

Fall 2016

The Past, Present, and Future of Predictive Coding

Matthew G. Kenney

Follow this and additional works at: <https://commons.law.famu.edu/famulawreview>

 Part of the [Computer Law Commons](#), [Courts Commons](#), [Legal Ethics and Professional Responsibility Commons](#), [Litigation Commons](#), and the [Science and Technology Law Commons](#)

Recommended Citation

Matthew G. Kenney, *The Past, Present, and Future of Predictive Coding*, 12 Fla. A&M U. L. Rev. 165 (2016).

Available at: <https://commons.law.famu.edu/famulawreview/vol12/iss1/7>

This Note is brought to you for free and open access by Scholarly Commons @ FAMU Law. It has been accepted for inclusion in Florida A & M University Law Review by an authorized editor of Scholarly Commons @ FAMU Law. For more information, please contact paul.mclaughlin@famu.edu.

THE PAST, PRESENT AND FUTURE OF PREDICTIVE CODING

Matthew G. Kenney, D.B.A

INTRODUCTION	165
I. STATISTICAL AND TECHNOLOGICAL CONCEPTS	167
A. <i>Statistical Basis for Predictive Coding</i>	167
B. <i>Relevant Technological Concepts</i>	168
II. HISTORY OF PREDICTIVE CODING IN LEGAL PRACTICE	170
A. <i>Early Scholarly Analysis</i>	170
B. <i>Earliest Use of Predictive Coding in Courts</i>	171
III. PRESENT STATE OF PREDICTIVE CODING IN LEGAL PRACTICE	174
A. <i>Adoption of Predictive Coding by Courts</i>	174
B. <i>Reluctance to Accept Predictive Coding Among Attorneys</i>	176
C. <i>Predictive Coding: Ethical and Professional Considerations</i>	177
IV. FUTURE APPLICATION OF PREDICTIVE CODING.....	179
A. <i>Special Masters: Application of Rule 53</i>	179
B. <i>Predictive Coding and International Application</i>	179
C. <i>Confidentiality and Work-Product Conflict</i>	181
D. <i>Teaching the Next Generation of Lawyers</i>	182
CONCLUSION	182

INTRODUCTION

Electronic discovery, or e-discovery, refers to the discovery of electronically stored documents and images.¹ Examples of e-discovery related documentation would include email, digital versions of paper documents (e.g. MS Word, PDF, Excel, and PowerPoint), social media postings, digital photos, Global Positioning System data, and content within computerized databases, etc. Digital data stored on computers, smartphones, tape drives, hard-drives, portable digital storage devices and the like would fall under the domain of e-discovery.² Collecting and sorting massive amounts of electronically stored data presents both opportunities and challenges for lawyers.

1. *Discovery in a World of Electronic Documents and Data*, SEDONA CONF., <https://the.sedonaconference.org/node/4413> (last visited Aug. 18, 2017).

2. *Id.*

For context: In 2015, electronic discovery was a \$10.2 billion global industry. Of this amount, \$8.2 billion flowed to e-discovery service providers (e.g. document review by contract attorneys and vendors); and \$2 billion flowed to the development of new software. The worldwide e-discovery market is expected to grow at 9.4% annually through 2019.³

Predictive coding—a dimension of e-discovery—is a process whereby attorneys train computer programs to identify potentially relevant documents within a large body of documents.⁴ This process begins with the attorney(s) selecting a “seed set” of documents and choosing keywords relevant to the case.⁵ This seed set is then searched, and re-searched, in an iterative process until the software recognizes patterns.⁶ For practical purposes, it is simply telling the software what to find. However, attorneys need to have a general understanding of statistics and computer assisted technology prior to engaging in predictive coding.

Consider the conflict in *Da Silva Moore*, a seminal case this article will explore. The parties agreed to use predictive coding but could not agree as to the size of the seed set to be drawn from roughly three million emails.⁷ While the court held a random sample of 2,399 email documents would suffice, a statistical expert for the plaintiff contended that a 16,555 document sample would be needed to produce statistically reliable results.⁸

What if the statistics expert was correct? Choosing the correct sample size is essential when seeking statistical validity.⁹ A lawyer might have little reason—and likely no opportunity in law school—to learn about advanced statistical techniques, let alone perform them.

3. Ed Silverstein, *Global E-Discovery Market Exceeds Ten Billion in 2015, Continues to Grow*, L. TECH. NEWS, Jan. 5, 2016, at 1.

4. Wallis M. Hampton, *Predictive Coding: It's Here to Stay*, PRAC. L., 2014 WL 9-567-0325.

5. *Id.*

6. *Id.*

7. John M. Barkett, *More on the Ethics of E-Discovery: Predictive Coding and Other Forms of Computer-Assisted Review*, 2012 DUKE UNIVERSITY SCHOOL OF LAW 20, https://law.duke.edu/sites/default/files/centers/judicialstudies/TAR_conference/Panel_5-Original_Paper.pdf (last visited Aug. 18, 2017).

8. *Id.* at 22.

9. Jeehyoung Kim & Bong Soo Seo, *How to Calculate Sample Size and Why*, 5(3) CLINICS IN ORTHOPEDIC SURGERY 235, 23 (2013), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3758995/pdf/cios-5-235.pdf>.

Yet, determining the appropriate sample size (i.e. seed set) is of paramount importance in using predictive coding effectively.¹⁰ This article will briefly explore those principles, and the computer-assisted technology facilitating the statistical analysis.

The remainder of this article focuses on the practical application of predictive coding in the past and present; and the role predictive coding will play in the future. This research is of timely importance because predictive coding is increasingly accepted by courts due to its ability to speed-up the discovery process and reduce costs.¹¹

Judge Andrew J. Peck noted in the *Da Silva Moore* case: “Computer-assisted review appears to be better than the available alternatives, and thus should be used in appropriate cases. While the Court recognizes that computer assisted review is not perfect, the Federal Rules of Civil Procedure do not require perfection.”¹²

This article will illustrate why Judge Peck is correct in his assessment. Because the use of predictive coding will likely be suggested by judges overseeing complex litigation in the future, it is important for attorneys to gain a level of comfort with this emerging technology. Said comfort will satisfy their ethical obligations under ABA Rule 1.1—Competency—which the article also explores.

While the vast majority of e-discovery is currently performed within the United States, it has been projected that Europe will represent twenty-three percent of the global e-discovery market by 2019; and Asia will represent seven percent.¹³ Thus, this article will also briefly explore the future application of predictive coding in foreign courts.

I. STATISTICAL AND TECHNOLOGICAL CONCEPTS

A. *Statistical Basis for Predictive Coding*

In statistical analysis, a population is an entire group of things a researcher is interested in studying.¹⁴ Within populations there are subpopulations. When conducting research, it is common for those inexperienced in statistical methods to confuse population and

10. Bill Dimm, *Comments on Rio Tinto v. Vale and Sample Size*, CLUSTIFY BLOG (Mar. 5, 2015), <https://blog.cluster-text.com/tag/sample-size/>.

11. See Barkett, *supra* note 7, at 4.

12. *Da Silva Moore v. Publicis Groupe & MSL*, 287 F.R.D. 182, 191 (S.D.N.Y. 2012).

13. Silverstein, *supra* note 3, at 1.

14. Valerie J. Easton & John H. McColl, *Statistics Glossary*, U. OF GLASGOW, http://www.stats.gla.ac.uk/steps/glossary/basic_definitions.html (last visited June 3, 2016).

subpopulation. Often, this is because the subpopulation is referred to as “the population” from which the sample will be drawn.

Any quantitative research methods—including predictive coding—will generate better results when the population is narrowed into subpopulations¹⁵ (e.g. relevant documents and irrelevant documents). Assume opposing counsel delivers five million email documents, but 90% of these documents are irrelevant. The documents of interest are within a 500,000 subpopulation only. The “seed set” sample, therefore, should be culled only from this relevant 500,000 “subpopulation,” which is referred to—somewhat confusingly—as the “population.” While population and subpopulations are distinct concepts with separate meanings, they are used synonymously in predictive coding practice.

Attorneys must take caution: in predictive coding, success is dependent on the quality of the seed set.¹⁶ Stated differently, the quality of the seed set will directly affect the quality of the results.¹⁷ Like any research, when it comes to predictive coding; garbage in will get garbage out.¹⁸ Because the initial seed set is so important for gaining accurate results, attorneys may want to consider using a predictive coding algorithm that allows for both a manually produced seed set and a list of documents culled from random sampling.¹⁹ From the randomly selected results an attorney would use his/her legal judgment to identify relevance.²⁰ Keywords identified from the random documents would then be incorporated into the seed set, resulting in better results.

B. Relevant Technological Concepts

Predictive coding is but a dimension of machine learning.²¹ Predictive coding should not be viewed as glorified keyword searching.

15. Cal. State Univ. Long Beach, *PPA 696: Sampling*, CAL. ST. U. LONG BEACH <http://web.csulb.edu/~msaintg/ppa696/696sampl.htm#Stratified%20random%20sample> (last visited Aug. 29, 2017).

16. John M. Facciola & Philip J. Favro, *Safeguarding the Seed Set: Why Seed Set Documents are Entitled to Work Product Protection*, 8 FED. CTS. L. REV. 3, 14 (2015).

17. *Id.*

18. Equivio, *Top 10 Best Practices for Predictive Coding*, DUKE U. SCH. OF L., https://law.duke.edu/sites/default/files/images/centers/judicialstudies/Panel_1-Background_Paper_4.pdf (last visited Aug. 29, 2017).

19. Herbert L. Roitblat, *Introduction to Predictive Coding* 7 (OrcaTec 2013), <http://www.theolp.org/Resources/Documents/Introduction%20to%20Predictive%20Coding%20-%20Herb%20Roitblat.pdf> (last visited Aug. 29, 2017).

20. *Id.*

21. Hampton, *supra* note 4, at 2.

It is more akin to teaching. Rather than training a student or junior associate to perform legal research, however, the attorney is training a software program. A well-designed seed-set and iterative process will deliver results far beyond keyword search results, and the value is limited only by the underlying algorithm. Attorneys can expect results sorted by concept, context, issue, metadata, probability, relevance, and clustering.²²

Essentially, the software will be able to do what seasoned attorneys do, infer meaning from what might not be explicitly stated.²³ While the technological specifics of predictive coding are beyond the scope of this article, attorneys should understand the general underlying technological principles; and monitor technological advances as such advancements will likely manifest as improved predictive coding product offerings in the legal market.

The relevant technological foundation for predictive coding includes algorithms related to the following:

Technological Concept	Description
Supervised Learning	Attorneys train the software through an iterative process.
Unsupervised learning	The software generates its own random review of the data and identifies patterns.
Linear or Non-Linear Review	The software delivers documents in chronological or other linear order, or can be grouped to improve sorting efficiency.
Prioritized Review	Documents of greater relevance are advanced to the top of the queue.
Automated Review	The algorithmic determination is considered final without opportunity for human intervention of a document into a certain classification.
Search Terms	Search engine and Boolean search techniques are incorporated into predictive coding software.

(Table Content Source: American Bar Association, 2012)²⁴

22. *Id.* at 3.

23. Roitblat, *supra* note 19, at 1.

24. ABA Section of Litigation, *Predictive Coding*, AM. B. ASS'N 3-4 (Apr. 20, 2012), http://www.americanbar.org/content/dam/aba/administrative/litigation/materials/sac_2012/14-1_predictive_coding_written_materials.authcheckdam.pdf.

While all predictive coding solutions use similar statistical concepts, they do not all use the same technological foundations.²⁵ Therefore, a good general understanding of the technological foundation of predictive coding would ensure attorneys choose the most advanced vendor.

II. HISTORY OF PREDICTIVE CODING IN LEGAL PRACTICE

A. *Early Scholarly Analysis*

Predictive coding in the legal field finds its genesis in 2008, when attorneys and computer scientists participating in the National Institute of Standards and Technology's Text Retrieval Conference (known as TREC) scientifically explored the inefficiency of Boolean searches.²⁶ The results showed that a Boolean search generated electronically discovered documents with only 22% relevancy.²⁷

In 2009, TREC participants sought to compare the effectiveness of manual discovery with Technology Assisted Review ("TAR").²⁸ Although not expressly referred to as predictive coding, the aforementioned "seed set" and iteration process was followed and it was shown that technology assisted review generated superior results both in terms of recall and precision.²⁹

The findings were replicated in another scientific analysis of the 2009 TREC data, with the researchers concluding, relative to Rule 26(g) of the Federal Rules of Civil Procedure's requirement that discovery be complete and accurate, that a TAR method was "all other things being equal, one that does as well as traditional practice would likely be considered reasonable."³⁰

The research conducted and published in 2009 and 2010 was noted by Federal Magistrate Judge Andrew Peck in an October 2011 article he wrote for Law Technology News that advocated the use of

25. Roitblat, *supra* note 19, at 3-4.

26. Maura R. Grossman & Terry Sweeny, *What Lawyers Need to Know About Search Tools*, NAT'L L. J. (Aug. 23, 2010), <https://www.law.com/nationallawjournal/almID/1202470870777/>.

27. *Id.*

28. Maura R. Grossman & Gordon V. Cormack, *Technology Assisted Review in E-Discovery can be More Effective and Efficient than Exhaustive Manual Review*, 17 RICH. J. L. TECH. 11, 2 (2011).

29. *Id.* at 3.

30. Herbert L. Roitblat et al., *Document Categorization in Legal Electronic Discovery: Computer Classification vs. Manual Review*, 61 J. AM. SOC'Y FOR INFO. SCI. AND TECH. 70, 77 (2010).

TAR.³¹ As a legal practitioner, Judge Peck was acutely aware of the limitations of keyword searching within electronic discovery. In 2009, he addressed the problems resulting from ineffective and uncooperative keyword searching by beginning his opinion with; “[t]his Opinion should serve as a wake-up call for the Bar”³²

B. Earliest Use of Predictive Coding in Courts

Judge Peck emerged as a leading jurist in issues related to Electronic Discovery,³³ and in 2012 presided over the first case involving the use of predictive coding.³⁴ In *Da Silva Moore*, the plaintiffs filed a gender discrimination suit against the defendant, a large advertising agency.³⁵ The parties ultimately agreed to a predictive coding protocol, which would use a seven-step iterative process whereby a dataset of approximately three million emails be reduced to a representative sample of 2,399 emails.³⁶ The discarded results would be randomly sampled to ensure potentially responsive documents were not erroneously excluded.³⁷

The historical significance of *Da Silva Moore*, however, was Judge Peck’s decree that “computer-assisted review is an acceptable way to search for relevant ESI in appropriate cases.”³⁸ Other judges quickly accepted this view. Shortly after Judge Peck’s decision, judges at both the state and federal level began accepting the legitimacy of predictive coding. In *Global Aerospace*, a case involving the collapse of three airline hangers in 2010, the Virginia Circuit Court allowed—over plaintiff’s objection—predictive coding due to the defendant’s production of 250 gigabytes of electronically stored data.³⁹

The court reasoned that it would take 20,000 man-hours and two million dollars to conduct a manual review that would return 60%

31. Andrew Peck, *Search, Forward*, LAW TECH. NEWS 2 (October, 2011), https://law.duke.edu/sites/default/files/centers/judicialstudies/TAR_conference/Panel_1-Background_Paper.pdf.

32. William A. Gross Constr. Assoc. v. Am. Mfr. Mut. Ins. Co., 256 F.R.D. 134, 134 (S.D.N.Y. 2009).

33. *Honorable Andrew J. Peck*, NAT’L EDISCOVERY LEADERSHIP INST., <http://ediscoveryleadership.org/index.php/speakers2016/90-hon-andrew-j-peck-u-s-d-c-southern-district-of-new-york-2> (last visited Jan. 20, 2017).

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

35. *Id.* at 183.

36. *Id.* at 187.

37. *See Glob. Aerospace Inc. v. Landow Aviation, L.P.*, 2012 WL 1419842 (Va. Cir. Ct. Apr. 9, 2012)

38. *Id.* at *8.

39. *Id.* at *2.

relevancy, while predictive coding would produce 75% relevancy in a fraction of the time and at a fraction of the cost.⁴⁰

Also in 2012, the U.S. District Court of Northern Illinois allowed the use of predictive coding in *Kleen Products LLC. v. Packaging Corporation of America*.⁴¹ This class action involved charges of price fixing in the containerboard industry, and was initiated several years prior to the *Da Silva Moore* decision.⁴² In 2011, plaintiffs requested the use of “content-based analytics” (i.e. predictive coding) in response to what they viewed as unresponsive results from the defendant’s keyword search activity.⁴³ The court held that predictive coding would not be applied to discovery requests prior to October 1, 2013 but could be used for data produced after October 1, 2013.⁴⁴

Kleen Products is an important case because the court allowed for the introduction of predictive coding after the initial meet-and-confer meetings had set the electronic discovery protocol, which did not include stipulations for predictive coding.⁴⁵ The parties collaborated and the court adapted to new technological realities of predictive coding.

In *Kleen Products*, the court also commended counsel for their collaboration and for working within the spirit of the Sedona Principles for electronic discovery,⁴⁶ and noted that courts should look to the Sedona Conference for guidance due to the lack of judicial frameworks to follow insofar as the use of predictive coding.⁴⁷

One aspect judges considered early in the development of predictive coding was proportionality.⁴⁸ Under Federal Rule of Civil Procedure 26(b)(1), all discovery must be proportional; therefore, judges must weigh the benefits of conducting predictive coding with the costs.⁴⁹ Of course, the question then becomes, what is proportional?

In October 2012, the Chancery Court of Maryland ordered that both parties in the case of *EORHB, Inc. v. HOA Holdings LLC*, use the

40. Glob. Aerospace Inc., 2012 WL 1419842 at *2.

41. Kleen Prods. L.L.C. v. Packaging Corp. of Am., No. 10 Civ. 5711, 2012 WL 4498465, at * 5 (N.D. Ill. Sept. 28, 2012).

42. *Id.* at *1.

43. *Id.* at *5.

44. *Id.*

45. *Id.* at *3-4.

46. *Id.* at *14.

47. *Id.*

48. *Sedona Conference Cooperation Proclamation: Resources for the Judiciary*, THE SEDONA CONF. 18 (Aug., 2011), <https://thesedonaconference.org/download-pub/425>.

49. *Id.*

same predictive coding vendor in their litigation for efficiency.⁵⁰ In May 2013, the same court issued an order that the plaintiff—after showing good cause—was not required to use the predictive coding vendor because the plaintiff did not have enough data that would justify the expense.⁵¹ Clearly, the court was considering the proportional costs and benefits.⁵²

To achieve proportionality, Federal Rule of Civil Procedure 26(c) allows courts to shift costs associated with electronic discovery.⁵³ Part (or all) of the cost to utilize predictive coding can be shifted to the requesting party. In *Independent Living Center of Southern California v. City of Los Angeles*, the court held that the plaintiff could require the defendant to utilize predictive coding; but also held the plaintiff would be required to share the \$50,000 expense.⁵⁴

In *Chevron v. Donzinger*, the issue dealt with cost sharing of discovery expenses for a nonparty.⁵⁵ The nonparty, Patton Boggs, requested cost sharing upon the contention that fulfilling the plaintiff's discovery request would take 15-20 lawyers 40-hours per week, for 40 weeks.⁵⁶ Patton Boggs did not calculate, for the court, the cost savings associated with predictive coding, and the court deemed this tacit recognition that predictive coding would have reduced their costs.⁵⁷ Patton Boggs' request for cost sharing was denied.⁵⁸

The EORHB created an important legal issue that the Chancery Court sidestepped when rescinding the court order on proportionality grounds: Can courts compel parties to use predictive coding? In 2015, the legal community again looked to Judge Andrew Peck for guidance.

In *Rio Tinto v. Vale*, Judge Peck noted that, as of 2015, predictive coding had become black letter law, and where parties wanted to use technology assisted review courts would allow it.⁵⁹ Stated differ-

50. EORHB, Inc. v. HOA Holdings LLC, C.A. No. 7409-VCL, 2013, at *1 (Del. Ch. May 6, 2013).

51. *Id.*

52. *See Id.*

53. Emily Madavo et al., *Recent Key Developments in Shifting E-discovery Costs*, 4 EDDE J. 2 (2013).

54. Civil Minutes – General at 3, *Indep. Living Ctr. of S. Cali. v. L. A.*, No. CV 12-551-FMO (C.D. Cal. June 26, 2014).

55. *Chevron Corp. v. Donziger*, No. 11 CIV. 0691(LAK), 2013 WL 1087236, at *8 (S.D.N.Y. Mar. 15, 2013).

56. *Id.* at *32.

57. *Chevron Corp.*, 2013 WL 1087236 at *32.

58. *Id.*

59. *Rio Tinto PLC v. Vale, S.A.*, 306 F.R.D. 125, 127 (S.D.N.Y. 2015).

ently, courts would allow predictive coding but could not require it. Therefore, had the initial *EORHB* order compelling the use of predictive coding been challenged, it likely would have been overturned. The traditional standard in American jurisprudence is that discovery be afforded “broad and liberal treatment”⁶⁰ and a court’s delegation of how to perform discovery—without a compelling reason—does not seem congruent with the precepts of Rule 26.

While predictive coding was enthusiastically embraced by many in its early years of adoption, it was not embraced by all legal practitioners, including practicing attorneys and computer forensic experts.⁶¹ Early critics contended that each software algorithm should be subject to independent peer-review; and that judges should use Rule 72 to set aside orders that might have been based on unproven technology.⁶² These arguments did not resonate with earlier courts.

III. PRESENT STATE OF PREDICTIVE CODING IN LEGAL PRACTICE

A. *Adoption of Predictive Coding by Courts*

Everett Rodgers, Ph.D., performed a meta-analysis of various studies related to human adoption of innovation. Human beings tend to fall within one of five categories; innovators, early adopters, early majority, late majority, and laggards.⁶³ The *Diffusions of Innovations* framework has been applied to technological adaptation across society, and will assist our understanding of why the legal community has seen a low adaptation rate of predictive coding since 2012.

60. *Hickman v. Taylor*, 329 U.S. 495, 507 (1947).

61. Sharon D. Nelson & John W. Simek, *Predictive Coding: A Rose by Any Other Name*, AM. B. ASS’N, http://www.americanbar.org/publications/law_practice_magazine/2012/july-august/hot-buttons.html (last visited Jan. 24, 2017).

62. *Id.*

63. EVERETT ROGERS, *DIFFUSION OF INNOVATIONS*, 11, 22 (The Free Press, 3rd ed., 1962).

Category	Percentage of Population	Brief Description
Innovators	2.5%	Those who create and apply new innovations. This group has a higher tolerance for risk and ambiguity.
Early Adopters	13.5%	Those who see value in the innovations and provide support to the innovators (e.g. investors, early customers etc.). Also have an above average tolerance for risk.
Early Majority	34%	This group has less tolerance for risk and will only act upon the innovation when seeing <i>Early Adopters</i> benefit.
Late Majority	34%	This group has a lower tolerance for risk, and will only act after seeing the <i>Early Majority</i> benefiting.
Laggards	16%	This group will never embrace the innovation.

(Table based on Everett Rodger's Diffusion of Innovation)

For the purposes of this article, we will use the 2015 *Rio Tinto* decision to mark the beginning of the modern era. Predictive coding appears stuck in the *Early Adopter* stage. Judge Peck—an *Early Adopter*—contends predictive coding has reached the status of black letter law.⁶⁴ Yet, there is clearly a need for more knowledge on the subject among legal practitioners.⁶⁵ As Attorney Ralph C. Losey noted: “The bottom line is that at this point in time, late May 2016, the Bench is waiting for the bar to catch up.”⁶⁶

The judicial decisions since *Rio Tinto* have been similar in result; an endorsement of predictive coding. In *Malone v. Kantner*, a 2015 case where the court held a defendant was not required to share the costs of electronic data reproduction due to human error in the manual review process, the court noted; “[p]redictive coding is now promoted (and gaining acceptance) as not only a more efficient and cost-effective method of ESI review, but a more accurate one.”⁶⁷

64. *Rio Tinto PLC*, 306 F.R.D., at 127.

65. Ralph C. Losey, *Explanation of the Legal Profession's Remarkably Slow Adaptation to Predictive Coding*, E-DISCOVERY L. TODAY (May 24, 2016), <http://www.ediscoverylawtoday.com/2016/05/explanation-of-the-legal-professions-remarkable-slow-adoption-of-predictive-coding/>.

66. *Id.*

67. *Malone v. Kantner Ingredients, Inc.*, No. 4:12CV3190, 2015 WL 1470334, at *3 n. 7 (D. Neb. Mar. 31, 2015).

In *Burd v. Ford Motor Co.*, the parties engaged in a contentious electronic discovery dispute in litigation involving an alleged design flaw that prevented the throttles in certain automobiles from closing adequately.⁶⁸ Ford used arguments similar to those used by Patton Boggs in *Cheveron v. Donzinger*, namely that production would be overly broad; an undue burden; and create work-product privilege conflicts,⁶⁹ but the court was similarly unsympathetic. The parties were ordered to involve their IT departments to generate the required documents, and encouraged—not compelled—to utilize predictive coding to resolve the impasse.⁷⁰

B. Reluctance to Accept Predictive Coding Among Attorneys

The expectation of some in the legal community that the use of predictive coding would accelerate quickly post *Da Silva Moore* has not manifested.⁷¹ Is this due to the value of predictive coding being overhyped and under-delivering, attorneys failing to embrace new technology, or a combination of both sprinkled with other factors? The answer seems to depend on who is being asked.

Attorney Maura Grossman and Professor Gordon Cormack, Ph.D., who are active leaders within TREC and the authors of an influential 2010 article on Technology Assisted Review—subsequently cited by Judge Peck in his influential 2011 article—recently opened an article by noting: “The complex vocabulary and rituals that have come to be associated with TAR, including statistical control sets, stabilization, F1 measure, overturns, and elusion have dissuaded many practitioners from embracing TAR.”⁷²

The remainder of their brief article reinforces the premise that Continuous Active Learning, a TAR protocol, will assist attorneys in getting good results without requiring an exhaustive understanding of Technology Assisted Review.⁷³ However, reading between-the-lines, the intimation is that predictive coding is not as difficult to use as attorneys are making it. Under the *Diffusion of Innovations* theory, Innovators are making the case to the Early Majority. . .who are not yet buying the sales pitch. In *Confessions of an E-Discovery Lawyer*,

68. *Burd v. Ford Motor Co.*, No. 3:13-CV-20976, 2015 WL 4137915, at *1 (S.D.W. Va. July 8, 2015).

69. *Id.* at *4.

70. *Id.* at *11.

71. Losey, *supra* note 65.

72. Maura R. Grossman & Gordon V. Cormack, *Continuous Active Learning for TAR*, PRACT. L. THE J., April/May 2016, at 32, 33.

73. *Id.*

Attorney Geoffrey A. Vance places the failure of law firms to embrace and implement predictive coding squarely on the shoulders of attorneys and law firm partners.⁷⁴

Mr. Vance makes a logical argument that vendors are providing good services; the courts are encouraging adaptation of the technology; established law firms have adequate in-house Information Technology staff to assist with the implementation; and clients are willing to invest. The problem, Mr. Vance contends, rests with attorneys who simply do not know enough about predictive coding to make it work for their firms and clients.⁷⁵

However, not all attorneys take this view. In a 2015 report, Gibson Dunn contended that the slow adoption of predictive coding by counsel may have less to do with technology, and more to do with “the perception that disclosing irrelevant documents and coding decisions from training sets will be required of those who wish to use predictive coding.”⁷⁶

Magistrate Judge Peck noted in *Rio Tinto* that there are alternatives to this type of disclosure,⁷⁷ but it is possible this misconception is reducing the rate of adoption.

Gibson Dunn Attorney Gareth Evans, in a substantive 2015 on-line post, provided insight into the slow adoption rate, which leads to an inference that the problem may be as much cultural as technological.⁷⁸ That is, the need for “transparency and cooperation” has turned into a myth that counsel “must provide” irrelevant documents to opposing counsel, which is incompatible with the traditional adversarial nature of the discovery process.⁷⁹

C. *Predictive Coding: Ethical and Professional Considerations*

On April 29, 2015, the United States Supreme Court approved several amendments to the Federal Rules of Civil Procedure, effective

74. Geoffrey A. Vance, *Confessions of an E-Discovery Lawyer: We're Light Years Behind*, N.J.L. J., July 1, 2015, <https://advance.lexis.com/document?crd=c4817407-64c7-4ed0-bb3c-23ae57ab3d3e&pddocfullpath=%2Fshared%2Fdocument%2Flegalnews%2Furn%3AcontentItem%3A5GBJ-92T1-JBM3-R278-00000-00&pdcontentcomponentid=8308&pdmfid=1000516&pdisurlapi=true>.

75. *Id.*

76. Gibson Dunn, 2015 *Mid-Year E-Discovery Update*, GIBSON DUNN (July 15, 2015), <http://www.gibsondunn.com/publications/Pages/2015-Mid-Year-E-Discovery-Update.aspx>.

77. *Id.*

78. Gareth Evans, *Predictive Coding: Can it Get a Break?*, LINKEDIN, (July 23, 2015) <https://www.linkedin.com/pulse/predictive-coding-can-get-break-gareth-evans>.

79. *Id.*

December 01, 2015.⁸⁰ The scope of Rule 1 was amended, insofar as the rule “should be construed, administered, and employed by the court and parties to secure the just, speedy, and inexpensive determination of every action and proceeding.”⁸¹

The relevant change in wording, “employed by the court and parties,” has direct application to the use of predictive coding.⁸² The proposed changes were developed by the Civil Rules Advisory Committee in May 2010, after the TREC findings had been published but almost two years prior to the *Da Silva Moore* ruling.

Therefore, we can see that as early as 2010, courts recognized not only the efficiency and economic benefits of technology aided review, but that courts have been willing to use predictive coding to ensure “just, speedy, and inexpensive” discovery. In 2012, the issue was raised as to whether the proposed change to Rule 1 would impose new ethical responsibilities towards predictive coding collaboration.⁸³

A take-away from *Rio Tinto* is that courts *cannot compel* parties to use predictive coding;⁸⁴ however, the later amendment to Rule 1 indicates that courts *could compel* such use where appropriate. It appears both judges and attorneys are faced with determining whether avoiding predictive coding unjustifiably creates an ethical rules violation at both the bench and the bar.

Additionally, Rule 1.1—Competence—of the ABA’s Model Rules of Professional Conduct, comment eight states: “To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology”⁸⁵

Failure to keep abreast of advancements in technology could be viewed as violating Rule 1.1.⁸⁶ This leads to an important question; are attorneys who neglect to educate themselves on issues related to predictive coding committing an ethical violation?

80. See Supreme Court of the United States, *Proposed Amendments to The Federal Rules of Civil Procedure*, SUPREME COURT.GOV (Apr. 29, 2015), [https://www.supremecourt.gov/orders/courtorders/frcv15\(update\)_1823.pdf](https://www.supremecourt.gov/orders/courtorders/frcv15(update)_1823.pdf).

81. *Id.* at 4.

82. See Barkett, *supra* note 7, at 6-7.

83. See *id.* at 6.

84. *Rio Tinto PLC*, 306 F.R.D. at 127.

85. MODEL RULES OF PROF’L RESPONSIBILITY r. 1.1 cmt. 8 (AM. BAR. ASS’N 2015).

86. Seth H. Row, *Technology-Related Ethics Rules Changes Litigators Must Understand*, AM. B. ASS’N (Feb. 14, 2013), <http://apps.americanbar.org/litigation/committees/pretrial/email/winter2013/winter2013-0213-technology-related-ethics-rules-changes-litigators-must-understand.html>.

IV. FUTURE APPLICATION OF PREDICTIVE CODING

A. *Special Masters: Application of Rule 53*

The present state of predictive coding could be synthesized as: predictive coding is a proven method of making electronic discovery faster, better, and in-expensive, with few lawyers using it. If accurate, then what can we envision for the future? Again, the judiciary may want to look to Magistrate Judge Peck for guidance. Several months after declaring predictive coding is accepted as black letter law in *Rio Tinto*, Judge Peck appointed Attorney Maura Grossman as a special master in that case to resolve predictive coding related disputes among the parties.⁸⁷

Under Rule 53 of the Federal Rules of Civil Procedure, a special master may (except when prevented by statute) intervene to “address pretrial and posttrial matters that cannot be effectively and timely addressed by an available district judge or magistrate judge of the district.”⁸⁸ In *Rio Tinto*, Judge Peck appointed a special master after the predictive coding protocols were established by the parties and accepted by the court.⁸⁹ There should have been no need for a special master at this stage, but the parties could not agree to the execution of the protocol (which defeats the purpose of a protocol).

Writing in explanation of his order appointing Attorney Grossman as special master, Judge Peck stated, “*Rio Tinto* . . . objects on the ground that Vale should have agreed much earlier to appointment of a special master. The Court certainly agrees, but as the saying goes, better late than never.”⁹⁰ A reasonable inference for attorneys going forward: Judges might be more proactive in appointing special masters from the onset in cases involving predictive coding.

B. *Predictive Coding and International Application*

In February 2016, the England and Wales High Court (Chancery Division) ruled that predictive coding could, for the first time, be used in an English court.⁹¹ In this case, *Pyrrho Investments v. MWB*

87. Jason Krause, *The Battle of Rio Tinto: Predictive Coding Hits Snag in Marquee Case*, ACEDS (July 24, 2015), <https://www.aceds.org/news/news.asp?id=287452&terms=%22battle+and+rio+and+tinto+and+predictive+and+coding+and+hits+and+sn%22>.

88. FED. R. CIV. P. 53 (a)(1)(C).

89. Krause, *supra* note 87.

90. *Tinto v. Vale*, No. 14 CIV. 3042 RMB AJP, 2015 WL 4367250, at *1 (S.D.N.Y. July 15, 2015).

91. *Pyrrho Inv. Ltd. v. MWB Bus. Exch., Ltd.*, [2016] EWHC 256 (Ch) [1] (Eng.).

Property Limited (an action of shareholders against company directors for alleged conflict-of-interest), the defendants produced over three million emails.⁹² Master Matthews, in announcing the decision, noted the proportionality benefits of predictive coding: “[E]stimates given in this case vary between £181,988 plus monthly hosting costs of £15,717, to £469,049 plus monthly hosting costs of £20,820. This is obviously far less expensive than the full manual alternative.”⁹³

The High Court of Ireland, in *Irish Bank Resolution Corporation v. Quinn*, was the first European court to allow predictive coding.⁹⁴ In *Quinn*, the parties agreed that there were 680,809 relevant documents, which would take ten lawyers nine months to review, at a cost of 2,000,000 Euros; while predictive coding would take ten lawyers 113 days, at a cost of 220,000 Euros.⁹⁵ The Court looked at the protocol outlined in *Da Silva Moore*, and determined it to be reasonable and proportionate.⁹⁶

The Irish court reaffirmed the defensibility of the technology; “[t]he evidence establishes, that in discovery of large data sets, technology assisted review using predictive coding is at least as accurate as, and, probably more accurate than, the manual or linear method in identifying relevant documents”⁹⁷; and arrived at the decision that predictive coding is appropriate under Irish law.⁹⁸

There is scant literature published, as of July 2016, on predicting the evolution of predictive coding outside the United States. At least insofar as England and Ireland, attorneys will likely see a similar trajectory as seen in the United States post-*Da Silva Moore*; increased acceptance by judges seeking efficiency and economy, and hesitation among attorneys uncomfortable with the technological and collaborative nature of the predictive coding process.

Where the use of technology-assisted review is not explicitly discussed in a foreign jurisdiction’s rules of civil procedure, lawyers can look to the *Quinn* decision for guidance. If the rules do not require that a manual review be carried out in the discovery process,⁹⁹ courts will likely accept it.

92. *Id.* ¶ 5.

93. *Id.* ¶ 33(7).

94. *Irish Bank Resolution Corp. v. Quinn* [2015] IEHC 175 (H. Ct.) (Ir).

95. *Id.* ¶ 12.

96. *Id.* ¶ 18.

97. *Id.* ¶ 66.

98. *Id.* ¶ 69.

99. *See Grossman & Cormack, supra* note 72.

C. Confidentiality and Work-Product Conflicts

Disclosing attorney-client privileged information and, or, privileged work product is a legitimate concern for attorneys considering predictive coding.¹⁰⁰ However, this issue harkens the reader back to the need to understand the technology underlying predictive coding. That is, predictive coding software will identify privileged documents as easily, and economically, as it finds responsive documents.

In November 2015, the American Bar Association identified predictive coding as a tool for mitigating the likelihood of inadvertently disclosing confidential information, or waiving privilege, rather than viewing it as a disclosure risk.¹⁰¹ This is important because Rule 1.6(c) of professional conduct states; “[a] lawyer shall make reasonable efforts to prevent the inadvertent or unauthorized disclosure of, or unauthorized access to, information relating to the representation of a client.”¹⁰² Therefore, the ABA considers the use of predictive coding a “reasonable” effort in mitigating the risk of inadvertent disclosure. A review of scholarly and practitioner literature leads to an inescapable conclusion; predictive coding will likely result in the inadvertent disclosure of privileged information.

The key is to take reasonable steps. Attorneys who use predictive coding in the future should familiarize themselves with two important rules; Federal Rule of Civil Procedure 26(b)(5)(B), and Federal Rule of Evidence 502(b).¹⁰³ Rule 26(b)(5)(B) provides the conditions and procedures relative to claw-back provisions, and Rule 502(b) relates to inadvertent disclosures where the attorney “took reasonable steps” to avoid and rectify the inadvertent disclosure of privileged materials.¹⁰⁴ However, attorneys who take the same level of reasonable care with predictive coding, as they do with manual review, should be reasonably well protected and in compliance with professional ethics.

100. Manfred Gabriel et al., *The Challenge and Promise of Predictive Coding for Privilege*, ICAIL DESI V WORKSHOP 1 (June 14, 2013), <https://pdfs.semanticscholar.org/aa13/dc9888c3913c308c1cd4319ccc1571e6e3f2.pdf>.

101. Lisa M. Gonzalo, *Inadvertent Disclosure in E-Discovery: How to Avoid Waiver of Privilege*, AM. B. ASS'N (Nov. 3, 2015), <http://apps.americanbar.org/litigation/committees/commercial/articles/fall2015-1115-inadvertent-disclosure-ediscovery-avoid-waiver-privilege.html>.

102. MODEL RULES OF PROF'L CONDUCT r. 1.6(c) (AM. BAR ASS'N 1983).

103. Gonzalo, *supra* note 101.

104. *Id.*

D. Teaching the Next Generation of Lawyers

Villanova University School of Law Professor Michele Pistone noted in a recent article: “. . . [L]awyers and law students will need to develop new skill sets in order to thrive professionally. Unfortunately, corporate clients are finding that many lawyers, including recent graduates, lack skills in the more sophisticated uses of technology for lawyering and law practice.”¹⁰⁵ Professor Pistone then delves into an explanation of predictive coding as an example of a sophisticated use of technology.¹⁰⁶ According to Professor Pistone, the legal academy has not kept pace with technology and seems stubbornly rooted to a “sage on the stage” culture, where professors and administrators do not appreciate that many of today’s law students have been engaging technology since pre-school.¹⁰⁷

Assuming this view is correct, it is fair to ask what role should law school professors and administrators play insofar as integrating predictive coding instruction into course curricula. Assuming predictive coding would be taught as a dimension of e-discovery, rather than a stand-alone course, let’s consider the most recent data. There are 204 ABA accredited law schools.¹⁰⁸ As of August 2015, a survey of 193 of those schools indicates: 123 law schools offer no e-discovery courses; sixty-nine (69) law schools offer at least one e-discovery course; and only eight (8) law schools incorporate hands-on learning using technological tools, which would include predictive learning.¹⁰⁹ Certainly, it appears law schools must be more proactive in teaching e-discovery in general, and predictive coding in particular.

CONCLUSION

This article began by defining predictive coding simply as a process whereby attorneys train a computer program to identify potentially relevant documents within a large body of documents.¹¹⁰

105. Michele Pistone, *Law Schools and Technology: Where We Are and Where We are Heading*, 64 J. LEGAL EDUC., 586, 589 (2015).

106. *See id.* at 590.

107. *Id.* at 592.

108. Am. Bar Ass’n, *ABA-Approved Law Schools*, AM. B. ASS’N, http://www.americanbar.org/groups/legal_education/resources/aba_approved_law_schools.html (last visited Sept. 19, 2017).

109. William Hamilton & Michelle C.S. Lang, *Law Schools Lag in Teaching E-Discovery*, TODAY’S GEN. COUNS., June/July 2015, at 28-29, <http://digital.todaysgeneralcounsel.com/?issueID=31&pageID=31>.

110. Hampton, *supra* note 4 at 29.

But predictive coding is not, fundamentally, a legal concept. Rather, it is a merging of statistics and computer science, which is applied to appropriate legal contexts. The article concludes by noting an inability or unwillingness to educate today's legal community as to the integration of statistics, technology, and modern discovery techniques, that will likely result in the promise and potential of predictive coding being unfulfilled.

The key word in the preceding paragraph may be the word "appropriate." When breaking legal ground in 2012, Judge Andrew Peck noted, "The technology exists and should be used where appropriate, but it is not a case of machine replacing humans: it is the process used and the interaction of man and machine that the courts need to examine."¹¹¹ Nobody—including the most ardent supporters of e-discovery—would suggest predictive coding is appropriate in all cases. Indeed, predictive coding is not warranted in the vast majority of cases.

However, in cases where litigants are exchanging hundreds of thousands, and potentially millions of documents, predictive coding is likely an appropriate, proportional, and defensible tool. As noted throughout this article; the technology is proportional, in larger cases, because it allows for vast amounts of documentation to be filtered for a fraction of the cost of manual discovery. The technology is defensible because scientific analysis indicates predictive coding generates a greater percentage of relevant documents than manual review. Thus, the proper use of predictive coding fulfills the Rule 26 requirement to perform reasonable discovery.

Concerns that predictive coding will produce documents that are privileged are well founded, as it likely will. However, there are proactive steps (e.g. claw-back agreements) that will mitigate this risk. Additionally, the initial iterative process whereby attorneys train the software to generate relevant documents can be used to identify, and segregate, privileged documents. Therefore, the American Bar Association has identified predictive coding as a proactive way to avoid the disclosure of privileged documents. For practical purposes, the rewards associated with the use of predictive coding clearly outweigh the risk.

The essence of this article was to inform those new to predictive coding that the model is new insofar as the practice of law, but the underlying statistics and computer science are well established and theoretically sound. Although predictive coding has a brief history in the law, it has quickly evolved to the point where judges will accept its

111. *Da Silva Moore v. Publicis Groupe & MSL*, 287 F.R.D. 182, 189 (S.D.N.Y. 2012).

validity as black letter law. Where appropriate, American judges will allow predictive coding, and recent cases in England and Ireland suggest foreign courts will follow suit. The impetus is upon law schools to teach predictive coding techniques to the next generation of lawyers and for licensed attorneys to develop a better understanding of predictive coding in daily practice.