

2016

Calling an End to Culling: Predictive Coding and the New Federal Rules of Civil Procedure

Stephanie Serhan
University of Richmond

Follow this and additional works at: <http://scholarship.richmond.edu/law-student-publications>



Part of the [Civil Procedure Commons](#)

Recommended Citation

Stephanie Serhan, *Calling an End to Culling: Predictive Coding and the New Federal Rules of Civil Procedure*, 23 Rich. J.L. & Tech. 5 (2016).

This Article is brought to you for free and open access by the School of Law at UR Scholarship Repository. It has been accepted for inclusion in Law Student Publications by an authorized administrator of UR Scholarship Repository. For more information, please contact scholarshiprepository@richmond.edu.

**CALLING AN END TO CULLING: PREDICTIVE CODING AND THE NEW
FEDERAL RULES OF CIVIL PROCEDURE**

Stephanie Serhan*

Cite as: Stephanie Serhan, *Calling an End to Culling: Predictive Coding and the New Federal Rules of Civil Procedure*, 23 Rich. J.L. & Tech. 5 (2016), http://jolt.richmond.edu/index.php/volume23_issue2_serhan/.

TABLE OF CONTENTS

I. INTRODUCTION	2
II. WHY TIMING MATTERS IN PREDICTIVE CODING.....	4
A. The Technical Difference Between the Two Methods.....	5
B. The Practical Implications in Applying the Two Methods.....	6
III. COURT DECISIONS AND THE NEW FEDERAL RULES	11
A. Court Decisions under the Old Rules	11
1. Ex-Ante Permissibility of Predictive Coding.....	11
2. Ex-Post Permissibility of Keyword Culling.....	15
B. Reinforcement of Court Decisions under the New Rules.....	21
1. Recent Amendments to the Rules	21
2. Subsequent Reactions to the New Rules	25
IV. ENCOURAGING PREDICTIVE CODING EX ANTE	28
A. Why Predictive Coding Ex Ante is Preferable	28

B. How Parties and Courts Should Proceed	30
V. CONCLUSION	35

I. INTRODUCTION

[1] In corporate litigation and dispute resolution, discovery is often a significant undertaking for both the producing and requesting parties. Each party's approach during discovery is usually guided by considerations regarding efficiency and accuracy during the process. One area of discovery in which parties prioritize these considerations is the implementation of predictive coding. Several studies have proven that the method of predictive coding is substantially more efficient and accurate than traditional methods of conducting discovery.¹

[2] The method of predictive coding begins with a senior attorney who is intimately familiar with the case identifying relevant and irrelevant

*J.D. Candidate 2017, University of Richmond School of Law. B.A., 2012, American University of Beirut. The author gratefully acknowledges Professor Jessica Erickson for her mentorship in the organization and articulation of arguments in this article, as well as Ms. Meghan Podolny for her assistance in the primary research phase of this topic. The author would also like to thank the editors and staff of the Richmond Journal of Law & Technology for their efforts in editing this article.

¹ See, e.g., Maura R. Grossman & Gordon V. Cormack, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, 17 RICH. J.L. & TECH. 1, 43, 48 (2011) (discussing benefits of predictive coding when conducting discovery); see also Joe Palazzolo, *Why Hire a Lawyer? Computers are Cheaper*, WALL STREET J., (June 18, 2012, 2:06 PM), <http://www.wsj.com/articles/SB10001424052702303379204577472633591769336>, archived at <https://perma.cc/FRN2-BTMW> (noting that predictive coding is one subset of technology-assisted review (TAR) processes); see Andrew Peck, *Search, Forward; Will Manual Document Review and Keyword Searches be Replaced by Computer-Assisted Coding?*, LAW TECH. NEWS (Oct. 2011), https://law.duke.edu/sites/default/files/centers/judicialstudies/TAR_conference/Panel_1-Background_Paper.pdf, archived at <https://perma.cc/7DDK-3HL5>.

documents to create a “seed set.”² This seed set is then fed into the predictive coding software, which trains the software to determine which documents are relevant, while suggesting other documents that may also be relevant.³ Additionally, the attorney might review a random sample of documents;⁴ or the attorney could feed in words, phrases, and concepts that are appropriate to the case, and the software can subsequently find similar phrases, with linguistic or sociological relevance.⁵ The aim of the method is to identify the most relevant documents to produce to the requesting party.

[3] Within predictive coding, tension between efficiency and accuracy frequently arises in deciding the appropriate time at which to apply predictive coding. This timing concern has sparked numerous debates, as well as a split between court opinions. The issue parties and courts address is whether predictive coding should be applied at the outset of discovery to an entire universe of documents, or if it should be applied after keyword culling.

[4] This issue has become increasingly addressed in virtually every important case that has large volumes of documents in discovery. Addressing this issue is important to the parties involved because it has profound implications regarding efficiency and accuracy. Courts have also been asked to address this question, but have offered little guidance

² COVINGTON & BURLING LLP, *The Duty to Produce ESI, in LITIGATING SECURITIES CLASS ACTIONS* § 13.04(2)(c) (Jonathan Eisenberg ed., 2016).

³ See Tonia Hap Murphy, *Mandating Use of Predictive Coding in Electronic Discovery: An Ill-Advised Judicial Intrusion*, 50 AM. BUS. L.J. 609, 618 (2013) (noting that predictive coding uses sophisticated technology to narrow down documents that are most relevant to a case).

⁴ See *id.*

⁵ See *id.* at 617.

regarding the time at which to implement predictive coding in a case. Rule 1 of the Federal Rules of Civil Procedure addresses this exact balance as a trade-off between the just resolution and the efficiency of a case, which has often arisen in issues concerning discovery.⁶ The recent amendments to the Federal Rules of Civil Procedure further emphasize this trade-off.⁷

[5] This paper examines the impact of the most recent amendments to the Federal Rules of Civil Procedure on the current split between courts about whether predictive coding should be applied at the outset or to a set of keyword-culled documents. Since the new Rules explicitly implement the concept of proportionality and a new set of standards in Rule 26, I argue that applying predictive coding at the outset is more compliant with the Federal Rules of Civil Procedure. Part II will explain the difference in timing between applying predictive coding after keyword culling or prior to it, and discuss the implications of accuracy and efficiency. Part III will first discuss the split between courts regarding the two methods prior to the recent amendments to the Rules, and subsequently, it will discuss reactions by courts and scholars regarding the applicability after the amendments to the Rules. Part IV will argue that the method of applying predictive coding at the outset is more compliant with the new amendments to the Rules since it is more accurate, and it will suggest that parties and courts should begin to implement these changes. Ultimately, this proposal will improve accuracy, without jeopardizing efficiency, with the goal of achieving the just resolution of a case.

II. WHY TIMING MATTERS IN PREDICTIVE CODING

[6] During the process of discovery, parties often face a choice regarding which method to use on large volumes of documents. Predictive coding has recently become a predominant method through which

⁶ See FED. R. CIV. P. 1.

⁷ See discussion *infra* Part III.B.1.

attorneys and parties alike may narrow down the universe of documents in an efficient and accurate manner.⁸ However, parties differ over the appropriate time at which predictive coding should be used in the discovery process, which has created two methods that differ only in timing. The two methods are: (i) the use of predictive coding at the outset, or (ii) the use of predictive coding after keyword culling documents. This Part explains the technical difference between these two methods, as well as the practical implications in applying each of these methods.

A. The Technical Difference Between the Two Methods

[7] Regarding the timing of when to apply predictive coding, the two methods are: (i) the use of predictive coding at the outset, or (ii) the use of predictive coding after keyword culling. The first method involves applying predictive coding at the beginning of the discovery phase; the second method involves keyword culling documents first, and subsequently applying predictive coding to the keyword-culled documents. Each of these methods will be explained separately.

[8] The first method provides the option of applying predictive coding to the entire universe of documents at the beginning of the discovery phase. All documents are gathered, and the predictive coding technology is applied to all of the documents at the outset as a whole.⁹ Applying

⁸ See *Da Silva Moore v. Publicis Groupe & MSL Group*, 287 F.R.D. 182, 193 (S.D.N.Y. 2012). The *Da Silva Moore* case has received a significant amount of attention, since it was the first case in which predictive coding was judicially approved. See also Bennett B. Borden & Jason R. Baron, *Finding the Signal in the Noise: Information Governance, Analytics, and the Future of Legal Practice*, 20 RICH. J.L. & TECH. 1, 7, 16 (2014) (providing an in-depth statistical analysis finding that predictive coding is abundantly more accurate and efficient than traditional methods of discovery); see generally Grossman & Cormack, *supra* note 1, at 3 (discussing the efficiency and effectiveness of predictive coding).

⁹ See *Most Important Documents Get Looked at First: Using Predictive Coding to Prioritize & Expedite Review*, CONSILIO (2016), <http://www.consilio.com/wp->

predictive coding to all documents means there is no previous method, such as keyword culling, to narrow down the universe of documents. The use of predictive coding will narrow down the universe of documents based on which documents are relevant, or predicted to be relevant, through a programmed algorithm.¹⁰ Alternatively, the second method allows a party to apply predictive coding to a set of documents that has already been reduced in size by keyword search techniques. These techniques are frequently referred to as “keyword culling.” In order to perform keyword culling on documents, a party would begin with the entire universe of documents that pertain to a case, and narrow down the universe of documents by searching for keywords. Through this method, documents are identified as relevant or irrelevant based on those search terms. The relevant documents remain, and these are a much smaller set of documents. These relevant documents are referred to as the keyword-culled documents, and predictive coding is subsequently applied only to these keyword-culled documents.¹¹

B. The Practical Implications in Applying the Two Methods

[9] These two methods have significant implications regarding a party’s monetary expenditures and time spent, which relates to important concerns of accuracy and efficiency in choosing between these two methods. Regarding accuracy, the use of predictive coding at the outset provides a much more accurate return of relevant documents than keyword

content/uploads/2016/01/Using-Predictive-Coding-to-Expedite-Review.pdf, *archived at* <https://perma.cc/8R9L-6N5V> (noting that if predictive coding were used at the outset it would have saved 70% of the time it took to conduct manual review).

¹⁰ See Murphy, *supra* note 3, at 621–22.

¹¹ See Jim Eidelman, *Best Practices in Predictive Coding: When are Pre-Culling and Keyword Searching Defensible?*, CATALYST, Jan. 9, 2012, <http://catalystsecure.com/blog/2012/01/best-practices-in-predictive-coding-when-are-pre-culling-and-keyword-searching-defensible/>, *archived at* <https://perma.cc/GG8K-3MMF>.

culling.¹² Applying predictive coding on the entire set of documents is the most accurate method in identifying relevant documents because it is applied to all documents, rather than the ones selected by keyword culling.¹³ Keyword culling is not as accurate because the party may lose many relevant documents if the documents do not contain the specified search terms, have typographical errors, or use alternative phraseologies.¹⁴ The relevant documents removed by keyword culling would likely have been identified using predictive coding at the outset instead.¹⁵ Therefore, keyword culling is not as accurate as predictive coding when used on the entire set of documents at the outset.

[10] Regarding efficiency, both methods provide efficient returns, depending on how efficiency is defined. The use of predictive coding at the outset can be beneficial in narrowing down documents based on even “‘linguistic’ or ‘sociological’” relevance.¹⁶ Another efficient benefit is that the technology is programmed at the outset and can identify the most relevant documents.¹⁷ Keyword culling, on the other hand, narrows down

¹² See *id.*; see also Barry Kazan & David Wilson, *Technology-Assisted Review Is a Promising Tool for Document Production*, N.Y. L.J. ONLINE, Mar. 18, 2013, <http://www.newyorklawjournal.com/id=1202592178481/TechnologyAssisted-Review-Is-a-Promising-Tool--for-Document-Production>, archived at <https://perma.cc/QZ6J-BVD6> (citing a case in which one party found that keyword culling only produces 20% of relevant documents, whereas predictive coding would be sufficient even when finding at a 75% responsive rate).

¹³ See Eidelman, *supra* note 11.

¹⁴ See Kazan & Wilson, *supra* note 12.

¹⁵ See John Hopkins, *Large Data and Document Production – Keyword Search and Predictive Coding*, SEARCY L. BLOG, May 31, 2013, <https://www.searcyllaw.com/large-data-and-document-production-keyword-search-and-predictive-coding/>, archived at <https://perma.cc/VA9V-HJXM>.

¹⁶ Murphy, *supra* note 3, at 617.

¹⁷ See *id.* at 620.

the universe of documents by conducting a keyword search that does not identify other potentially-relevant documents, but simply searches through the documents using the keywords that are chosen.¹⁸ The keyword search can be quickly applied to a set of documents to determine which documents to keep and which to remove.¹⁹ Keyword culling can be useful since it narrows down the universe of documents to a much smaller number, as it does not predict other potentially-relevant documents.²⁰ It may be quicker for the technology to simply apply keyword searches prior to predictive coding to limit the number of documents that need to be coded, but once again, it comes at the cost of accuracy in revealing responsive documents.²¹

[11] Furthermore, prior to keyword culling, the parties often spend significant amounts of time discussing which keywords to employ in the search.²² This back and forth between the parties frequently results in

¹⁸ See *id.* at 614–16, 620.

¹⁹ The traditional way to employ keyword culling is run keywords through documents to retain the documents, which contain those keywords. See Ralph C. Losey, *Predictive Coding and the Proportionality Doctrine: A Marriage Made in Big Data*, 26 REGENT U. L. REV. 7, 58–59 (2013) (arguing that keyword culling could instead be used to cull documents out that are least likely to be relevant).

²⁰ See Jacob Tingen, *Technologies-That-Must-Not-Be-Named: Understanding and Implementing Advanced Search Technologies in E-Discovery*, 19 RICH. J.L. & TECH. 1, 33, 37 (2012); see Kate Mortensen, *E-discovery Best Practices for Your Practice, Step 4: Search and Review*, INSIDE COUNSEL, May 20, 2014, <http://www.insidecounsel.com/2014/05/20/e-discovery-best-practices-for-your-practice-step>, archived at <https://perma.cc/Q7JW-XTTZ>.

²¹ See Joseph H. Looby, *E-Discovery – Taking Predictive Coding Out of the Black Box*, FTI J. (Nov. 2012), <http://ftijournal.com/article/taking-predictive-coding-out-of-the-black-box-deleted>, archived at <https://perma.cc/4T49-CRTS>.

²² See Mark F. Foley, *Expert Testimony May Be Needed for E-Discovery Keyword Searches*, VONBREISEN, Mar. 1, 2008, <http://www.vonbriesen.com/legal->

disagreement.²³ The danger is that the inputted terms for searching might be “over- or underinclusive, either returning large amounts of irrelevant documents or failing to capture relevant ones.”²⁴ Consequently, “...the requesting party may ask for additional search terms or request that the producing party takes steps to verify the completeness of production.”²⁵

[12] Since predictive coding would be employed under each of the two methods, the costs associated with each are not significantly different. The majority of costs associated with predictive coding come from: the time of a senior attorney who is intimately familiar with the case, the cost of employing a company that has the available technology and software to run predictive coding, and the time associated with training the software to identify relevant documents.²⁶ These three categories of costs will be incurred regardless of which of the two methods is employed.

[13] The point at which the monetary costs and time spent may vary between the two methods is a senior attorney’s identification of potentially relevant documents or training of the software on a larger universe of documents. In predictive coding, there may be a larger universe of potentially relevant documents, simply because the software is more accurate in predicting which documents may be potentially

news/2098/expert-testimony-may-be-needed-for-e-discovery-keyword-searches, *archived at* <https://perma.cc/2TGW-9KV9>.

²³ See Murphy, *supra* note 3, at 614.

²⁴ *Id.* at 615–16.

²⁵ *Id.* at 614–15.

²⁶ See Matt Miller, *Making Sure Your Predictive Coding Solution Doesn’t Cost More*, DISCOVERREADY BLOG, Apr. 30, 2013, <http://discoverready.com/blog/making-sure-your-predictive-coding-solution-doesnt-cost-more/>, *archived at* <https://perma.cc/ZH6T-CZFN>.

relevant.²⁷ Keyword culling, on the other hand, eliminates many documents, even if they may be potentially relevant.²⁸ The reason is that the method of searching by keywords does not have that “predictive” feature; it merely eliminates any documents that do not contain the inputted words and phrases.²⁹ Accordingly, the cost differential between these two methods is not in the cost of the technology of predictive coding, but in the time it takes to identify the potentially relevant documents, as well as the resulting production of those documents.

[14] In sum, both methods employ predictive coding but at different stages in the discovery process. Predictive coding at the outset is abundantly more accurate than applying predictive coding after keyword culling.³⁰ The main costs associated with predictive coding will be the same, but since predictive coding at the outset is applied to more documents than keyword-culled documents, there may be additional time spent in training the software.³¹ Therefore, the actual cost of predictive coding will likely be substantially equal in both methods since the majority of the costs will be incurred in both methods.

[15] The remainder of this paper will discuss how this trade-off between accuracy and efficiency has been approached by several courts, litigating parties, and the Federal Rules of Civil Procedure in choosing the appropriate time to apply predictive coding.

²⁷ *See id.*

²⁸ *See Eidelman, supra* note 11.

²⁹ *Id.*

³⁰ *See id.*

³¹ *See Miller, supra* note 26.

III. COURT DECISIONS AND THE NEW FEDERAL RULES

[16] This Part will first address how courts have dealt with the issue, which developed a split in court decisions between applying predictive coding at the outset versus applying it on keyword-culled documents. Second, this Part will describe the recent amendments to the Federal Rules of Civil Procedure, as well as the subsequent reactions of courts and scholars.

A. Court Decisions under the Old Rules

[17] Prior to the recent amendments to the Federal Rules of Civil Procedure, parties and courts were aware of the concept of proportionality, but there have been various outcomes in different cases. In the past few years, the split in authority regarding the timing of predictive coding has spurred important realizations of accuracy and efficiency. The discussion below will reveal that some courts encouraged predictive coding at the outset, while some have allowed defendants to employ keyword culling first. These perspectives often depend on what the parties had mutually agreed on, what the parties had already accomplished, and the specific issue in the case. The arguments for each method are usually party-driven, as requesting parties argue for a broader scope of discovery to find the maximum amount of relevant documents, whereas producing parties tend to argue for a narrower scope of discovery to produce fewer documents.³²

1. Ex-Ante Permissibility of Predictive Coding

[18] Courts have routinely found that the application of predictive coding at the outset is appropriate. For example, in the 2012 landmark decision of *Da Silva Moore v. Publicis Groupe SA*, the court of the Southern District of New York found that predictive coding at the outset was appropriate.³³ The discovery issue in this case was whether predictive coding should be used at the outset, compared to other methods of discovery, including keyword culling.³⁴ The defendants had gathered approximately three million emails, a sizable amount of documents.³⁵

³² See, e.g., *In Re Biomet M2a Magnum Hip Implant Prods. Liab. Litig.*, No. 3:12MD2391, 2013 U.S. Dist. LEXIS 84440, at *1 (N.D. Ind. Apr. 18, 2013) (Order Regarding Discovery of ESI) (noting that the requesting party expected about 10 million documents, but the producing party only produced 2.5 million documents).

³³ See *Da Silva Moore v. Publicis Groupe & MSL Group*, 287 F.R.D. 182, 193 (S.D.N.Y. 2012).

³⁴ See *id.* at 184–85.

[19] The defendants sought to use predictive coding, and although the plaintiffs voiced their concerns, the plaintiffs were not opposed to predictive coding.³⁶ Magistrate Judge Peck allowed the use of predictive coding and emphasized the concept of proportionality from the Federal Rules of Civil Procedure.³⁷ Subsequently, the plaintiffs raised objections, which fell under the purview of the district judge.³⁸ The district judge found that the magistrate judge's decision was not clearly erroneous, denied the plaintiffs' objections, and accordingly adopted the magistrate judge's opinion.³⁹ The district judge noted that "the use of the predictive coding software as specified in the ESI protocol is more appropriate than keyword searching."⁴⁰ In this case, the defendants used, and the court allowed, predictive coding at the outset instead of keyword culling.

[20] A circuit court in Virginia upheld a similar ruling in *Global Aerospace, Inc. v. Landow Aviation, L.P.* in the same year.⁴¹ The court addressed whether the defendants would be permitted to use predictive coding at the outset instead of keyword culling. The defendants urged for the application of predictive coding at the outset instead of keyword

³⁵ *See id.* at 184.

³⁶ *See id.* at 184–86.

³⁷ *See id.* at 186, 188.

³⁸ *See Da Silva Moore v. Publicis Groupe SA*, No. 11 Civ. 1279, 2012 U.S. Dist. LEXIS 58742, at *2 (S.D.N.Y. Apr. 25, 2012).

³⁹ *See id.* at *8–9.

⁴⁰ *Id.* at *8.

⁴¹ *See Global Aerospace Inc. v. Landow Aviation, L.P.*, No. CL 61040, 2012 Va. Cir. LEXIS 50, at *2 (Va. Cir. Ct. Apr. 23, 2012).

culling.⁴² Although the plaintiffs objected to the use of predictive coding at the outset,⁴³ the judge allowed it, stating that the defendants “shall be allowed to proceed with the use of predictive coding for purposes of the processing and production of electronically stored information.”⁴⁴

[21] Similar to the rulings in *Da Silva Moore* and *Global Aerospace, Inc.*, the court in *In Re Actos (Pioglitazone) Products Liability Litigation* also allowed the parties to employ predictive coding at the outset.⁴⁵ The parties worked together and collaborated in choosing which method to employ. The high level of transparency and cooperation between the parties enabled the successful implementation of predictive coding at the outset on the entire universe of documents.⁴⁶ The parties agreed to review document samples collaboratively, meet and confer, and reveal their respective methodologies to each other.⁴⁷ The court allowed the parties to proceed in this manner because it was a mutually agreed upon method and proportional under the Rules.⁴⁸

⁴² See Brief in Opposition of Plaintiffs, Motion for Protective Order Regarding Electronic Documents and “Predictive Coding” at 2, *Global Aerospace Inc. v. Landow Aviation, L.P.*, No. CL 61040, 2012 Va. Cir. LEXIS 50 (Va. Cir. Ct. Apr. 16, 2012), 2012 WL 1419848, at *1–2.

⁴³ See *id.* at *2–3.

⁴⁴ *Global Aerospace Inc.*, 2012 Va. Cir. LEXIS 50, at *2.

⁴⁵ See *In re Actos (Pioglitazone) Prods. Liab. Litig.*, No. 6:11-MD-2299, 2012 U.S. Dist. LEXIS 187519, at *20, *34 (W.D. La. July 27, 2012).

⁴⁶ See *id.* at *20.

⁴⁷ See *id.* at *21.

⁴⁸ See *id.* at *43.

[22] A slightly different case reveals a court's hesitation in applying simplistic keyword searches. In *McNabb v. City of Overland Park*, the defendant produced about 20,000 e-mails after it unilaterally redacted the information that it thought was "confidential or irrelevant."⁴⁹ The plaintiff also submitted a list of about thirty-five search terms for the defendant to use, but the defendant argued that the requests were "overly broad and would encompass a significant number of documents."⁵⁰ The court agreed with the defendant and denied the plaintiff's motion, on grounds of proportionality. In other words, the court denied the implementation of these broad, general keyword searches.⁵¹ The motion papers in this case indicate "that the parties considered using predictive coding[.]" but the defendant decided not to.⁵² The outcome may have been different if the parties agreed to employ predictive coding at the outset because the plaintiff may have received more of the relevant data it was searching for, and the defendant may have been able to protect other documents as well.⁵³

[23] Overall, when presented with the issue at the outset, courts have routinely held that predictive coding is appropriate. The courts in *Da Silva Moore v. Publicis Groupe SA*, *Global Aerospace, Inc. v. Landow Aviation, L.P.*, and *In Re Actos* all allowed the parties to proceed with the

⁴⁹ See *McNabb v. City of Overland Park*, No. 12-CV-2331 CM/TJJ, 2014 U.S. Dist. LEXIS 37312, at *5 (D. Kan. Mar. 21, 2014).

⁵⁰ See *id.* at *2.

⁵¹ See Adam Kuhn, *The Interplay Between Proportionality and Predictive Coding in e-Discovery*, RECOMMIND, June 12, 2014, <http://www.recommind.com/blog/interplay-proportionality-predictive-coding-ediscovery>, archived at <https://perma.cc/LQX8-HYQM> [hereinafter *Interplay Between Proportionality and Predictive Coding*].

⁵² *Id.*

⁵³ See *id.*

application of predictive coding at the outset.⁵⁴ The judge's reluctance and refusal to allow simplistic keyword searches in *McNabb* also points in the same direction, suggestive of the possibility that predictive coding may have been an appropriate approach from the outset.⁵⁵ Accordingly, parties and courts have been supportive of the use of predictive coding at the outset.

2. Ex-Post Permissibility of Keyword Culling

[24] Courts have only permitted the use of predictive coding on previously keyword-culled documents after the fact, meaning after the documents had *already* been culled. In one example, the Northern District of Illinois court allowed the defendants to first employ keyword culling in *Kleen Products, LLC v. Packaging Corporation of America* in 2012.⁵⁶ The defendants had already produced "more than three million pages of documents" through keyword culling,⁵⁷ but plaintiffs requested the judge to order redoing discovery by employing predictive coding at the outset instead.⁵⁸ After several months of disputing these discovery issues, the

⁵⁴ See *Da Silva Moore v. Publicis Groupe SA*, 287 F.R.D. 182, 193 (S.D.N.Y. 2012); see *Global Aerospace Inc. v. Landow Aviation, L.P.*, No. CL 61040, 2012 Va. Cir. LEXIS 50, at *2 (Va. Cir. Ct. Apr. 23, 2012); *In re Actos (Pioglitazone) Prods. Liab. Litig.*, No. 6:11-MD-2299, 2014 U.S. Dist. LEXIS 187519, at *12, *20 (W.D. La. July 27, 2012) (Case Management Order: Protocol Relating to the Production of Electronically Stored Information).

⁵⁵ See *McNabb v. City of Overland Park*, No. 12-CV-2331 CM/TJJ, 2014 U.S. Dist. LEXIS 52534, at *7, *9 (D. Kan. Apr. 16, 2014).

⁵⁶ See *Murphy*, *supra* note 3, at 629 (noting that the district judge allowed the discovery issue to be decided separately by the magistrate judge).

⁵⁷ *Id.* at 629–30 (citing the Joint Status Conference Report No. 3, at 3, *Kleen Prods., LLC v. Packaging Corp. of Am.*, Civil Case No. 1:10-cv-05711 (N.D. Ill. May 17, 2012)).

⁵⁸ See *id.* at 630 (quoting Defendants' Brief on Discovery Issues at 1, *Kleen Prods., LLC v. Packaging Corp. of Am.*, No. 1:10-cv-05711 (N.D. Ill. Feb. 6, 2012)).

parties reached an agreement.⁵⁹ The plaintiffs withdrew their demand to restart and apply predictive coding at the outset on the entire universe of documents in the case.⁶⁰ In other words, the defendants kept the documents that were already culled down using keyword searches and were not required to restart the discovery process with predictive coding.⁶¹ The magistrate judge approved their agreement to employ keyword culling at the outset and restated Sedona Principle 6, “responding parties are best situated to evaluate” the appropriate method, with deference to the producing party.⁶²

[25] In the same year, the court in *In Re Biomet M2a Magnum Hip Implant Products Liability Litigation* also permitted keyword culling prior to the application of predictive coding.⁶³ The party had already employed keyword culling and reduced the universe of documents from “19.5 million to 3.9 million.”⁶⁴ The court stated that if the party was ordered to restart and apply predictive coding on the entire universe of documents, it

⁵⁹ *See id.*

⁶⁰ *See id.*

⁶¹ *See* Murphy, *supra* note 3, at 630–31.

⁶² *See* Matthew Verga, *Predictive Coding Cases, Part 2 – Kleen Products*, MODUS, Mar. 5, 2015, <http://discovermodus.com/blog/predictive-coding-cases-2-kleen-products/>, archived at <https://perma.cc/6PHG-D49Z>.

⁶³ *See In Re Biomet M2a Magnum Hip Implant Prods. Liab. Litig.*, No. 3:12MD2391, 2013 U.S. Dist. LEXIS 84440, at *1 (N.D. Ind. Apr. 18, 2013) (Order Regarding Discovery of ESI).

⁶⁴ Bob Ambrogi, *In Praise of Proportionality: Judge OKs Predictive Coding After Keyword Search*, CATALYST, Apr. 29, 2013, <http://www.catalystsecure.com/blog/2013/04/in-praise-of-proportionality-judge-oks-predictive-coding-after-keyword-search/>, archived at <https://perma.cc/2W7M-ZNHM>.

would not have been proportional under the previous version of Rule 26.⁶⁵ The court said this approach was reasonable under the circumstances.⁶⁶ The judge stated that the issue is not whether predictive coding is better than keyword culling, but whether the party satisfied its discovery obligations.⁶⁷ Furthermore, the judge stated that regardless of the other proportionality factors, the additional cost of going back to do the predictive coding on all documents would have outweighed the benefit of potentially finding more relevant documents.⁶⁸

[26] In a related line of cases, two courts have allowed keyword culling after the parties had agreed to it, but courts and parties have disagreed as to the proper approach after keyword culling. For example, in *Progressive Casualty Insurance Company v. Delaney*, the parties agreed to use keyword culling at the outset.⁶⁹ The producing party employed keyword culling which reduced the amount of documents from 1.8 million to 565,000.⁷⁰ For the remaining 565,000 documents, after employing keyword culling, the parties disagreed as to the appropriate method that

⁶⁵ See *Citing Proportionality, Court Declines to Require Defendant to Redo Discovery Utilizing Only Predictive Coding*, K&L GATES, Apr. 23, 2013, <https://www.ediscoverylaw.com/2013/04/citing-proportionality-court-declines-to-require-defendant-to-redo-discovery-utilizing-only-predictive-coding/>, archived at <https://perma.cc/5YUM-U6CY> (citing Order Regarding Discovery of ESI, *In Re Biomet M2a Magnum Hip Implant Prods. Liab. Litig.*) [hereinafter *Citing Proportionality*].

⁶⁶ See *Keyword Filtering Prior to Predictive Coding Deemed Reasonable*, EDISCOVERY WIRE, Dec. 6, 2013, <http://www.ediscoverywire.com/keyword-filtering-prior-to-predictive-coding-deemed-reasonable/>, archived at <https://perma.cc/P8S8-2WQZ>.

⁶⁷ See Ambrogi, *supra* note 64.

⁶⁸ See *id.*

⁶⁹ See *Progressive Cas. Ins. Co. v. Delaney*, No. 2:11-CV-00678-LRH-PAL, 2014 U.S. Dist. LEXIS 69166, at *5 (D. Nev. May 20, 2014).

⁷⁰ See *id.* at *6–7.

should be used.⁷¹ The producing party found that subsequently performing manual review would take a significant amount of time and money.⁷² To circumvent these costs, the producing party unilaterally chose to employ predictive coding instead of manual review on the remaining 565,000 documents.⁷³ After the producing party made this decision, it informed the requesting party, and the requesting party filed a motion to compel.⁷⁴ The court did not allow this change from manual review to predictive coding because it was not originally agreed upon by the parties, and it would result in more disputes and delays.⁷⁵ This case demonstrates that other disputes may arise after keyword culling is used because it calls into question the accuracy of subsequent methods. Predictive coding is contemplated but disagreed upon after keyword culling since the parties had already agreed upon manual review, although it is a time-consuming approach.⁷⁶ Instead, when predictive coding is used at the outset, these disputes are eliminated.

[27] Another example in which keyword culling was permitted at the outset is in *Bridgestone Americas, Inc. v. International Business Machines Corp.*⁷⁷ The plaintiff had already employed keyword culling and wanted to

⁷¹ *See id.*

⁷² *See id.* at *6.

⁷³ *See id.*

⁷⁴ *See Progressive Cas. Ins. Co.*, 2014 U.S. Dist. LEXIS 69166, at *3–4.

⁷⁵ *See id.* at *31.

⁷⁶ *See id.*

⁷⁷ *See Bridgestone Ams., Inc. v. Int'l Bus. Machs. Corp.*, No. 3:13-1196, 2014 U.S. Dist. LEXIS 142525, at *3 (M.D. Tenn., July 24, 2014) (Order Regarding use of Predictive Codes in Discovery) (explaining that the Magistrate Judge may permit the Plaintiff to use predictive coding on the documents).

proceed to use predictive coding. The defendant argued it would be unfair for the plaintiff to use predictive coding after documents had already been keyword culled, relying on *Progressive Casualty Insurance Company*.⁷⁸ However, because of concerns regarding proportionality and efficiency, the judge allowed the use of predictive coding on the previously keyword-culled documents.⁷⁹ This case also stands for the proposition that the parties should be the ones to try to resolve this issue.⁸⁰ The court believed that the use of keyword culling prior to predictive coding can be appropriate under Rule 26, but it depends on many factors, including “the type of data, the value of the case juxtaposed to the cost of using advanced analytics, and other factors that are matter specific.”⁸¹

[28] As demonstrated by *Bridgestone Americas, Inc.* and *Progressive Casualty Insurance Company*, when parties agree on keyword culling at the outset, parties and courts are left confused as to the appropriate method to use going forward to review the remaining documents. The reason is that the accuracy of the remaining relevant documents is already called

⁷⁸ See Adam Kuhn, *Bridgestone v. IBM Approves Predictive Coding Use, Rejects Progressive*, RECOMMIND, Aug. 12, 2014, <http://www.recommind.com/blog/bridgestone-v-ibm-approves-predictive-coding-use-rejects-progressive>, archived at <https://perma.cc/NXY6-JX64>.

⁷⁹ See *Bridgestone Ams., Inc.*, 2014 U.S. Dist. LEXIS 142525, at *3.

⁸⁰ See Gilbert S. Keteltas, *Predictive Coding After Keyword Screening!? Don't Miss the Point of Bridgestone Americas*, BAKERHOSTETLER: DISCOVERY ADVOCATE, Aug. 21, 2014, <http://www.discoveryadvocate.com/2014/08/21/predictive-coding-after-keyword-screening-dont-miss-the-point-of-bridgestone-americas/>, archived at <https://perma.cc/YTR5-9UGX>.

⁸¹ Jason Bonk, *Reasonableness and Proportionality Win Another Fight for Predictive Coding*, E-DISCOVERY L. REV. (Sept. 17, 2014), <http://www.ediscoverylawreview.com/2014/09/17/reasonableness-and-proportionality-win-another-fight-for-predictive-coding/>, archived at <https://perma.cc/98EY-ASU4> (quoting Eric Seggebruch).

into question since keyword culling is not as accurate as predictive coding.⁸² Furthermore, concerns of time, cost, and efficiency going forward in deciding between manual review and predictive coding become prominent issues for the parties.

[29] All four of these cases share a common denominator of one part of their holding regarding the discovery issue.⁸³ All four courts in *Kleen Products, LLC, In Re Biomet, Progressive Casualty Insurance Company, and Bridgestone Americas, Inc.* permitted the parties to employ keyword culling at the outset only after they had already performed keyword culling, or after it was already agreed upon by the parties.⁸⁴ Although the parties disagreed as to the proper method to apply after keyword culling was employed,⁸⁵ the courts found that ordering the parties to restart discovery and employ predictive coding would have been disproportional under the Rules.⁸⁶

⁸² See discussion *supra* Part II.B.

⁸³ See Edward Schoenecker Jr., *Nine Cases on Predictive Coding from Modus*, LINKEDIN, April 14, 2015, <https://www.linkedin.com/pulse/nine-cases-predictive-coding-from-modus-edward-schoenecker>, archived at <https://perma.cc/N4ZY-VCRW>.

⁸⁴ See *Bridgestone Ams., Inc. v. Int'l Bus. Machs. Corp.*, No. 3:13-1196, 2014 U.S. Dist. LEXIS 142525, at *1–2 (M.D. Tenn., July 24, 2014) (Order Regarding use of Predictive Codes in Discovery); see also *Progressive Cas. Ins. Co. v. Delaney*, No. 2:11-CV-00678-LRH-PAL, 2014 U.S. Dist. LEXIS 69166, at *31 (D. Nev. May 20, 2014); see *In Re Biomet M2a Magnum Hip Implant Prods. Liab. Litig.*, No. 3:12MD2391, 2013 U.S. Dist. LEXIS 84440, at *1 (N.D. Ind. Apr. 18, 2013) (Order Regarding Discovery of ESI); see *Kleen Prods. LLC v. Packaging Corp. of Am.*, No. 10 C 5711, 2012 U.S. Dist. LEXIS 139632, at *14–19 (N.D. Ill. Sept. 28, 2012).

⁸⁵ See *Bridgestone Ams., Inc.*, 2014 U.S. Dist. LEXIS 142525, at *1–2; see *Progressive Cas. Ins. Co.*, 2014 U.S. Dist. LEXIS 69166, at *31.

⁸⁶ See *Bridgestone Ams., Inc.*, 2014 U.S. Dist. LEXIS 142525, at *5; see *Kleen Prods., LLC*, 2012 U.S. Dist. LEXIS 139632, at *28.

B. Reinforcement of Court Decisions under the New Rules

[30] Recently, the drafters of the Federal Rules of Civil Procedure and the Supreme Court rebalanced the priorities of discovery and set a legislative-like answer in the amendments to the Rules. This Part discusses those amendments, as well as the subsequent reactions of courts and scholars.

1. Recent Amendments to the Rules

[31] The Federal Rules of Civil Procedure were recently amended and deemed effective as of December 1, 2015. The new revisions can be found in the 2016 edition of the Federal Rules of Civil Procedure.⁸⁷ Many rules were amended, but the revisions to Rules 1 and 26 directly impact this discussion. Through these revisions, the rule drafters and the Supreme Court chose to highlight proportionality, as well as the responsibility of parties and courts in making these decisions.

[32] Rule 1 was amended to emphasize that parties are just as responsible as courts in applying the Federal Rules of Civil Procedure to ensure the efficiency of every action in a case.⁸⁸ The previous version of Rule 1 stated that the rules “should be construed and administered to secure the just, speedy, and inexpensive determination of every action and

⁸⁷ See *2015-2016 Federal Rules of Civil Procedure Amendments Released*, FEDERAL RULES OF CIVIL PROCEDURE UPDATES, May 13, 2015, <https://www.federalrulesofcivilprocedure.org/2015-2016-federal-rules-of-civil-procedure-amendments-released/>, archived at <https://perma.cc/54GY-2XKK> [hereinafter *2015-2016 Federal Rules Amendments*]

⁸⁸ *Id.*; see also *Federal Rule Changes Affecting E-Discovery Are Almost Here - Are You Ready This Time?*, K&L GATES, Oct. 1, 2015, <http://www.ediscoverylaw.com/wp-content/uploads/2015/10/Rules-Amendment-Alert-100115.pdf>, archived at <https://perma.cc/H7A3-2C7T> [hereinafter *Rule Changes*].

proceeding.”⁸⁹ The new version of Rule 1 states that the rules “should be construed, administered, and employed by the court and the parties to secure the just, speedy, and inexpensive determination of every action and proceeding.”⁹⁰

[33] Rule 26 was amended to emphasize factors of proportionality in defining the scope of discovery.⁹¹ The previous version of Rule 26(b)(1) stated:

Scope in General. 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 persons who know of any discoverable matter. For good cause, the court may order discovery of any matter relevant to the subject matter involved in the action. Relevant information need not be admissible at the trial if the discovery appears reasonably calculated to lead to the discovery of admissible evidence. All discovery is subject to the limitations imposed by Rule 26(b)(2)(C).⁹²

[34] The amended version of Rule 26(b)(1) now states:

⁸⁹ FED. R. CIV. P. 1 (2014) (amended 2015).

⁹⁰ FED. R. CIV. P. 1 (emphasis added).

⁹¹ See 2015-2016 Federal Rules Amendments, *supra* note 87.

⁹² FED. R. CIV. P. 26(b)(1) (2014) (amended 2015).

Scope in General. 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 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 the discovery in resolving the issues, and whether the burden or expense of the proposed discovery outweighs its likely benefit. Information within this scope of discovery need not be admissible in evidence to be discoverable.⁹³

[35] The concept of proportionality appeared in Rule 26(b)(2)(C) in the previous version and has always been present; however, it now appears at the beginning of Rule 26(b)(1), which makes it more explicitly applicable to the entire scope of discovery.⁹⁴ Specifically, the proportionality factors moved from Rule 26(b)(2)(C)(iii) to the new location at the beginning of Rule 26(b)(1).⁹⁵ The Committee's intention in moving these factors is to "make them an explicit component of the

⁹³ FED. R. CIV. P. 26(b)(1) (emphasis added).

⁹⁴ See *Just Follow the Rules! FRCP Amendments Could be E-Discovery Game Changer*, METROPOLITAN CORPORATE COUNSEL (July 17, 2015, 11:49 PM), <http://www.metrocorpocounsel.com/articles/32726/just-follow-rules-frcp-amendments-could-be-e-discovery-game-changer>, archived at <https://perma.cc/A9U7-3CHY> [hereinafter *Just Follow the Rules!*].

⁹⁵ See *E-Discovery Update: Federal Rules of Civil Procedure Amendments Go into Effect*, MCGUIREWOODS, Dec. 1, 2015, <https://www.mcguirewoods.com/Client-Resources/Alerts/2015/12/E-Discovery-Update.aspx>, archived at <https://perma.cc/J5H6-4XET>.

scope of discovery, requiring parties and courts alike to consider them when pursuing discovery and resolving discovery disputes.”⁹⁶

[36] It is important to note that the Committee made revisions to the actual factors that pertain to proportionality as well. They amended the order of the factors; the “importance of the issues at stake” now precedes the “amount in controversy” which places an emphasis on proportionality *related to the issues, not only the dollar amount*.⁹⁷ They also added one additional factor: “the parties’ relative access to relevant information.”⁹⁸

[37] The other change to Rule 26 is the removal of the language “reasonably calculated to lead to the discovery of admissible evidence.”⁹⁹ This means that the previous guidance in discovery, to find evidence that might lead to admissible evidence, has been taken out. Since it is no longer a requirement to potentially lead to admissible evidence, there may be a push from attorneys to narrow the scope of discovery.¹⁰⁰ The reason is that the previous requirement did not require a direct nexus to the case as discoverable evidence only had to

⁹⁶ *Rule Changes*, *supra* note 88, at 2 (quoting THE COMMITTEE ON RULES OF PRACTICE AND PROCEDURE, REPORT OF THE JUDICIAL CONFERENCE COMMITTEE ON RULES OF PRACTICE AND PROCEDURE TO THE CHIEF JUSTICE OF THE UNITED STATES AND MEMBERS OF THE JUDICIAL CONFERENCE OF THE UNITED STATES app. at B-8 (2014), <http://www.uscourts.gov/rules-policies/archives/committee-reports/reports-judicial-conferencesseptember-2014>).

⁹⁷ *Just Follow the Rules!*, *supra* note 94 (arguing that although a case may not have an amount in controversy, it could still be a significant issue that deserves the concern of proportionality, such as discrimination or First Amendment cases).

⁹⁸ *Rule Changes*, *supra* note 88, at 2.

⁹⁹ FED. R. CIV. P. 26(b)(1).

¹⁰⁰ *See Just Follow the Rules!*, *supra* note 94.

potentially lead to other admissible evidence. In this application, it might be a call to highlight the most relevant evidence in discovery.

[38] In sum, Rule 1 now explicitly makes it the priority of parties and courts to ensure that a case proceeds in a just and expedient manner. Rule 26 now explicitly prioritizes proportionality to dictate the scope of discovery. Both of these rules impact the decision of when it is the right time to apply predictive coding for several reasons. Predictive coding and keyword culling, as discussed above, have important implications regarding the accuracy and efficiency of the discovery process.

2. Subsequent Reactions to the New Rules

[39] Courts have begun to apply these recent amendments of the Federal Rules of Civil Procedure, and there has not been a drastic change in the past few months. Many courts are finding that the priority of proportionality has been present since the prior version of the Rules, but the courts are able to more easily point to this priority as it is explicitly referred to first in Rule 26 regarding the scope of discovery.

[40] For instance, just six days after the amendments went into effect, the court in *Carr v. State Farm Mutual Automobile Insurance Co.* found that the burdens on the parties have not fundamentally changed.¹⁰¹ In that case, the defendant's motion to compel was granted since the burden on the plaintiff to resist the motion to compel had not changed under the new rules, as evidenced by the Committee's notes on the amendments.¹⁰²

¹⁰¹ See *Court Applies Amended Rule 26 Concludes Burdens on Parties Resisting Discovery have not Fundamentally Changed*, K&L GATES, Dec. 17, 2015, <http://www.ediscoverylaw.com/2015/12/court-applies-amended-rule-26-concludes-burdens-on-parties-resisting-discovery-have-not-fundamentally-changed/>, archived at <https://perma.cc/A8W8-QQRK>.

¹⁰² See *Carr v. State Farm Mut. Auto. Ins. Co.*, No.3:15-CV-1026-M, 2015 U.S. Dist. LEXIS 163444, at *15–17 (N.D. Tex. Dec. 7, 2015).

Another court has concluded the application of predictive coding was disproportional under the new rules in *Gilead Sciences, Inc. v. Merck & Co., Inc.*, but it stated that the result would have been the same even under the prior version of the Rules.¹⁰³ In that patent infringement case, the defendant's motion to compel additional discovery was denied because the plaintiff would have needed to produce an excessive amount of information regarding the contents of tubes of compounds that were not at issue in the case.¹⁰⁴

[41] The court stated that the amendments now first require an inquiry into whether the additional discovery would be proportional, rather than whether it might lead to something admissible.¹⁰⁵

[42] Similarly, the court of the Southern District of Florida allowed the defendants to redact information that was irrelevant from documents that were considered responsive.¹⁰⁶ The court based its opinion on the concept of proportionality in Rule 26.¹⁰⁷

¹⁰³ See *Court Concludes Defendant's Request was "Precisely the Kind of Disproportionate Discovery That Rule 26—Old Or New—Was Intended to Preclude,"* K&L GATES, Jan. 19, 2016, <https://www.ediscoverylaw.com/2016/01/court-concludes-defendants-request-was-precisely-the-kind-of-disproportionate-discovery-that-rule-26-old-or-new-was-intended-to-preclude/>, archived at <https://perma.cc/V8T8-WJHG> (citing *Gilead Sciences, Inc. v. Merck & Co., Inc.*, No. 5:13-CV-04057-BLF, 2016 U.S. Dist. LEXIS 5616 (N.D. Cal. Jan. 13, 2016)) [hereinafter *Court Concludes*].

¹⁰⁴ *Gilead Sciences, Inc. v. Merck & Co., Inc.*, No. 5:13-CV-04057-BLF, 2016 U.S. Dist. LEXIS 5616, at *7 (N.D. Cal. Jan. 13, 2016).

¹⁰⁵ See *Court Concludes*, *supra* note 103 (citing *Gilead Sciences, Inc. v. Merck & Co., Inc.*, No. 5:13-CV-04057-BLF, 2016 U.S. Dist. LEXIS 5616 (N.D. Cal. Jan. 13, 2016)).

¹⁰⁶ See *Court Approves Proposal to Redact or Withhold Irrelevant Information from Responsive Documents and Document Families*, K&L GATES, Mar. 3, 2016, <http://www.ediscoverylaw.com/2016/03/court-approves-proposal-to-redact-or-withhold-irrelevant-information-from-responsive-documents-and-document-families/>, archived at <https://perma.cc/26E4-UXLH> (citing *In re Takata Airbag Prods. Liab. Litig.*, 2016 U.S. Dist. LEXIS 131746 (S.D. Fla. Mar. 1, 2016)).

[43] The Year-End Report of the Federal Judiciary argues that the amendments have had a profound impact on the expected efficiency of parties and courts.¹⁰⁸ Magistrate Judge John M. Facciola believes the Rules were significantly modified in that the scope of discovery does not regard whether an item is “reasonably calculated to lead to the discovery of admissible evidence,”¹⁰⁹ but rather regards the issues at stake and proportionality concerns.¹¹⁰ Because of this, lawyers may argue to narrow the scope of discovery.¹¹¹

[44] The courts that have begun to apply the new amendments to the Rules are finding that the outcome would have been similar even under the old Rules. The courts are only able to more easily point to the primary concerns of proportionality, justness, and expediency through the new amendments.

IV. ENCOURAGING PREDICTIVE CODING EX ANTE

[45] In light of the court decisions and recent amendments to the Federal Rules of Civil Procedure, predictive coding should be encouraged at the outset of the discovery process to be applied on the entire universe of documents in a case. This Part will first explain the

¹⁰⁷ *See id.*

¹⁰⁸ *See 2015 Year-End Report on the Federal Judiciary*, SUPREMECOURT.GOV 1, 6, 9 (Dec. 31, 2015), <http://www.supremecourt.gov/publicinfo/year-end/2015year-endreport.pdf>, archived at <https://perma.cc/5RU7-DCF7>.

¹⁰⁹ *Just Follow the Rules!*, *supra* note 94.

¹¹⁰ *See id.*

¹¹¹ *See id.*

reasons why predictive coding should be used at the outset, and second, it will suggest how parties and courts should proceed in implementing this method.

A. Why Predictive Coding Ex Ante is Preferable

[46] Employing predictive coding at the outset provides significantly more accurate results in identifying relevant documents than keyword culling.¹¹² Predictive coding employs sophisticated technology which can more accurately predict relevant documents, beyond the simplistic search terms used in keyword culling.¹¹³ The method of keyword culling is not as accurate because many relevant documents slip through the cracks when keyword searches are employed.¹¹⁴ In terms of accuracy, predictive coding is significantly more accurate than keyword culling when used on the entire set of documents at the outset.

[47] Since predictive coding would be employed under each of the two methods, the costs associated with either method are not significantly different. The majority of costs associated with predictive coding come from the time of a senior attorney who is intimately familiar with the case training the software, and the cost of employing a company that has the available technology and software to run predictive coding.¹¹⁵ However, these costs will be expended in both methods since predictive coding is used in both methods. The point at which the monetary costs and time spent may vary between the two methods is in the senior attorney identifying potentially relevant documents and

¹¹² See Eidelman, *supra* note 11; see also Kazan & Wilson, *supra* note 12.

¹¹³ See Eidelman, *supra* note 11; see also Kazan & Wilson, *supra* note 12.

¹¹⁴ See Eidelman, *supra* note 11; see also Kazan & Wilson, *supra* note 12.

¹¹⁵ See Miller, *supra* note 26.

training the software on a larger volume of documents.¹¹⁶ Accordingly, the cost differential between these two methods is in the time it takes to identify these potentially relevant documents, as well as the resulting production of documents. There has not been enough empirical research done on this inquiry, but no courts have held, and no parties have argued, that predictive coding would cost more at the outset. Although there is currently no proof that the costs are steeper, even if that were the case, it is likely not substantial enough to outweigh the benefit of accuracy in identifying relevant documents.

[48] Furthermore, as discussed in Part III.A, courts have routinely upheld and encouraged the use of predictive coding at the outset. The courts that held keyword culling is permissible at the outset only found it permissible after the documents had already been keyword culled, and found it too burdensome and costly to restart discovery.¹¹⁷

[49] The recent amendments to the Federal Rules of Civil Procedure further reinforce the concepts of proportionality and the responsibilities of the parties and courts to ensure the just and efficient resolution of a case. Rule 1 now mandates that the rules “should be construed, administered, and employed by *the court and the parties* to secure the just, speedy, and inexpensive determination of every action and proceeding.”¹¹⁸ There is now an explicit emphasis on both courts and the parties to work justly and efficiently all throughout a case from the beginning to the end, which includes the discovery phase. More specifically, Rule 26(b) now highlights that the scope of discovery must begin with an inquiry of proportionality.¹¹⁹ The Rule mandates that the parties and courts consider

¹¹⁶ *See id.*

¹¹⁷ *See* discussion, *supra* Part III.A.

¹¹⁸ FED. R. CIV. P. 1 (emphasis added).

¹¹⁹ *See* FED. R. CIV. P. 26(b)(1).

several factors of proportionality, including “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 the discovery in resolving the issues, and whether the burden or expense of the proposed discovery outweighs its likely benefit.”¹²⁰

[50] The Rules explicitly emphasize proportionality with a list of many factors. This legislative-like answer set by the rules’ drafters and the Supreme Court was a deliberate decision to refocus the attention of discovery to the issues at stake as well as the importance of discovery in finding a resolution to those issues. As discussed above, the cost differential between both methods is likely insignificant. Proportionality, as applied in a discovery issue, concerns both accuracy and efficiency because it impacts time, cost, and the just resolution of a case. Since cost is not a determinative factor, the parties will gain accuracy in employing predictive coding at the outset, which is particularly proportional in the scope of discovery under the Rules. In this way, the parties gain accuracy without sacrificing efficiency.

B. How Parties and Courts Should Proceed

[51] At the beginning of discovery, parties should opt to employ predictive coding on the entire universe of documents in a case, in light of the benefits regarding accuracy and proportionality. Even under the previous version of the Rules, parties were encouraged to collaborate regarding discovery methods and to consider each step of predictive coding at the outset.¹²¹ This collaboration is essential because the parties

¹²⁰ *Id.*

¹²¹ See Karl Schieneman & Thomas C. Gricks III, *The Implications of Rule 26(g) on the Use of Technology-Assisted Review*, 7 FED. CTS. L. REV. 239, 273–74 (2013) (noting that even under the old Rules, counsel was encouraged to consider each step of technology-assisted review under Rule 26(g) and 26(b)(2)(C)(iii)).

are usually the ones that are in the best position to initially evaluate the method rather than courts.¹²²

[52] The ideal protocol is that which was employed by the parties in *In Re Actos*.¹²³ In that case, the parties cooperated and collaborated at the beginning of the discovery phase and were able to successfully implement predictive coding.¹²⁴ At the opposite end of the spectrum, the parties in *Kleen Products*, demonstrated how destructive it was to dispute the methodology of discovery for several months, wasting both time and money on the dispute.¹²⁵ Further, the plaintiffs withdrew their demand which allowed the defendants to keep their previously keyword-culled documents.¹²⁶ This end result of accepting the keyword-culled documents was not a judicial decision, nor was it a collaborative effort by the parties. Rather, it was the easier solution after several months of dispute, and a result that was brought on by the plaintiffs' withdrawal of the demand.¹²⁷ If parties are encouraged to collaborate at the outset and practice transparency by sharing the predictive coding methodology with the other party, there is little left for the other party to object to.¹²⁸ The reason is that costs are already being saved by employing predictive coding regardless

¹²² See Charles Yablon & Nick Landsman-Roos, *Predictive Coding: Emerging Questions and Concerns*, 64 S.C. L. REV. 633, 674 (2013) (citing Sedona Principle 6).

¹²³ See *In re Actos (Pioglitazone) Prods. Liab. Litig.*, No. 6:11-MD-2299, 2012 U.S. Dist. LEXIS 187519, at *27 (W.D. La. July 27, 2012).

¹²⁴ See *id.* at *27.

¹²⁵ See *Kleen Prods. LLC v. Packaging Corp. of Am.*, No. 10-C-5711, 2012 U.S. Dist. LEXIS 139632, at *60–62 (N.D. Ill. Sept. 28, 2012).

¹²⁶ See *id.* at *62–63.

¹²⁷ See *id.* at *58, *62.

¹²⁸ See *id.* at *58.

of the time at which it is applied, and the method of employing predictive coding is overwhelmingly more accurate in producing relevant documents than keyword culling.¹²⁹

[53] Subsequently, all that is left that the parties may dispute is the input to the predictive coding software. Parties may disagree about the inputs in training the software, but it does not have to be a daunting task, as the parties in *In Re Actos* planned for that and allowed options to work together on the inputs and scheduled for times to meet and confer.¹³⁰ Therefore, it is more proportional and worthwhile to start with predictive coding at the outset.¹³¹

[54] The courts in *McNabb* and *Progressive Casualty Insurance Company* also teach an important lesson about the importance of collaboration between the parties at the outset.¹³² Since the court in *McNabb* rejected the plaintiff's motion to compel and employ further keyword searches,¹³³ the parties could have both benefitted from predictive coding at the outset. The producing party in *Progressive Casualty Insurance Company* unilaterally decided to switch to predictive

¹²⁹ See Grossman & Cormack, *supra* note 1, at 44, 48.

¹³⁰ See *In re Actos (Pioglitazone) Prods. Liab. Litig.*, No. 6:11-MD-2299, 2014 U.S. Dist. LEXIS 86101, at *20–34 (W.D. La. June 23, 2014).

¹³¹ See *Interplay Between Proportionality and Predictive Coding*, *supra* note 51 (“[A] party who unilaterally decides later on in discovery that its search tactics were too imprecise could find that proportionality standards prevent the use of more advanced, accurate, and targeted searches with predictive coding technologies.”).

¹³² See, e.g., *McNabb v. City of Overland Park*, No. 12-CV-2331 CM/TJJ, 2014 U.S. Dist. LEXIS 37312, at *2–14 (D. Kan. Mar. 21, 2014); see also *Progressive Cas. Ins. Co. v. Delaney*, No. 2:11-CV-00678-LRH-PAL, 2014 U.S. Dist. LEXIS 69166, at *30–32 (D. Nev. May 20, 2014).

¹³³ See *McNabb*, 2014 U.S. Dist. LEXIS 37312, at *5.

coding, which instigated a motion to compel from the requesting party.¹³⁴ These situations could have been avoided if there were collaborative efforts at the outset, as well as transparency throughout the process.

[55] As discussed in Part III.A.2, courts allowed predictive coding to be used after keyword culling, primarily because keyword culling had already been employed by the producing party, and it would have been costly to start over with predictive coding on the entire universe of documents in the case. The judges reasoned that it would have been highly inefficient and disproportional to require that party to start over at the beginning, especially if the parties agreed on the use of the keyword search method at the outset.¹³⁵ In *Kleen Products, LLC v. Packaging Corporation of America*, the “defendants [had] [already] produced more than three million pages of documents” through keyword culling,¹³⁶ but plaintiffs requested the judge to order redoing discovery using predictive coding.¹³⁷ The parties eventually reached an agreement, with the plaintiffs withdrawing their demand.¹³⁸ The court in *In Re Biomet M2a Magnum Hip Implant Products Liability Litigation* allowed keyword culling prior to the application of predictive coding because if the party was ordered to restart and apply predictive coding on the raw data, it would have been expensive and disproportional under Rule 26.¹³⁹

¹³⁴ See *Progressive Casualty*, 2014 U.S. Dist. LEXIS 69166, at *2, *30–32.

¹³⁵ See *id.* at *30–32.

¹³⁶ Murphy, *supra* note 3, at 629–30.

¹³⁷ See *id.* at 630.

¹³⁸ See *id.*

¹³⁹ See *Citing Proportionality*, *supra* note 65 (discussing the court’s decision in *Kleen* regarding ESI searches).

[56] As shown by these cases, producing parties continually employ keyword culling at the outset, possibly because it is quicker or because it produces a smaller amount of documents.¹⁴⁰ Regardless of the motive, once this discovery issue is before the courts and the producing party has already employed keyword culling, courts have been hesitant to order the party to start the discovery process again. In effect, the producing parties are permitted to retain their keyword culling methods.

[57] Courts need to lead the change. If the parties do not begin to employ predictive coding at the outset and continue to employ keyword culling, courts should suggest the use of predictive coding at the outset. It will be relatively simple for courts to encourage or mandate predictive coding at the outset, as the courts discussed in Part III.A did. Courts may be more reluctant to order a producing party to abandon its keyword culling and restart the discovery process to employ predictive coding at the outset, but at this point, it is necessary. Proportionality is a primary concern under the Federal Rules of Civil Procedure. When predictive coding will be used in a case, it should be used at the outset in order to obtain the most accurate documents. It may only take one court in one case to capture the attention of parties and other courts, in order to lead the change for a more accurate and proportional discovery process in the cases to come.

V. CONCLUSION

[58] Predictive coding has been proven to be more accurate and efficient than traditional methods of discovery. There has been a split in

¹⁴⁰ See Eidelman, *supra* note 11.

authority as to the point at which predictive coding should be applied. The issue that courts have been facing is whether predictive coding should be applied at the outset to the entire universe of documents, or if it should be applied to keyword-culled documents. Courts have gone both ways on this issue, but as of December 1, 2015, the drafters of the Federal Rules of Civil Procedure and the Supreme Court approved amendments to the Rules. Primarily, the amendments to Rules 1 and 26(b)(1) directly impact this discussion, as these rules emphasize the responsibility of parties and courts to ensure that a case proceeds justly and efficiently, while highlighting the importance of proportionality in the scope of discovery. Considering these amendments, predictive coding should be applied at the outset on the entire universe of documents in a case. The reason is that it is far more accurate, and is not more costly or time-consuming, especially when the parties collaborate at the outset. As seen in prior cases, this is the best method to identify more relevant documents. The point at which it becomes costly and inefficient is if a party had already used keyword culling and must restart the discovery process to employ predictive coding. However, if parties collaborate and participate in transparency at the outset, they will often find that it is significantly more effective and in the interest of both parties to employ predictive coding to identify the most relevant documents. If parties cannot agree or fall back on old ways of keyword culling, courts can and should lead the change by encouraging predictive coding at the outset of the discovery process, with the recent amendments to the Federal Rules of Civil Procedure on their side.