after substantial prodding from the court and with the plaintiffs taking the unusual step of “objecting” to the agreement.⁹⁹ The relationship between the parties subsequently deteriorated into protracted motion practice over minor issues of relevance on individual documents.¹⁰⁰

The resulting acrimony between the parties was the result of various factors, including the nature of the stipulated protocol and the plaintiffs’ obstreperous conduct.¹⁰¹ With respect to the protocol, it allowed the plaintiffs to review all non-privileged documents – both responsive and non-responsive – used to train the predictive coding process, together with the defendant’s document coding designations.¹⁰² The plaintiffs took advantage of the protocol’s transparency to challenge nominal issues of relevance and litigate other collateral matters, all of which led to protracted motion practice and increased attorney fees.¹⁰³

Da Silva Moore teaches that parties must exercise care in entering into a stipulated use protocol. Allowing an adversary to help train the predictive coding process may create difficulties in discovery that previously were non-existent.¹⁰⁴ While providing complete transparency into the predictive coding

⁹⁸ Robert Douglas Brownstone, *Collaborative Navigation of the Stormy e-Discovery Seas*, 10 RICH. J.L. & TECH. 53, ¶ 29 (2004).

⁹⁹ *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182, 187, n.6, 204 (S.D.N.Y. 2012).

¹⁰⁰ *See, e.g.*, Plaintiffs’ Rule 72(a) Objection to the Magistrate’s April 25, 2012 Discovery Rulings, *Da Silva Moore v. Publicis Groupe*, No. 11-cv-1279 (S.D.N.Y. May 9, 2012), ECF No. 190 (seeking reconsideration of various court rulings on discovery issues relating to the defendant’s use of predictive coding). The motion was later withdrawn after the court granted the plaintiffs’ request to stay a certain production of documents by the defendants, thereby mooting the remainder of the plaintiffs’ objections. *See Da Silva Moore v. Publicis Groupe*, No. 11-cv-1279 (S.D.N.Y. May 14, 2012), ECF No. 196.

¹⁰¹ *Id.*; *Da Silva Moore*, 287 F.R.D. at 193-204.

¹⁰² *Da Silva Moore*, 287 F.R.D. at 199-203 (“[Defendant] will provide Plaintiffs’ counsel with all of the non-privileged documents and will provide, to the extent applicable, the issue tag(s) coded for each document, as described above. Plaintiffs’ counsel shall promptly review and provide notice as to any documents with which they disagree where they do not understand the coding. If necessary, counsel will meet and confer to attempt to resolve any disagreements regarding the coding applied to the documents in this seed set.”).

¹⁰³ *See, e.g.*, Plaintiffs’ Rule 72(a) Objection to the Magistrate’s April 25, 2012 Discovery Rulings, *supra* note 101 (requesting reconsideration of multiple court rulings on issues of relevance and proportionality relating to the defendant’s use of predictive coding).

¹⁰⁴ *See, e.g.*, Transcript of Record at 110-111, *Fed. Hous. Fin. Agency v. UBS Americas*, 11-cv-06188 (S.D.N.Y. July 31, 2012) ECF No. 134. Counsel for defendant JPMorgan Chase & Co. summarized the difficulties of working jointly with the plaintiff in developing a predictive coding workflow: “We meet every day with the plaintiff to have a status report, get input, and do the best we can to integrate that input. It isn’t always easy, not just to carry out those functions but to work with the plaintiff. The suggestions we have had so far have been unworkable and by and large would have swamped the project from the outset and each day that a new suggestion gets made. But we do our best to explain that and keep moving forward.” *Id.*

process may at some level be more convenient for the court and opposing counsel,¹⁰⁵ it may not serve the client's interests to do so.¹⁰⁶ As with any strategic decision, whether to enter into a stipulation and its precise terms must be carefully considered to determine the best course of action for the client.¹⁰⁷

In re Actos (Pioglitazone) Products Liability Litigation

Just like the *Da Silva Moore* case, *In re Actos* involved a scenario where the parties entered into a predictive coding use protocol.¹⁰⁸ However, that is where the similarity between these two cases ends. *In re Actos* involved an unusual level of cooperation between the parties involved in that multidistrict litigation. Indeed, the court lauded the parties' collaborative efforts on multiple occasions, praising them as both "exceptional" and "impressive."¹⁰⁹

Given the parties' cooperative efforts, the court entered a case management order reflecting the parties' stipulation regarding the production of ESI.¹¹⁰ That order, which encompassed the use protocol, provided for an extraordinary level of transparency in the predictive coding process.¹¹¹ Among other things, the parties agreed to jointly prepare, develop, and carry out the following: the sample set of documents to determine prevalence, the predictive coding training set, the cut-off relevance probability score assigned by the algorithm to a document that would determine which documents the defendants would manually review for relevance, and the validation process.¹¹² To address the defendants' concern regarding the possible disclosure of non-responsive trade secrets, the court entered an additional order memorializing

¹⁰⁵ See *Da Silva Moore*, 287 F.R.D. at 192 ("transparency allows the opposing counsel (and the Court) to be more comfortable with computer-assisted review, reducing fears about the so-called 'black box' of the technology. This Court highly recommends that counsel in future cases be willing to at least discuss, if not agree to, such transparency in the computer-assisted review process.").

¹⁰⁶ MODEL RULES OF PROF'L CONDUCT PREAMBLE & SCOPE ¶ 2 (2013).

¹⁰⁷ *Id.*

¹⁰⁸ *In re Actos (Pioglitazone) Products Liab. Litig.*, 6:11-MD-2299, 2012 WL 7861249 (W.D. La. July 27, 2012).

¹⁰⁹ *In re Actos (Pioglitazone) Products Liab. Litig.*, 6:11-MD-2299, 2014 WL 2921653, at *2 (W.D. La. June 23, 2014) ("As a general matter, this case has progressed in an exceptional manner due to the continued and professional cooperation of the parties, aided by the Special Masters and with input of the magistrate judge, resulting in the resolution of most discovery issues and disputes without the need for formal court intervention."); *In re Actos (Pioglitazone) Products Liab. Litig.*, 6:11-MD-2299 (W.D. La. July 13, 2012) ECF No. 1413 (noting the "impressive levels of cooperation, coordination, and negotiation that have taken place among counsel for the plaintiffs and the defense").

¹¹⁰ *In re Actos (Pioglitazone) Products Liab. Litig.*, 6:11-MD-2299, 2012 WL 7861249 (W.D. La. July 27, 2012).

¹¹¹ *Id.*

¹¹² *Id.*

the parties' stipulation that would essentially preserve the confidentiality of those irrelevant trade secrets.¹¹³

Unlike *Da Silva Moore*, the result of the parties' highly collaborative process was that most discovery issues in *In re Actos* were resolved without motion practice.¹¹⁴ As a result, the *In re Actos* defendants, which possessed the majority of the discoverable documents in this asymmetrical products liability case, likely benefited from this arrangement with reduced search and review costs.¹¹⁵

The apparent success of the *In re Actos* protocol is instructive on multiple levels for lawyers seeking to use predictive coding. On the one hand, *In re Actos* demonstrates that lawyers can successfully implement a highly transparent and cooperative predictive coding workflow that can expedite discovery while reducing the costs of the search and review process. Moreover, it shows that this process can be accomplished without devolving into tangential motion practice that characterized *Da Silva Moore*.¹¹⁶ On the other hand, because many lawyers might not engage in such "exceptional" cooperation, entering a stipulated use protocol may not be in the client's best interests. Much like *Da Silva Moore*, *In re Actos* ultimately teaches that counsel should evaluate all of the circumstances in a given case – particularly the temperament of its litigation adversary – before deciding to enter into a stipulated protocol.¹¹⁷

Progressive Casualty Insurance Co. v. Delaney

In *Progressive*, the court took the unusual step of prohibiting the plaintiff from using predictive coding to search through and review its documents.¹¹⁸ The parties had initially entered into a keyword search protocol that was reflected in the court's case management order.¹¹⁹ Pursuant to that order, the plaintiff agreed to run a series of keyword searches jointly developed with the defendants against its universe of potentially responsive information.¹²⁰ Once it completed those searches, the plaintiff would either manually review the subset of potentially responsive documents or simply produce all of the documents except for those that were potentially privileged.¹²¹ While this protocol seemed acceptable at first, the

¹¹³ *In re Actos* (Pioglitazone) Products Liab. Litig., 6:11-MD-2299 (W.D. La. July 30, 2012) ECF No. 1540.

¹¹⁴ *See supra* note 109 and accompanying text.

¹¹⁵ Despite the ostensible success of the parties' predictive coding protocol, one of the defendants suffered a difficult discovery setback when the court imposed severe remedial sanctions for its alleged spoliation of electronic information. *See In re Actos* (Pioglitazone) Products Liab. Litig., 6:11-MD-2299, 2014 WL 2921653, at *59-61 (W.D. La. June 23, 2014) (issuing an adverse inference instruction against one of the defendants and ordering, among other things, that defendant to "continue at its sole cost to reconstruct all deleted files").

¹¹⁶ *See supra* note 109 and accompanying text.

¹¹⁷ *See* MODEL RULES OF PROF'L CONDUCT PREAMBLE & SCOPE ¶ 2 (2013).

¹¹⁸ *Progressive Cas. Ins. Co. v. Delaney*, No. 2:11-cv-00678, 2014 WL 3563467, at *10-12 (D. Nev. July 18, 2014).

¹¹⁹ *Id.* at *6-9.

¹²⁰ *Id.*

¹²¹ *Id.*

plaintiff quickly determined that it would be too expensive to manually review the over 500,000 documents that remained after running the keyword searches.¹²²

The plaintiff then turned to predictive coding given its potential to expedite the search process at a reduced cost.¹²³ However, the plaintiff did not initially disclose to opposing counsel its decision to modify its keyword search protocol and neglected to produce any responsive documents in the meantime.¹²⁴ When the parties subsequently failed in their attempt to reach an agreement on the use of predictive coding, the court placed the blame on the plaintiff.¹²⁵ Not only did the court forbid the plaintiff from using predictive coding, it ordered the plaintiff to make a blanket production of the potentially responsive documents within two weeks, with the exception that it could withhold arguably privileged materials.¹²⁶

The fundamental lesson from *Progressive* concerns unilaterally violating an agreement entered as an order of the court.¹²⁷ While the *Progressive* case dealt with a keyword search protocol, its holding is equally applicable to predictive coding use stipulations. Once such a stipulation is reached and reduced to a court order, it may be difficult to change.¹²⁸ If counsel becomes dissatisfied with the framework governing the predictive coding process, e.g., the level of its adversary's involvement in the training of the process, it may be foreclosed from unilaterally modifying its strategy and tactics just like the plaintiff from *Progressive*.¹²⁹ Instead, counsel would likely have to negotiate with its adversary or ask the court to modify the agreement.¹³⁰ All of which could be costly to the client and could still fail to yield the sought-after results.

While a stipulated predictive coding use protocol could have cost benefits, there are potentially significant risks associated with entering into such agreements. To ensure that the client's interests are properly represented in discovery, counsel should carefully investigate at the outset of litigation which search and review protocols will best address the demands of discovery in the case.¹³¹ Counsel should also gauge the

¹²² *Id.* at *2-5.

¹²³ *Id.*

¹²⁴ *Id.* at *9.

¹²⁵ *Id.* at *10-12.

¹²⁶ *Id.*

¹²⁷ *Id.*

¹²⁸ *Id.*

¹²⁹ *Contra Bridgestone*, at *1-2.

¹³⁰ See IBM's Response to Bridgestone's Request to Add Predictive Coding to the ESI Protocol, *Bridgestone Americas, Inc. v. Int'l Bus. Mach. Corp.*, No. 3:13-cv-1196 (M.D. Tenn. July 14, 2014) ECF No. 85; Bridgestone's Reply to IBM's Opposition to Use of Attorney Trained Technology-Assisted Review, *Bridgestone Americas, Inc. v. Int'l Bus. Mach. Corp.*, No. 3:13-cv-1196 (M.D. Tenn. July 18, 2014) ECF No. 88; *Progressive Cas. Ins. Co. v. Delaney*, No. 2:11-cv-00678, 2014 WL 3563467, at *4, *8-11 (D. Nev. July 18, 2014).

¹³¹ See Byram, *supra* note 34, at 697.

cooperativeness of opposing counsel and determine the level of involvement (if any) that its adversary should have in the search and review process.¹³² Finally, counsel should ensure that its client's interests, including its secrets and confidences, are protected, particularly if it decides to enter into a stipulated use protocol.¹³³

By addressing these issues in a prudent fashion, counsel will better ensure that it avoids the risks exemplified in the above referenced cases.

V. CONCLUSION

The complexities and confusion surrounding the use of predictive coding need not obscure the point that predictive coding is a viable, judicially-approved, and dynamic method for conducting discovery. So long as its workflow meets the Rule 1 mandate and the related notions of relevance, proportionality, and reasonableness, use of predictive coding should be defensible. To adequately satisfy these standards does not always require counsel to enter into a stipulated use protocol or to disclose its use of predictive coding. Those decisions may or may not be advisable depending on the circumstances of a case.

To establish defensibility, counsel must accurately determine the prevalence of responsive information, ensure that its training process yields acceptable levels of recall and precision for its production of documents, and validate its final production results. By addressing these issues and by deploying safeguards to protect against disclosures of non-responsive information, attorney work product, or attorney-client privileged materials, counsel should have the elements in place to successfully use predictive coding in discovery.

¹³² See *Kleen Products LLC v. Packaging Corp. of Am.*, No. 10-cv-5711, 2012 WL 4498465, at *5 (N.D. Ill. Sept. 28, 2012); *Ford Motor Co. v. Edgewood Props.*, 257 F.R.D. 418, 427 (D.N.J. 2009); *SEDONA PRINCIPLES*, at 38.

¹³³ See MODEL RULES OF PROF'L CONDUCT R. 1.6 (2013); STATE BAR OF CALIFORNIA STANDING COMM. ON PROF'L RESPONSIBILITY AND CONDUCT, PROPOSED FORMAL OP. INTERIM NO. 11-0004 (2014), available at <http://www.calbar.ca.gov/AboutUs/PublicComment/201404.aspx>.