ORRICK

The Trend Towards Greater Adoption of Predictive Coding: the Good, the Bad, and the Ugly

Wendy Butler Curtis and Jeffrey W. McKenna[i]

Introduction

Predictive coding, also known as automated review, represents an evolving technology that provides litigants with a computer-assisted alternative to manual review of large document sets. Although numerous implementations of the technology exist, nearly all involve a team of attorneys manually reviewing a small sampling of documents to "train" the computer how to code the larger pool of remaining documents.

Recently, a number of opinions have been issued permitting, or even encouraging, the use of predictive coding. As a result, some pundits and jurists have gone so far as to suggest that search terms are dead, or at best, an antiquated method for culling documents. Yet, while the use of predictive coding can reduce the cost of review, it is not a magic bullet or a good fit for every case. Indeed, in many cases, search terms paired with a well-managed human review will still provide the most cost-effective option without sacrificing accuracy.

This alert highlights developing trends in predictive coding, explores the range of associated consequences and suggests appropriate scenarios for using predictive coding in lieu of other forms of computer-assisted review.

Contact a Team Member

Wendy Butler Curtis, eDiscovery Of Counsel
Washington, D.C.
(202) 339-8584
wcurtis@orrick.com

Cathleen Peterson, eDiscovery Of Counsel and Legal Director, Document Review Service
Global Operations Center
(304) 231-2850
cpeterson@orrick.com

Jeffrey W. McKenna, eDiscovery and Privacy Attorney
San Francisco, CA
(415) 773-4152

PREDICTIVE CODING IN 30 SECONDS

Predictive coding provides litigants with a computer-assisted alternative to manual review of large document sets. Nearly all implementations involve a team of attorneys manually reviewing a small sampling of documents to "train" the computer how to code the bulk of the documents at issue.

Within the past six months, case law has evolved, including:

- One of the first court orders endorsing the use of predictive coding.
- The first state court order holding that defendants can use predictive coding despite a plaintiff's objections.
- A ruling by Judge Scheindlin supporting the use of predictive coding in both civil discovery and in the FOIA context.

The Good:

- Courts are embracing technology and cost-effective approaches.
- Predictive coding technology has tremendous potential to reduce costs and improve accuracy.
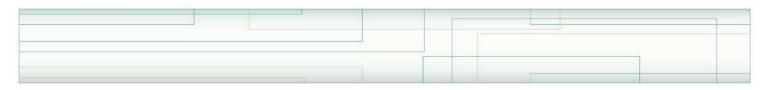
The Bad:

- Too much emphasis on the benefits of predictive coding has led to inaccurate criticism of other technologies, especially search terms.
- When using predictive coding it may be necessary to provide all non-privileged training documents and their coding, regardless of relevance or responsiveness, to opposing counsel.

The Ugly:

- Litigating the need for predictive coding and negotiating predictive coding protocols can be more expensive than other options.

Bottom line: Search terms are not obsolete, and parties should be cautious of costs and delays associated with negotiating predictive coding protocols.

Trends

Over the past six months, the case law addressing predictive coding developed considerably. In February, Magistrate Judge Andrew Peck issued one of the first court orders endorsing the use of predictive coding in *Da Silva Moore v. Publicis Groupe*.[ii] Two months later, in *Global Aerospace v. Landow Aviation*,[iii] Judge Chamblin issued the first state court order holding that defendants can use predictive coding despite a plaintiff's objections. On July 13, 2012, Judge Shira Scheindlin issued an opinion in *National Day Laborer v. U.S. Immigration & Customs Enforcement*,[iv] endorsing the use of predictive coding both for civil eDiscovery[v] and responses to Freedom of Information Act ("FOIA") requests. Most recently, on August 21, 2012, Magistrate Judge Nan Nolan approved a stipulation in *Kleen Products v. Packing Corp. of America* resolving a contentious dispute regarding the purported advantages of predictive coding over traditional keyword searching and related discovery obligations.[vi]

The Good

- **Courts Are Embracing New Technology and Cost-Effective Approaches.**

In *Da Silva Moore*, Judge Peck emphasized the importance of choosing an eDiscovery method that, consistent with Rule 1, would help "secure the just, speedy, *and inexpensive*" resolution of the case at bar.[vii] He also emphasized the importance of selecting a discovery method consistent with Rule 26(b)(2)(C)'s proportionality requirements.[viii] Indeed, Judge Peck's support for predictive coding is based in no small part on its perceived potential to reduce discovery costs.[ix] Similarly, in suggesting predictive coding as an option in *National Day Laborer*, Judge Scheindlin specifically mentions its potential to permit more efficient and cost-effective review, recognizing that the appropriate use of various technologies such as search terms and predictive coding—either independently of one another or in combination—can achieve more accurate and cost-effective results.[x] Such a willingness on the part of courts to embrace new technologies to improve the quality of document productions while reducing their cost is commendable and offers great promise.

According to a recent study by the Rand Corporation, predictive coding could result in a time savings of 75 percent compared to traditional review methods.[xi]

Similarly, the Electronic Discovery Institute recently conducted a survey of eleven eDiscovery vendors that found, "on average, predictive coding saved forty-five percent of the costs of normal review."[xii] Orrick is currently participating in the Electronic Discovery Institute's Computer Assisted Document Review Study.

Moreover, as the defendants argued in *Global Aerospace*,[xiii] predictive coding can also improve accuracy and consistency.[xiv]

The Bad

- **Too Much Emphasis on the Benefits of Predictive Coding Has Led to Inaccurate Criticism of Other Technologies, Especially Search Terms.**
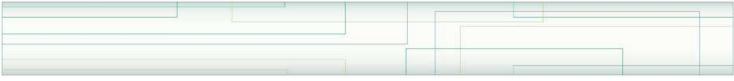
Although Judge Peck warned, "computer-assisted review is not a magic, Staples-Easy-Button, solution appropriate for all cases[,]"[xv] there has been a tendency to see predictive coding as a cure-all or magic bullet for all eDiscovery ailments. That is not the case. Parties and courts need to separate the hype from the reality.[xvi]

One of the downsides to the hype surrounding predictive coding is that it has led some pundits to assume that all other preexisting technologies are inferior or "antiquated," when in reality nothing could be further from the truth. For example, despite recent criticisms, keywords remain an effective and accepted method for culling documents.[xvii]

Moreover, as suggested by Judge Scheindlin, sometimes the best results will actually be achieved by using a combination of technologies.[xviii] For example, when using predictive coding, it is often most cost-effective to begin the process with keyword searching to isolate relevant documents.[xix] Accordingly, the pundits and vendors who are trumpeting the end of keyword searches are doing everyone a disservice.

- **Undesirable Disclosures May Be Required.**

Many predictive coding protocols used to date have required all non-privileged training documents and their coding, regardless of relevance or responsiveness, to be provided to opposing counsel. In fact, Judge Peck encouraged this practice, stating that this level of transparency enables opposing counsel and the Court to be more comfortable with predictive coding. While transparency certainly has its benefits, not every case or every client may benefit from the same degree of it. If you or your client are concerned about providing too much information to opposing counsel, predictive coding may not be the way to go unless you can agree on a more restrictive protocol that does not require such disclosures.

The Ugly

- **Litigating Predictive Coding Can Be More Expensive Than Other Options.**

The emerging trend in which courts are encouraging extensive meet and confer negotiations and disclosures regarding the choice of predictive coding tools and protocol to be followed is of great concern. Repeated, open-ended meet and confer sessions and hearings involving attorneys and outside experts can cause costs to add up quickly, but still fail to produce a workable agreement. For example, in *Da Silva Moore*, even though all parties purportedly liked the idea of using predictive coding, after multiple meet and confer sessions and hearings they were unable to reach an agreement on the details, and instead ended up in contentious collateral litigation over the supposedly "agreed" predictive coding protocol ordered by Judge Peck. Similarly, in *Kleen Products*, the defendants had already produced around three million pages of documents and incurred approximately 1,300 hours of outside expert time developing and validating their search methodology in consultation with plaintiffs, when plaintiffs requested that the court order defendants to redo the entire production using predictive coding. Eight months later, plaintiffs largely dropped their demands, but by then the Court had heard approximately two full days of expert testimony regarding the advantages/disadvantages of each competing approach. There had also been numerous hearings and case management conferences on the subject, as well as extensive briefing. Thus, as these two cases illustrate, without careful judicial oversight, the potential cost savings produced by using predictive coding can easily be overshadowed by increased attorney and expert fees for related negotiations and litigation. Indeed, if the desire on the part of one party to use predictive coding results in the sort of extensive mini-trials or appeals seen in *Kleen Products* and *De Silva Moore*, any potential benefits in cost-efficiency will be lost.

- **The Availability of Predictive Coding Does Not Justify Requiring Re-Review.**

As demonstrated by the *Kleen Products* case, some opposing parties are not shy about asserting their preference for discovery methods and technologies on the other side, even after significant time and expense has been incurred utilizing another means of discovery.[xx] In *Kleen Products*, for example, the plaintiff requested that the defendant be *forced* to recollect and review data using predictive coding rather than search terms.[xxi] As noted above, these demands came after discovery was largely completed and without a challenge to the completeness of the productions.[xxii] The technology and methods used for discovery should be within the control of the producing party, provided the producing party uses methods that meet the basic requirements of Rule 26 and Rule 34. Arguments over which method might in theory have been more ideal are not appropriate in an adversary system, lest every case turn into a mini-trial over eDiscovery technology. Indeed, many large companies and law firms have already made considerable investments in existing technologies and infrastructure – often after careful consideration of each institution's unique discovery needs and preferences. Thus, to expect companies and law firms to change their technology or preferred review strategies based on the opposing parties' technological preferences is simply not appropriate.

Conclusion

Predictive coding clearly offers great promise. None of these issues are fatal, but they should be considered when deciding what the most appropriate review method will be in your case. Moreover, to the extent the excitement and hype surrounding predictive coding leads to opinions and commentary denigrating search terms, it will have done the eDiscovery community a great disservice.

---

[i] We would like to thank Orrick summer associate Whitney H. Coble for her significant contributions to this article.

[ii] *Da Silva Moore v. Publicis Groupe*, No. 1:11-cv-01279 (S.D.N.Y. Feb. 24, 2012) [Doc. 96].

[iii] *Global Aerospace, Inc. v. Landow Aviation, L.P.*, No. CL 61040 (Va. Cir. Ct. April 23, 2012).

[iv] *Nat'l Day Laborer Organizing Network v. U.S. Immigration & Customs Enforcement Agency*, No. 1:10-Civ-3488, 2012 U.S. Dist. LEXIS 97863, (S.D.N.Y. July 13, 2012) [Doc 197].

[v] Although the procedural context of the opinion is somewhat unique to FOIA, the opinion is suggests that Judge Scheindlin's favorable discussion of predictive coding is not limited to the FOIA context, but rather would apply in the civil discovery as well. *See Id.* at *37-*40, *45, n. 110.

[vi] *Kleen Products, LLC v. Packaging Corp. of Am.*, Case No. 1:10-cv-05711 (N.D. Ill. Aug. 21, 2012) [Doc. 385].

[vii] Fed. R. Civ. P. 1, cited in *Da Silva Moore v. Publicis Groupe, et al.*, No. 1:11-cv-1279, 22 (S.D.N.Y. Feb. 24, 2012) (emphasis added).

[viii] Fed. R. Civ. P. 26(b)(2)(C), cited in *Da Silva Moore* at 22.

ix *Da Silva Moore v. Publicis Groupe*, No. 1:11-cv-01279 (S.D.N.Y. Feb. 24, 2012) [Doc. 96] at 11, 25 ("[C]omputer-assisted review works better than most of the alternatives, if not all of the [present] alternatives. . . . [It] should be seriously considered for use in large-data-volume cases where it may save the producing party (or both parties) significant amounts of legal fees in document review.").

x *Nat'l Day Laborer*, 2012 U.S. Dist. LEXIS 97863, at *37-45.

xi Nicholas M. Pace & Laura Zakaras, "Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery," RAND Institute for Civil Justice, xviii (2012).

xii Seven of the eleven vendors in the survey also reported savings of 70 percent or more in individual cases. *See* Anne Kershaw & Joseph Howie, "Crash or Soar? Will the legal community accept 'predictive coding?'" Law Technology News (Online), Oct. 1, 2010, available at **http://www.akershaw.com/articles/LTN_CrashOrSoar_2010_Oct.pdf**; *see also*, Axcelerate eDiscovery Suite with Predictive Coding, http://www.recommind.com/products/axcelerate-ediscovery; Kroll Ontrack, Inc. case study, **http://www.krollontrack.com/library/irtcostsavingcasestudy_kroll-ontrack2012.pdf** (2012).

xiii *See* Case Briefs section, *infra*, for additional details.

xiv Although few independent studies have been done on the accuracy of predictive coding, those that have suggested the technology can be at least as effective as human linear review. For example, in a well-known article titled "Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient than Exhaustive Manual Review," Maura Grossman and Gordan Cormack compared the results of manual reviews with the results of five technology-assisted reviews. The results demonstrated that the technology-assisted reviews were more efficient, on average, than the five manual reviews. *See* Pace, *supra* note 24, at xviii; 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. 11, 35, Spring 2011, **http://jolt.richmond.edu/v17i3/article11.pdf**; *see also*, 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, 79 (2010) (reaching similar conclusion).

xv *Da Silva Moore v. Publicis Groupe*, No. 1:11-cv-01279 (S.D.N.Y. Feb. 24, 2012) [Doc. 96] at 17.

xvi Predictive coding is not a replacement for good judgment and careful discovery planning. Indeed, the cardinal rule remains that the single most effective cost reduction method is the focused collection of records most likely to contain relevant information. Parties and courts still need to think about the collection and review strategy that makes the most sense in each case, and as discussed below, that will not always be predictive coding. *See* Patrick Oot, Anne Kershaw, and Herbert L. Roitblat, *Practitioners' View: Mandating Reasonableness in a Reasonable Inquiry*, 87 DENV. U.L. REV. 533 (2010).

xvii In December 2011, the District of Delaware adopted new ESI policies favoring use of search terms. *See Default Standard for Discovery, Including Discovery of Electronically Stored Information*. Similarly, the Federal Circuit and United States District Court for the Eastern District of Texas each recently adopted model eDiscovery orders based entirely on the concept of keyword searching. *See Federal Circuit Model Order Regarding E-Discovery in Patent Cases; Eastern District of Texas Model Order Regarding E-Discovery In Patent Cases*. Several studies also suggest that a well-designed keyword-based culling strategy can reduce costs over manual, linear review and achieve a high rate of accuracy. *See* Herbert L. Roitblatt, Anne Kershaw & Patrick Oot, *Document Categorization in Legal Electronic Discovery: Computer Classification vs. Manual Review*, 61 J.AM. SOC'Y FOR INFO. SCI. & TECH. 70, 79 (2010); Maura R. Grossman & Gordon V. Cormack, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, RICH. J.L. & TECH., Spring 2011, at 43, 48; David L. Blair & M.E. Maron, *An Evaluation of Retrieval Effectiveness for A Full-Text Document Retrieval System*, 28 COMM. ACM 289 (1985).

xviii *Nat'l Day Laborer*, 2012 U.S. Dist. LEXIS 97863, at *37-45.

xix Ed. Jason R. Baron. *The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery*. The Sedona Conference, Vol. 8, 200 (Fall 2007) (defined as "set-based searching using simple words or word combinations, with or without Boolean and related operators").

xx *Kleen Products*, 775 F. Supp. 2d 1071.

xxi 2/6/12 Pls.' Mem. [Doc. 290] at 8.

xxii 2/6/12 Defs.' Br. [Doc. 288] at 8.

CASE BRIEFS

Da Silva Moore

On February 24, 2012, Andrew J. Peck of the United States District Court for the Southern District of New York entered an order authorizing the parties in *Da Silva Moore v. Publicis Groupe*[xxiii] to rely on predictive coding for identification of responsive documents during discovery in lieu of traditional document review and search terms. Until recently, few courts had addressed the issue, making Judge Peck's order one of the first. The decision speaks very positively regarding the benefits of predictive coding and cites statistics indicating that human review is no more accurate, and perhaps less accurate, than predictive coding. The order is also highly critical of using keywords to filter documents for review, despite the wide acceptance of this practice.[xxiv]

- The Protocol

In order to "train" the software used for predictive coding, the parties in *Da Silva Moore* were required to generate multiple training sets using random sampling from a corpus of 3 million documents.[xxv] These training sets amounted to thousands of documents and were to be reviewed by senior attorneys on the case team, not junior associates, contract attorneys or paralegals. Indeed, the training process ordered involved several distinct stages, including seven iterative rounds of review involving 500 documents per round, with additional rounds added as needed.[xxvi]

It was also ordered that non-privileged training documents and their coding, regardless of relevance or responsiveness, had to be provided to opposing counsel to permit validation of the defendants' training method.[xxvii]

Finally, defendants had requested that Judge Peck limit their total production to 40,000 of the most relevant documents as identified by the predictive coding software.[xxviii] However, he refused to place any limit on the total number of documents to be produced. Instead, the Court explained that "where [the] line will be drawn [as to review and production] is going to depend on what the statistics show for the results," since "[p]roportionality requires consideration of results as well as costs. And if stopping at 40,000 is going to leave a tremendous number of likely highly responsive documents unproduced, [defendant's proposed cutoff] doesn't work."[xxix]

- Procedural Context

Despite the groundbreaking nature of Judge Peck's order, the fact that the parties purportedly agreed to use predictive coding on their own and Judge Peck merely helped them resolve some lingering disputes as to its precise implementation made the decision of limited precedential value. Indeed, a subsequent and unpleasant controversy regarding the circumstances of the order and the purported "agreement" on which it was based further muddied the waters. Thus, while the order was ultimately affirmed by U.S. District Judge Andrew Carter in New York, standing alone it remained a limited precedent.[xxx]

Global Aerospace

On April 23, 2012, Virginia Circuit Court Judge James H. Chamblin issued the first state court order embracing predictive coding for eDiscovery.[xxxi] The order held that the defendants could use predictive coding despite plaintiff's objections that human-based review was more accurate and effective.[xxxii] The order also specified that processing and production were to be completed within 120 days, with "processing" to be completed within 60 days and "production to follow as soon as practicable and in no more than 60 days."[xxxiii] The order does not address whether the parties are required to decide upon a mutually agreeable protocol, nor does it address whether plaintiffs will be allowed to review any of the non-privileged training documents as permitted in *Da Silvia Moore*. However, Judge Chamblin does note that the receiving party will still have the opportunity to question "[the] completeness of the contents of the production or the ongoing use of predictive coding."[xxxiv] The order does not elaborate on what this meant.

In support of their motion to use predictive coding, defendants argued that a first pass manual review of approximately two million documents would cost two million dollars and successfully locate only 60 percent of the potentially responsive documents. Interestingly, defendants acknowledged in their briefing that keyword searching might be *more* cost-effective under the circumstances than predictive coding, but they claimed that it would only retrieve 20 percent of potentially relevant documents. In contrast, they contended that predictive coding would locate 75 percent or more of potentially relevant documents while still costing less than a traditional linear human review. Plaintiffs took issue with the notion that defendants might produce only 75 percent of potentially relevant documents, arguing instead that they should have to produce "all responsive documents located upon a reasonable inquiry." Plaintiffs did not address defendants' assertion that even a fully manual review was unlikely to locate more than 60 percent of such documents.

Thus, this opinion is not only the first time a state court found predictive coding to be an acceptable option, but also appears to be the first order resulting from a bona fide dispute between parties.

National Day Laborer Organizing Network

On Friday, July 13, 2012, Judge Shira Scheindlin of the United States District Court for the Southern District of New York issued an opinion exploring the reasonableness of self-collection by government entities responding to FOIA requests.[xxxv] The plaintiffs—the National Day Laborer Organizing Network, the Center for Constitutional Rights, and the Immigration Justice Clinic of the Benjamin N. Cardozo School of Law—were trying to obtain information from federal agencies, including the Federal Bureau of Investigation and the Department of Homeland Security.

Although the procedural context of the opinion is somewhat unique to FOIA, the opinion endorses the use of predictive coding, notes its potential benefits, and suggests that Judge Scheindlin's support of the technology is not limited to the FOIA context, but rather would apply in civil discovery as well.[xxxvi] In fact, Judge Scheindlin suggests that the parties consider utilizing predictive coding for future productions in the case, if they can agree on a protocol.[xxxvii]

Significantly, the opinion also reaffirms the validity of using search terms. Although Judge Scheindlin cautions that, like any technology, search terms must be used carefully, in accordance with best practices,[xxxviii] the opinion is clear that the occasional failure to uncover all documents is not fatal to the adequacy of the search. Rather, it is the "failure to design a search that is reasonably calculated to uncover all documents" that poses problems.[xxxix] To address this issue, the opinion recommends adhering to the Sedona Principles when using keyword searching, including "careful thought, quality control, testing, and cooperation with opposing counsel."[xl]

Kleen Products

Recently, Magistrate Judge Nan Nolan in the United States District Court for the Northern District of Illinois entered a stipulated order in *Kleen Products v. Packaging Corp. of America* resolving a dispute over the relative merits of predictive coding and keyword searching.[xli] As set forth more fully below, after discovery was largely finished and without a challenge to the completeness of defendants' productions, the plaintiffs in *Kleen* had demanded that the defendants be *forced* to recollect and re-review data using predictive coding, rather than search terms as had been used previously. Plaintiffs contended that predictive coding technology would provide added assurances that defendants' productions were thorough.[xlii]

To put this dispute and the recent stipulation in better context, it is worth considering the number of documents produced and the protocol for search terms followed by the defendants prior to plaintiffs' objections.

- By the Numbers

At the time plaintiffs brought their motion, defendants had collectively produced more than three million pages of documents even after application of industry-standard de-duplication technology.[xliii] Productions included both hard copy documents and ESI including emails, Word documents, electronic calendars, and spreadsheets. In addition, defendant Georgia-Pacific's efforts to develop, validate, and test search terms involved over 1300 hours of work by outside consultants, not including time contributed by employees and outside counsel.[xliv] In contrast, the plaintiffs produced 25,000 pages of documents.[xlv]

- The Arguments

Notwithstanding the robust search term protocols implemented by defendants, plaintiffs argued that the protocol concerning search terms did not satisfy Rule 34. Specifically, plaintiffs claimed that the defendants' keyword, custodian-based search methodology was, at best, likely to find less than 24 percent of responsive documents rather than their proposed approach based on predictive coding. Plaintiffs asserted their approach could locate, at worst, 70 percent of responsive documents at no greater burden.[xlvi] Plaintiffs also attacked both defendants' specific search term mythology and the concept of search terms more broadly.

Defendants countered by pointing out, among other things, that the Seventh Circuit's current eDiscovery principles expressly contemplate the use of "keyword searching."[xlvii] Defendants also pointed out that no court has ever suggested a well-developed search term strategy is legally deficient. In fact, they pointed out that in December 2011, the District of Delaware adopted new ESI policies favoring use of search terms, and similarly, the Federal Circuit recently adopted a Model E-Discovery Order for Patent Cases based entirely on the concept of keyword searching.[xlviii] Defendants also submitted voluminous testimony regarding the efficacy of their protocol. In fact, by August 2012, the Court had heard approximately two full days of expert testimony from both sides regarding the advantages/disadvantages of both search terms and predictive coding. There had also been numerous hearings and case management conferences, as well as extensive briefing, where these issues were addressed by counsel.

- The Stipulation

Ultimately, with encouragement from the Court, the parties reached a stipulation largely resolving their differences. On August 21, 2012, after slight modification at the Court's request, this was entered as an order.

In the stipulation, the plaintiffs agree to drop their demands that defendants use predictive coding on collections and productions related to discovery requests served prior to October 2013, and to meet and confer over the best ESI search and review methodology for any new discovery requests served after October 2013. The plaintiffs do, however, reserve their right to challenge other aspects of the defendants' collections and productions, including prior collections and productions. The defendants, in turn, reserved their rights to argue all collections and productions met or exceeded relevant legal standards. The stipulated order does not take a position on whether those standards have, in fact, been met.

Accordingly, while the ending may seem anticlimactic and inconclusive to some, the hours of evidentiary hearings and extensive briefing regarding the defensibility of search terms and potential advantages of predictive coding still serve to make this a significant case on predictive coding. Indeed, that the plaintiffs' arguments in favor of predictive coding were taken so seriously by the Court despite being made relatively late in the case, and that so much time and money was spent by both sides advocating their positions, is itself a significant development. Never before have the relative merits of search terms and predictive coding been discussed at such length in a judicial proceeding.

[xxiii] *Da Silva Moore v. Publicis Groupe*, No. 1:11-cv-01279 (S.D.N.Y. Feb. 24, 2012) [Doc. 96].

[xxiv] *See, e.g., Federal Circuit Model Order Regarding E-Discovery in Patent Cases; Eastern District of Texas Model Order Regarding E-Discovery In Patent Cases.*

[xxv] 12/2/11 Conf. Tr. [Doc. 51] at 7-8.

[xxvi] 2/8/12 Conf. Tr. [Doc. 88] at 73-74.

[xxvii] 12/2/11 Conf. Tr. [Doc. 51] at 20-21.

[xxviii] 1/4/12 Conf. Tr. [Doc. 71] at 47-48, 51.

[xxix] 1/4/12 Conf. Tr. [Doc. 71] at 51-52; *see also id.* at 57-58; 2/8/12 Conf. Tr. [Doc. 88] at 84.

[xxx] *Da Silva Moore v. Publicis Groupe*, No. 1:11-cv-01279 (S.D.N.Y. Apr. 26, 2012) [Doc. 175].

[xxxi] Matthew Nelson, *First State Court Issues Order Approving the Use of Predictive Coding*, e-discovery 2.0 (Apr. 26 2012), http://www.clearwellsystems.com/e-discovery-blog/2012/04/26/first-state-court-issues-order-approving-the-use-of-predictive-coding-in-ediscovery/.

[xxxii] *Global Aerospace, Inc. v. Landow Aviation, L.P.*, No. CL 61040 (Va. Cir. Ct. April 23, 2012).

[xxxiii] *Id.*

[xxxiv] *Id.*

[xxxv] *Nat'l Day Laborer Organizing Network v. U.S. Immigration & Customs Enforcement Agency*, No. 1:10-Civ-3488, 2012 U.S. Dist. LEXIS 97863 (S.D.N.Y. July 13, 2012).

[xxxvi] *Id.* at *40, *45.

[xxxvii] *Id.* at *45.

xxxviii *Id.* at *46-47 (warning that keyword searches are "'not nearly as effective at identifying relevant information as many lawyers would like to believe'") (quoting Maura R. Grossman & Terry Sweeney, *What Lawyers Need to Know About Search Tools: The Alternatives to Keyword Searching Include Linguistic and Mathematical Models for Concept Searching*, Nat. L. J. (Aug. 23, 2010)).

xxxix *Id..* at *46 n.110.

xl *Id.* at *49-50 & n.115 (citing *Sedona's Commentary on Achieving Quality in the E-Discovery Process* (2009) *and Commentary on Search & Retrieval Methods* (2007)).

xli *Kleen Products, LLC v. Packaging Corp. of Am.*, 775 F. Supp. 2d 1071 (N.D. Ill. 2011).

xlii It should be noted that the Sedona Conference and other authorities emphasize cooperation, and if necessary, court intervention to settle eDiscovery disputes at the outset in an effort to avoid eleventh hour challenges to an entire production.

xliii 5/22/12 Conf. Tr. at 3.

xliv 2/6/12 Defs.' Br. [Doc. 288] at 8.

xlv 5/22/12 Conf. Tr. at 3.

xlvi 2/6/12 Pls.' Mem. [Doc. 290] at 5-6.

xlvii *See, e.g.*, Seventh Circuit Electronic Discovery Committee, Principles Relating to the Discovery of Electronically Stored Information ("Seventh Circuit Principles" at Principle 2.05, revised Aug. 1, 2010).

xlviii *Federal Circuit Model Order Regarding E-Discovery in Patent Cases; Eastern District of Texas Model Order Regarding E-Discovery In Patent Cases*.