

**UNITED STATES DISTRICT COURT  
SOUTHERN DISTRICT OF NEW YORK**

**MONIQUE DA SILVA MOORE,** )  
**MARYELLEN O'DONOHUE,** )  
**LAURIE MAYERS, HEATHER** )  
**PIERCE, and KATHERINE** )  
**WILKINSON on behalf of themselves** )  
**and all others similarly situated,** )  
 )  
**PLAINTIFFS,** )  
 )  
**v.** )  
 )  
**PUBLICIS GROUPE SA and** )  
**MSLGROUP,** )  
 )  
**DEFENDANTS.** )  
\_\_\_\_\_ )

**Civ No. 11-CV-1279 (ALC) (AJP)**

**PLAINTIFFS' RULE 72(a) OBJECTION TO THE MAGISTRATE'S  
FEBRUARY 8, 2012 DISCOVERY RULINGS**

**TABLE OF CONTENTS**

**I. INTRODUCTION.....1**

**II. BACKGROUND.....3**

**A. Predictive Coding in General.....3**

**B. Predictive Coding in this Employment Discrimination Case.....4**

**III. LEGAL STANDARD.....6**

**IV. ARGUMENT.....7**

**A. The Magistrate’s Decision to Adopt MSL’s Predictive Coding Method Is Clearly Erroneous and Contrary to Law.....7**

**1. The Magistrate’s Decision Sanctioning the Use of Predictive Coding in this Employment Discrimination Case Violates Rule 26 of the Federal Rules of Civil Procedure.....7**

**2. The Magistrate’s Decision to Abandon a Federal Court’s Gatekeeping Function and Adopt a Novel Discovery Methodology Without Supporting Evidence was Clearly Erroneous and Contrary to Law.....9**

**3. The Magistrate’s Decision to Adopt a Novel Discovery Methodology Without Standards for Assessing Reliability was Clearly Erroneous and Contrary to Law.....13**

**4. The Measurements Set Forth in MSL’s Method are Not Reliable Standards.....17**

**5. The Magistrate’s Additional Rulings on February 8, 2012 were Clearly Erroneous and Contrary to Law.....18**

**V. CONCLUSION.....21**

## I. INTRODUCTION

On February 8, 2012, as a matter of first impression, Judge Peck ordered the use of “predictive coding”<sup>1</sup> -- a novel discovery tool -- as the sole determinate of what electronic discovery will be produced and withheld in this employment discrimination class action.<sup>2</sup> Adopting Defendant MSL’s electronically stored information (“ESI”) protocol without real modification, the Magistrate condoned the use of Recommind’s (MSL’s e-discovery vendor) unique application of its predictive coding software that lacks any generally accepted reliability standards (hereinafter “MSL’s predictive coding method” or “MSL’s method”). As explained in detail below, no evidence was presented to the Magistrate supporting MSL’s predictive coding method established in its protocol and had the Magistrate required such evidence to be formally presented, its unreliability would have been clearly evident.

The Text Retrieval Conference (TREC), a project of the National Institute of Standards and Technology (NIST), a federal agency within the U.S. Department of Commerce and the Department of Defense, has established reliability standards for e-discovery search methods.<sup>3</sup> TREC has tested predictive coding in general. *See* Decl. of

---

<sup>1</sup> Predictive coding is software used to conduct automated or computer-assisted document review.

<sup>2</sup> The Magistrate Judge ordered Plaintiffs to file their objections to his February 8, 2012 rulings on February 22, 2012. *See* Dkt. No. 90. However, on February 22, 2012, the Magistrate Judge issued a written order in regards to the February 8, 2012 discovery hearing. Plaintiffs reserve their right to appeal other aspects of the Magistrate’s February 22, 2012 Order pursuant to Rule 72.

<sup>3</sup> “NIST’s mission is to promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life.” *See* [http://www.nist.gov/public\\_affairs/general\\_information.cfm](http://www.nist.gov/public_affairs/general_information.cfm).

NIST has, for many years, been studying the effectiveness of a variety of search methodologies, in cooperation with the Department of Defense. *See Victor Stanley, Inc. v. Creative Pipe, Inc.*, 250 F.R.D. 251, 260 n.10 (D. Md. 2008). “This project, known as the Text Retrieval Conference (TREC), evolved into the Trec LegalTrack, a research effort aimed at studying the e-discovery review process to evaluate the effectiveness of a wide array of search methodologies.” *Id.* As recognized by numerous courts, “[t]he goal of the project is to create industry best practices for use in electronic discovery. This project can be expected to identify both cost effective and reliable search information retrieval methodologies and best

Siham Nurhussein in Support of Plaintiffs' Rule 72(a) Objections ("Decl.") at ¶ 8. According to TREC, predictive coding results vary in accuracy from 10 - 67%. *Id.*

Here, Plaintiffs will never know if MSL's predictive coding method is accurate because their proposal lacks any defined standards regarding: (1) relevance; (2) accuracy (i.e., recall and precision); or (3) acceptance tests. *See e.g.* Decl. of Paul J. Neale, DOAR Litigation Consulting ("Neale Decl."). Moreover, at MSL's direction, Recommind abandoned its existing predictive coding method to create a new method for this case, contrary to the requirements of its own patent, its promotional material, and its involvement in TREC.<sup>4</sup> Indeed, MSL's counsel admitted to the Court that the purported "standards" used in MSL's method have "no science to it." *See* Decl. at ¶ 2 - Feb. 8, 2012 Tr. at 74. ("That was a number that we picked. **There is no science to it**") (emphasis added). When presented with this troubling information, Judge Peck accepted MSL's method without modification.

This Court should reverse Judge Peck's decision to order the use of MSL's unreliable method. It is clearly erroneous and contrary to law. First, the decision is contrary to Federal Rule of Civil Procedure 26. Applying MSL's method, Plaintiffs will be denied relevant discovery and MSL will be largely released of its discovery obligations, including MSL's duty under Rule 26(g) to certify that its production is "complete and correct." Second, the decision is contrary to Federal Rule of Evidence 702. It does not rely on any evidence (either in testimony, affidavits, or independent

---

practice recommendations, which, if adhered to, certainly would support an argument that the party employing them performed a reasonable ESI search, whether for privilege review or for other purposes." *Id.*

<sup>4</sup> As noted by Plaintiffs' expert at the hearing, "[MSL's method] is also inconsistent with [Recommind's] patent, which suggests that you do the iterations until the system tells you it's got it right. Speaking to the limit on that without having done it is **not consistent with your own patent and with what is generally accepted as best practice.**" (Feb. 8, 2012 Tr. at 86)(emphasis added).

studies) supporting either the use of the novel predictive coding technology or MSL's particular method. MSL's method is so lacking in scientific reliability that judicial acceptance of the method at any stage in federal litigation violates the gatekeeping function underlying Rule 702. Finally, if the decision stands, it will have broad implications beyond this case.<sup>5</sup>

## II. BACKGROUND

### A. Predictive Coding in General

Predictive coding is a novel approach to the production of ESI. Indeed, "no reported case (federal or state) has ruled on the use of computer-assisted coding. While anecdotally it appears that some lawyers are using predictive coding technology, it also appears that many lawyers (and their clients) are waiting for a judicial decision approving of computer-assisted review." Magistrate Judge Andrew Peck, *Search Forward*, Legal Technology News, October 1, 2011 (referenced by the Court in Dkt. 58 at 2).

Put simply, predictive coding aims to replace manual document review by attorneys with a review by a computer. Predictive coding is different than keyword searches, which involve picking certain words as your "keywords" and then having the computer search for either a keyword or some combination of keywords within a set of

---

<sup>5</sup> See LegalTech New York 2012, January 30 - February 1, 2012, *Man vs. Machine: The Promise/Challenge of Predictive Coding and Other Disruptive Technologies*, (panelists include MSL's counsel from Jackson Lewis LLP and Magistrate Judge Andrew Peck) [http://www.legaltechshow.com/r5/cob\\_page.asp?category\\_id=72043&initial\\_file=cob\\_page-ltech\\_agenda.asp](http://www.legaltechshow.com/r5/cob_page.asp?category_id=72043&initial_file=cob_page-ltech_agenda.asp); <http://www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=1202542221714&slreturn=1> <http://ediscoveryjournal.com/2012/02/the-honorable-andrew-j-peck-on-the-record-with-predictive-coding-early-headlines-get-it-wrong/>; [http://www.abajournal.com/news/article/is\\_federal\\_magistrate\\_the\\_first\\_to\\_require\\_computerized\\_predictive\\_coding\\_p/](http://www.abajournal.com/news/article/is_federal_magistrate_the_first_to_require_computerized_predictive_coding_p/); <http://www.edupdate.com/2012/02/predictive-coding-in-andrew-j-pecks-court.html>; <http://postmodern-ediscovery.blogspot.com/2012/02/draft-litigators-technical-primer-for.html>; <http://www.legaltechtoday.com/2012/02/14/judge-peck-orders-predictive-coding-in-federal-case-law-technology-news-sean-doherty/>; <http://electronicdiscovery.info/predictive-coding-encouraged-by-the-courts-electronic-discovery/>; <http://www.pinhawkblog.com/peck-picked-peck-predictive-peppers/>; <http://www.ediscoverylawreview.com/2012/02/articles/southern-district-of-new-york-poised-to-address-predictive-coding/>; <http://abovethelaw.com/tag/andrew-j-peck/>.

documents. Predictive coding assumes a computer system can be “trained,” using a set of exemplars, to review a party’s ESI and predict what of that ESI is relevant. Although the system uses attorneys at the beginning to identify exemplars that will “train” the system, and attorneys are involved in a few checks when the system is being “trained,” once the system is “trained,” the attorneys only review the subset of documents that the computer determines are relevant. Attorneys do not review the entire universe of documents. Thus, if objectively relevant documents are not determined to be relevant by the system, they are withheld from the requesting party.

**B. Predictive Coding in this Employment Discrimination Case**

During the parties’ first hearing before Judge Peck on December 2, 2011, Defendant MSL raised the use of predictive coding in this case:

[MSL]: . . . I think right now there are two core disputes as relates to discovery. The first is plaintiffs’ reluctance to utilize predictive coding . . . .

THE COURT: You must have thought you died and went to Heaven when this was referred to me.

(Dec. 2, 2011 Tr. at 8.)

THE COURT: Now if you want any more advice, for better or for worse on the ESI plan and whether predictive coding should be used, or anything else, if the case – I will say right now, what should not be a surprise, I wrote an article in the October Law technology News called Search Forward, which says predictive coding should be used in the appropriate case.

(*Id* at 20.) In light of these statements and Defendant MSL’s unilateral decision to use and implement predictive coding prior to any real agreement between the parties,

Plaintiffs were left with limited choices: either consider the use of predictive coding in this case or seek relief from Judge Peck.

Most notably, Plaintiffs' reluctance about predictive coding is related to the asymmetrical e-discovery sources in employment discrimination cases like this one. Like most employment discrimination cases, the employer here, Defendant MSL, controls nearly all the discovery Plaintiffs will need to prove their case. Indeed, Plaintiffs have virtually no access to relevant evidence except through the discovery methods available under the Federal Rules of Civil Procedure.

As noted by Judge Peck on December 2, 2012:

What matter[s] is whether [Defendants] have completed production. Yes, **I understand that as the defendant in an employment case, [Plaintiffs are] going to have virtually nothing, [Defendants] have everything**, and it is more expensive, et cetera, et cetera. That's what happens when you work for Jackson Lewis, you represent defendants. I am being facetious, but the question is not how much [Defendants] have produced, but **what haven't [Defendants] produced.**

*See* Dec. 2, 2011 Tr. at 6 (emphasis added.) Unfortunately, MSL has taken advantage of this one-sided reality, and thus far has produced only a few hundred documents despite numerous requests served by Plaintiffs. Indeed, seemingly confident that Judge Peck would agree, MSL has withheld the vast majority of documents (even if they exist in paper format or were compelled by Judge Sullivan) based on its proposed predictive coding protocol. Thus, in this case, MSL's reliance on this novel and unproven e-discovery tool to fulfill its every discovery obligation has taken on a heightened importance.

Although the parties raised these issues with Judge Peck, on February 8, 2012, the Court adopted MSL's method virtually wholesale -- without conducting any check of its

reliability standards, or Plaintiffs' ability to verify the accuracy of the results. Predictive coding, like other electronic search tools, is only as good as its design. Much like keyword searches, not all types of predictive coding are built the same. Because MSL's method lacks the necessary standards to ensure its accuracy and thus its reliability as an electronic information retrieval methodology, the use of MSL's method is contrary to the generally accepted standards of other developers in this fairly small market, and to Recommind's own patent. In addition, MSL's method in its current form lacks any means for the parties and the Court to check that the methodology and its results are accurate. Thus, Judge Peck's adoption of MSL's unreliable method is clearly erroneous and contrary to law.

### **III. LEGAL STANDARD**

Rule 72(a) provides that "a district judge in the case must consider timely objections (to a nondispositive order) and modify or set aside any part of the order that is clearly erroneous or contrary law." Fed. R. Civ. P. 72(a). Similarly, "the Federal Magistrates Act, 28 U.S.C. § 636(b)(1)(A), provide[s] that a district court shall reverse a magistrate judge's ruling regarding a non-dispositive matter only where the order is 'clearly erroneous or contrary to law.'" *In re Consol. RNC Cases*, [No Number in Original], 2009 U.S. LEXIS 40293, at \*16 (S.D.N.Y. Jan. 8, 2009) (quoting 28 U.S.C. § 636(b)(1)9A)).

"A ruling is 'clearly erroneous' if the reviewing court is 'left with the definite and firm conviction that a mistake has been committed.'" *Id.* (quoting *Easley v. Cromartie*, 532 U.S. 234, 242 (2001)). "A ruling is 'contrary to law' 'when it fails to apply or misapplies relevant statutes, case law or rules of procedure.'" *Id.* (quoting *Thompson v.*



*Keane*, No. 95 Civ. 2442 (SHS), 1996 U.S. Dist. LEXIS 6022, at \*1 (S.D.N.Y. May 6, 1996) (internal quotation marks omitted)).

#### **IV. ARGUMENT**

##### **A. The Magistrate’s Decision to Adopt MSL’s Predictive Coding Method Is Clearly Erroneous and Contrary to Law**

###### **1. The Magistrate’s Decision Sanctioning the Use of Predictive Coding in this Employment Discrimination Case Violates Rule 26 of the Federal Rules of Civil Procedure**

MSL’s predictive coding method, as adopted by the Court, is so lacking in reliability that MSL’s ultimate email production will be alarmingly under-inclusive. Rule 26(b)(1) permits discovery, absent a valid objection, of “any non-privileged matter that is relevant to any party’s claim or defense . . . [;] Relevant information need not be admissible at trial if the discovery appears reasonably calculated to lead to the discovery of admissible evidence.” Fed. R. Civ. P. 26(b)(1) (emphasis added). Courts have construed the term “relevance” broadly to encompass “any matter that bears on, or that reasonably could lead to other matters that could bear on, any issue that is or may be in the case.” *Oppenheimer Fund, Inc. v. Sanders*, 437 U.S. 340, 351 (1978). Before limiting discovery, Rule 26 requires that the Court balance Plaintiffs’ need to discover relevant evidence with the burden on MSL to review and produce responsive electronically stored information.<sup>6</sup>

Here, the Court failed to properly balance MSL’s burden against Plaintiffs’ need to obtain evidence to prove their claims. Notably, the Court did not require MSL to submit formal evidence demonstrating that Recommind will reliability capture a

---

<sup>6</sup> See Fed. R. Civ. P. 26 Advisory Committee’s Note. (providing that “subparagraph (B) [of FRCP 26 was] added to regulate discovery from [electronic] sources[;] [u]nder this rule, a responding party should produce electronically stored information that is relevant, not privileged, and reasonably accessible, subject to the (b)(2)(C) limitations that apply to all discovery.”)

sufficient number of relevant documents from the total universe of documents in existence. Instead, the Court advocated for use of predictive coding as a viable discovery tool, disregarding most if not all of Plaintiffs' evidence and arguments, demonstrated in great detail below, that MSL's proposal will *grossly* exclude a large volume of responsive email from MSL's ultimate production. In so doing, the Court's ruling runs afoul of Rule 26 and Magistrate's Peck's prior holdings in another matter, that electronic document search methodologies be carefully crafted and tested for quality assurance.<sup>7</sup> *See Victor Stanley, Inc.*, 250 F.R.D. 250, 260-62.

Importantly, Magistrate Peck's ruling is also contrary to law given the fact predictive coding is being used "[i]n the context of [an] employment discrimination case" where courts have traditionally favored "'liberal civil discovery rules,' giving plaintiffs 'broad access to employers' records in an effort to document their claims.'" *Vuona v. Merrill Lynch \* Co.*, No. 10 Civ. 6529, 2011 U.S. Dist. LEXIS 131491, at \*9 (S.D.N.Y. Nov. 15, 2011) (quoting *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642, 657 (1989)).<sup>8</sup> "This is especially the case with regard to discrimination claims, where the imposition of unnecessary discovery limitations [are] to be avoided." *Chan v. NYU Downtown Hosp.*, No. 03 Civ. 3003, 2004 U.S. Dist. LEXIS 16751, at \*12 (internal citation omitted) (overturning magistrate judge's prior ruling regarding discoverability of employment

---

<sup>7</sup> Magistrate Peck's ruling also provides unlawful "cover" for MSL's counsel, who has a duty under FRCP 26(g) to "certify" that their client's document production is "complete" and "correct" as of the time it was made. FRCP 26(g)(1)(A). Essentially, since Magistrate Peck sanctioned the use of predictive coding in this case without sufficient reliability and quality controls in place, defense counsel can then "certify" that the resulting production is complete without taking the steps necessary to ensure that such a certification meets the standard under the Federal Rules.

<sup>8</sup> *See also EEOC v. Caesars Entm't, Inc.*, 237 F.R.D. 428, 431 (D. Nev. 2006) (internal citation omitted) ("In deciding whether to restrict discovery under Rule 26(b)(2), 'the court should consider the totality of the circumstances, weighing the value of the material sought against the burden of providing it, and taking into account society's interest in furthering the truth-seeking function in the particular case before the court'").

discrimination complaints in a Title VII matter) (J. Baker Motley). This is so in part because, civil rights litigants, unlike litigants in the intellectual property context, for example, largely depend on *circumstantial* evidence to prove Title VII violations.<sup>9</sup> For example, in the seminal e-discovery case, *Zubulake v. UBS Warburg LLC*, which involved an age discrimination litigant, the court noted that email discovery is particularly necessary in the electronic age and certainly in discrimination cases, where Plaintiffs frequently seek to use e-mails as evidence of discriminatory behavior. *See e.g., Zubulake v. UBS Warburg LLC*, 220 F.R.D. 212, 215 (S.D.N.Y. 2003). Indeed, if restrictive and unreliable predictive coding proposals, like the one proposed by MSL and adopted by the Court, were to gain acceptance in the employment discrimination context, Title VII's mandate to eradicate discrimination could be chilled.<sup>10</sup> Based on these public policy considerations alone, Magistrate Peck's ruling should be reversed.

## **2. The Magistrate's Decision to Abandon a Federal Court's Gatekeeping Function and Adopt a Novel Discovery Methodology Without Supporting Evidence was Clearly Erroneous and Contrary to Law**

---

<sup>9</sup> In a gender discrimination case, in the absence of direct evidence of discrimination, a plaintiff proves a prima facie case of gender discrimination by showing, via circumstantial evidence, that: (1) she belongs to a protected class; (2) she performed her job satisfactorily; (3) she suffered an adverse employment action; and (4) her employer treated her differently than similarly situated employee(s). *See Gerving v. Opbiz, LLC*, 324 Fed. Appx. 692, 694 (9th Cir. 2009).

<sup>10</sup> *See e.g., Rivera v. NIBCO, Inc.*, 364 F.3d 1057, 1065-66 (9th Cir. 2003) (recognizing that discovery that would cause a chilling effect on plaintiffs seeking to enforce their employment rights was a concern that should be evaluated when deciding discovery disputes); *EEOC v. Bice of Chicago*, 229 F.R.D. 581 (N.D. Ill. 2005) (taking into account whether the discovery tactics of the defendant-employer would place a substantial burden on the plaintiffs and the public interest, and would have a chilling effect on victims of employment discrimination coming forward to assert claims). Additionally, because Title VII entrusts private litigants to enforce federal civil rights law, this case "involves issues important to the public." *See Glenmede Trust*, 56 F.3d at 483; *see e.g. Marek v. Chesny*, 473 U.S. 1, 32 (1985) ("[C]ivil rights plaintiffs 'appear before the court cloaked in a mantle of public interest' . . . to promote the 'vigorous enforcement of modern civil rights legislation.'"); *Albermarle Paper Co. v. Moody*, 422 U.S. 405, 415 (1975) ("There is, of course, an equally strong public interest in having injunctive actions brought under Title VII, to eradicate discriminatory employment practices); Section-By-Section Analysis of H.R. 1746, accompanying The Equal Employment Opportunity Act of 1972 Conference Report, 118 Con. Rec. 7166, 7168 (1972) ("[C]laims under the Title VII involve the vindication of a major public interest.").

MSL's predictive coding method is a new technology whose efficacy "certainly requires knowledge beyond the ken of a lay person (and a lay lawyer) . . . ." *Equity Analytics, LLC v. Lundin*, 248 F.R.D. 331, 333 (D.D.C. 2008); *see also Victor Stanley, Inc. v. Creative Pipe, Inc.*, 250 F.R.D. 251, 260 n.10 (D. Md. 2008) ("[R]esolving contested issues of whether a particular search and information retrieval method was appropriate . . . involves scientific, technical or specialized information."). Indeed, Judge Peck's own opinion, to which all litigants are referred in his individual practices, appears to endorse Judge Facciola's statement that given the complexities of electronic search protocols, "[t]his topic is clearly beyond the ken of a layman and requires that any such conclusion be based on evidence that, for example, meets the criteria of Rule 702 of the Federal Rules of Evidence." *Gross Construction Assocs.*, 256 F.R.D. at 135 (quoting *United States v. O'Keefe*, 537 F. Supp. 2d 14, 24 (D.D.C. 2008)).

When a trial judge is asked to resolve conflicts regarding the reliability or appropriateness of a novel methodology, the trial judge must rely on the aid of experts. *Victor Stanley, Inc.*, 250 F.R.D. at 260 n.10 (citing *Equity Analytics*, 248 F.R.D. at 333 and *United States v. O'Keefe*, 537 F. Supp. 2d 14, 24 (D.D.C. 2008))("[C]hallenges to the sufficiency of keyword search methodology unavoidable involves scientific, technical and scientific subjects, and *ipse dixit* pronouncements from lawyers unsupported by an affidavit or other showing that the search methodology was effective for its intended purpose are of little value to a trial judge . . . ."). In other words, "the trial judge must decide a method's appropriateness with the benefit of information from some reliable source – whether an affidavit from a qualified expert, a learned treatise, or, if appropriate, from information judicially noticed." *Id.* Indeed, judicial reliance on methodologies

developed by judges that do not rely on expert assistance has been “characterized as arbitrary and lacking logical foundation.” *Id.* (citing *Am. Nat’l Bank & Trust Co. v. Equitable Life Assurance Society*, 406 F.3d 867, 879 (7th Cir. 2005)) (internal quotation marks omitted).

Judges are well equipped to determine the reliability of such evidence; after parties submit evidence which supports their methodology – whether it be in the form of in person testimony, affidavits, treatises, or requests for judicial notice – a judge can turn to the standards established by Federal Rule of Evidence 702, “to determine whether the information is helpful” to the court’s “factual determinations involving disputed areas of science, technology or other specialized information.” *Id.* Using Rule 702, in conjunction with the principles outlined in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), forces the district court to function as a gatekeeper for expert testimony that the court itself will rely upon in making a factual determining about the reliability of an information retrieval methodology. *Cf. Major League Baseball Props., Inc. v Salvino, Inc.*, 542 F.3d 290, 311 (2d Cir. 2008). For example, using *Daubert* principles, a court might permit the parties to submit additional information that suggests that the expert evidence submitted to the Court is unreliable, perhaps because the purported expert lacks qualifications or has a financial interest in the case, because the opinions are unsupported by sufficient facts, or because the opinions are based on demonstrably unreliably methodology. *See Victor Stanley, Inc.*, 250 F.R.D. at 260 n.10.

Here, at no point did the Magistrate review any evidence to support his decision. The Magistrate took no judicial notice of any documents or studies that support the reliability of MSL’s method, nor did he receive any affidavits or declarations from

purported experts that supported the methodology of MSL's method. To his credit, the Magistrate did ask the parties to bring the ESI experts they had hired to advise them regarding the creation of an ESI protocol. These experts, however, were never sworn in, and thus the statements they made in court at the hearings were not sworn testimony made under penalty of perjury. The Magistrate judge never asked for or evaluated the qualifications of these experts, nor were the parties given an opportunity to question or cross-examine the experts in order for the Court to make a finding regarding the reliability of the experts' opinions. Thus, the Magistrate's decision relies only on the arguments made by counsel.

Even if the Court disagrees and find that the Magistrate's reliance on expert argument qualifies as reliance on expert testimony, the standard the Magistrate used to evaluate the admissibility of MSL's expert's statements was clearly erroneous and contrary to law, because he deployed no standards whatsoever. To that end, Plaintiffs have included a declaration from their experts that speak to the reliability of the methodology proposed by MSL and the statements made by MSL's experts. *See e.g.* Neale Decl. Moreover, as will be further discussed below, statements made by MSL's experts about the reliability of its product are not supported by Recommind's own patent or generally accepted industry practices.

In short, the Magistrate's ruling does not rely on evidence that was determined to be "information from some reliable source," either because no evidence was admitted, or because no standards of admissibility were considered by the Court, therefore, the ruling is clearly erroneous and contrary to law.

**3. The Magistrate’s Decision to Adopt a Novel Discovery Methodology Without Standards for Assessing Reliability was Clearly Erroneous and Contrary to Law**

MSL’s method lacks the necessary standards for assessing whether its results are accurate; in other words, there is no way to be certain if MSL’s method is reliable. The Magistrate has, himself, previously recognized the importance of such checks, “[when conducting searches of electronic discovery] the proposed methodology must be quality control tested to assure accuracy in retrieval and elimination of false positives.” *William A. Gross Construction Assocs., Inc. v. Am. Manufactueres Mut. Ins. Co.*, 256 F.R.D. 134, 136 (S.D.N.Y. 2009). “Common sense dictates that sampling and other quality assurance techniques must be employed to meet requirements of completeness.” *Id.* at 135-36 (quoting *In re Seroquel Products Laibility Litig.*, 244 F.R.D. 650, 662 (M.D. Fla. 2007)). Moreover, the Magistrate supports Judge Grimm’s opinion that *Gross Construction Assocs.*, 256 F.R.D. at 135 (quoting *Victor Stanley, Inc. v. Creative Pipe, Inc.*, 250 F.R.D. 251, 262 (D. Md. 2008)) (“[s]election of the appropriate search and information retrieval technique requires careful advance planning by persons qualified to design effective search methodology. The implementation of the methodology selected should be tested for quality assurance; and **the party selecting the methodology must be prepared to explain the rationale for the method chosen to the court, demonstrate that it is appropriate for the task, and show that it was properly implemented**”) (emphasis added)<sup>11</sup> Unfortunately, the Magistrate’s acceptance of MSL’s method on February 8, 2012, does not accord with his prior holdings.

---

<sup>11</sup> Indeed, in an article he published advising as to what he would look for in predictive coding, The Magistrate wrote, “Proof of a valid ‘process,’ including quality control testing, also will be important.” Andrew Peck, *Search, Forward*, Law Technology News, October 1, 2011.

MSL's method does not have sufficient quality assurance measurements in place to determine whether the methodology is reliable. Unlike other e-discovery protocols that use predictive coding methodologies, MSL's method fails to include explicit and defined standards that the protocol must meet, in order for the parties and the Court to conclude that the technology reached a level of accuracy that was acceptable to the parties in producing relevant documents as is required under Rule 26. In other words, MSL's method asks the parties to simply trust in its reliability, and provides no metrics by which it can be judged. Such a lack of explicit defined standards is contrary to generally accepted industry practices. *See e.g.* Neale Decl.

As noted by Plaintiffs' expert, Paul J. Neale, there are three essential elements that are missing from MSL's method; the absence of any one of these essential elements makes MSL's method fundamentally flawed. *See e.g.* Neale Decl. First, MSL's method fails to include an agreed-upon standard of relevance that is transparent and accessible to all parties. *See e.g.* Neale Decl. Actual relevancy of documents is the basis for all of the calculations that allow the parties and the Court to determine whether MSL's method is accurate in differentiating between relevant and non-relevant documents. (Neale Decl.) Without this standard, there a high-likelihood of delay as the parties resolve disputes with regard to individual documents on a case-by-case basis. *See e.g.* Neale Decl. Moreover, it makes it impossible for a third party investigator or auditor to replicate the results, should that be necessary. *See e.g.* Neale Decl.

Second, MSL's method fails to set forth in advance how it will make any of the calculations that are necessary to determine whether the system is accurate in identifying responsive documents. *See e.g.* Neale Decl. As noted by Plaintiffs' expert, researchers



in the field of information retrieval have developed effective and accurate approaches to measuring the variables necessary to calculate the accuracy of a given system. *See e.g.* Neale Decl. Highly detailed descriptions of how one measures and calculates those variables are standard in any adequate description of a protocol used in retrieving e-discovery. *See e.g.* Neale Decl. These descriptions help the parties set goals of accuracy, force accountability of the methodology, and prevent later manipulations. As noted by Plaintiffs' expert, these descriptions are missing from MSL's method. A simple review of the MSL's method reveals only vague language that uses the words involved in taking such measurements, but does not set forth anything specific about how such a measure will be taken.

Indeed, as set forth in his declaration, Plaintiffs' expert uses his knowledge to determine any potential ways of measuring the variables needed to calculate accuracy. Even if Plaintiffs' expert gives MSL's method the benefit of the doubt and assumes that it is suggesting validation techniques, as Plaintiffs' expert notes, such validation techniques are completely inadequate. *See e.g.* Neale Decl.

Finally, MSL's method fails to state the standard of acceptance that they are trying to achieve. In other words, how the parties determine whether, at the end of the day, MSL's method actually works. *See e.g.* Neale Decl. MSL's method states it will take a random sample of the documents that the system has deemed to be irrelevant and attorneys will review the sample to see if, in fact, any of these non-produced documents are relevant. MSL's method fails to state what happens next. For example, if the attorneys find out of a random sample of 2,399 deemed-irrelevant documents, there are twenty that are relevant, is that considered acceptable under the protocol and is the

methodology deemed successful and accurate? Must there be no relevant documents or a low number found in the random sample? If so, what should that number be? As is immediately apparent, without any decision about this made in advance, the Court is simply kicking the can down the road in terms of disputes of failures to produce under Rule 26.

Indeed, the lack of standards is compounded by additional information that suggests that MSL's method is not a reliable information retrieval methodology. As stated above, the National Institute of Standards and Technology ("NIST") is a federal agency within the U.S. Department of Commerce. "NIST's mission is to promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life." [http://www.nist.gov/public\\_affairs/general\\_information.cfm](http://www.nist.gov/public_affairs/general_information.cfm). NIST has, for many years, been studying the effectiveness of a variety of search methodologies, in cooperation with the Department of Defense. *See Victory Stanley, Inc.*, 250 F.R.D. at 261 n.10. "This project, known as the Text Retrieval Conference ("TREC"), evolved into the Trec LegalTrack, a research effort aimed at studying the e-discovery review process to evaluate the effectiveness of a wide array of search methodologies." *See* [http://www.nist.gov/public\\_affairs/general\\_information.cfm](http://www.nist.gov/public_affairs/general_information.cfm) As recognized by numerous courts, "[t]he goal of the project is to create industry best practices for use in electronic discovery. This project can be expected to identify both cost effective and reliable search information retrieval methodologies and best practice recommendations, which, if adhered to, certainly would support an argument that the party employing them performed a reasonable ESI search, whether for privilege review or for other purposes."

*Id.* Just as the courts believe that testing by TREC might support a methodology's reliability, so too should the Courts be willing to use TREC's measurements of accuracy to support the fact that a methodology is not reliable.

Finally, as Defendants will no doubt argue, the Magistrate held that many of the issues raised by Plaintiffs could be revisited at some later point in the litigation, but that he wanted to "see how it works." (Feb. 8, 2012 Tr. at 83-84, 87.) Without standards in the protocol to check on the accuracy of MSL's method, however, the Court is without any ability to determine "how it works" and thus without any evidence to revisit such issues. Therefore, the Court should find that the adoption of MSL's method, without any standards to review its reliability, is clearly erroneous and contrary to law.

#### **4. The Measurements Set Forth in MSL's Method are Not Reliable Standards**

MSL also admitted that some of the methods by which it purports to "train" the system are based not on science, but on limiting costs. For example, MSL's expert stated: "[W]e will stop either at the end of the seventh round or if, between two rounds, the number of new documents being brought back is less than 5 percent. That was a number that we picked. **There is no science to it.**" *See* Decl. at ¶ 2 - Feb. 8, 2012 Tr. at 74. Indeed, there is no science to reviewing at least 500 documents, or stopping after either less than 5% of new documents are returned or after seven iterations. As pointed out by Plaintiffs' expert at the hearing, "[i]t is also inconsistent with your patent, which suggests that you do the iterations until the system tells you it's got it right. Speaking to the limit on that without having done it is **not consistent with your own patent and with what is generally accepted as best practice.**" *Id.* at p. 86.

Indeed, Defendants appeared to admit to creating a methodology based on costs rather than on science or reliability:

[PLAINTIFFS' EXPERT]: [Recommind's] patent itself suggests that as a result of this process you should be reviewing 10 to 35 percent of your total document collection, which is supposed to indicate a significant savings, which in this case would be about 300 [thousand] to 1 million documents. They keep talking about 40,000 to 75 as being burdensome and disproportional. If they don't understand the result of the system, what to expect, I don't understand why they are proposing it in the first place.

[MSL]: Your Honor, one of the reasons why we developed this work flow was, again, this is not a case where we are prepared to review a million documents during this first phase. We worked with our vendor and came up with a modified work flow that we believe is defensible but is also reviewing a more reasonable number of documents for this case.

*Id.* at pp. 87-88. The Magistrate rightly held that issues of proportionality and cost shifting are not appropriately reviewed at this point in the process. However, the Magistrate failed to follow his own ruling by accepting a protocol that is based on costs and not on science. A purportedly scientific methodology that is based on costs, rather than science, is inherently unreliable.

##### **5. The Magistrate's Additional Rulings on February 8, 2012 were Clearly Erroneous and Contrary to Law**

Nearly nine months after Plaintiffs served their First Set of Document Demands, MSL has failed to produce complete and accurate payroll data. Initially, MSL refused to produce **any** responsive documents in response to Plaintiffs' request for company-wide employment data (Request No. 1 from Plaintiffs' First Set of Document Demands),<sup>12</sup>

---

<sup>12</sup> **“REQUEST FOR PRODUCTION 1:** Any database or computerized information (to be provided with metadata) used, created or maintained by Defendants, or now in the possession, custody or control of Defendants, from which information can be obtained concerning the following data for each Public Relations professional employed by Defendants in the United States, including but not limited to Account

forcing Plaintiffs to seek Court intervention. After Judge Sullivan ordered the production of Company-wide employment data. *See* Sept. 14, 2011 Order - Dkt. No. 40. MSL produced some employment data from its PeopleSoft database, which contains computerized HR data, but not the payroll data, which indicates what employees are actually paid.

After MSL defied the Court's September 14, 2011 Order compelling production, Plaintiffs requested a pre-motion conference to file a motion for sanctions. At the conference, MSL initially stated that it had produced the compelled payroll data, but after further questioning by the Court, conceded that **Defense counsel had not even asked MSL's payroll department about payroll.** (*Compare* Jan. 4, 2012 Tr. at 6-7 ("We have provided the payroll data and other electronic data." *with id.* at 30 ("They have a payroll department, and I haven't posed this question").) Defense counsel also told the Court and Plaintiffs that MSL did not keep electronic copies of W-2s. (*See, e.g., id.* at 6 ("And we advised plaintiffs' counsel that the Box 5 data from the W-2s were not available electronically.")) The Court ordered MSL to produce the data by January 25 or face sanctions, and on January 26, MSL produced an incomplete version of the compelled

---

Executive through Director, from February 2008 forward (February 2005 for compensation data): a. Name of the employee; b. employee ID or unique identifier; c. gender of the employee; d. age and number of children; e. dates of employment; f. job history, including all job assignment(s)/job title(s) held; g. job function, job sub-function and/or job family, including unit, region, office, division or department worked in; h. compensation, including salary, bonuses, benefits and other forms of compensation by year, including data from Box 5 of W-2 form; i. any promotion received, applied for, or denied, including date on which employee received, applied for or was denied a promotion, and any reason for non-selection for promotion (if applicable); j. any demotion, including date of demotion and any reasons for demotion; k. any resignation or termination, including date of resignation or termination, any reasons for resignation or termination, and whether the employee signed a severance agreement and/or received severance pay; l. any transfer received, including those associated with any restructuring, the date of the transfer, and any reasons for the transfer; m. dates and nature of any leave of absence or other extended absence from employment, including maternity or paternity leave.

payroll data – the document produced was partially redacted and contained many omissions – and failed to produce any W-2 information.

Defense counsel’s willful noncompliance with the Court’s Order and continuous misrepresentations about the way in which data is stored, was confirmed by MSL’s corporate designee on payroll, on February 1, 2012. MSL’s designee testified that Defense counsel **never asked** the payroll department’s litigation contact (who happened to be the designee herself) about payroll data in connection with this litigation. Moreover, MSL’s designee testified that MSL retains W-2 information electronically dating back to 2001, and is in possession of W-2 data in CD format – one CD for all MSL employees per year. Therefore, by simply providing 7 CDs of data dating back to 2005, MSL could fulfill its discovery obligations and comply with the Court’s order to produce W-2 information.

Rather than order Defendants to turn over these CDs, without any explanation or reasoning, Judge Peck shifted the entire burden onto Plaintiffs: “THE COURT: Fine. Here is what you are going to do. You are going to read them at defense counsel's office. No notes can be taken. You will print out what you want to print out page by page only for who you are entitled to the information on. You will then show those copies to Mr. Brecher or his colleagues. Then you will get the copies, assuming they are for the right people.” *See* Feb. 8, 2012 Tr. at 16. Of course, then the Plaintiffs had to retype all of the printed out information into a computer. Thus far, Plaintiffs have already spent over several days reviewing and printing out W-2 forms. Such hidden cost-shifting tactics, with regard to evidence that has been compelled by the Court multiple times, is clearly erroneous and contrary to law.

Moreover, Plaintiffs had the opportunity to receive all such data through a third-party source, on whom a subpoena had previously been filed. The Court, without any motion before it and without making any showing under Rule 45(c)(3), *sua sponte* quashed Plaintiffs' subpoena. *See* Dkt. No. 83. This ruling should be reversed as clearly erroneous and contrary to law.

## V. CONCLUSION

For all of the foregoing reasons, the Court should reverse the Magistrate's February 8, 2012 rulings adopting MSL's ESI protocol and condoning the use of predictive coding, as well as *sua sponte* quashing third party subpoena Plaintiffs' served on ADP, MSL's payroll services company, to obtain payroll data that was compelled by Judge Sullivan on September 14, 2011.

DATED: February 22, 2012

SANFORD WITTELS & HEISLER, LLP

/s/ Janette Wipper

Janette Wipper Esq.

Steven L. Wittels, Esq.

Siham Nurhussein, Esq.

Deepika Bains, Esq.

*Attorneys for the Plaintiffs and the Class*

**CERTIFICATE OF SERVICE**

I, Sheri Pan, hereby certify under penalty of perjury that on this 22<sup>nd</sup> day of February 2012 true and correct copies of the foregoing Rule 72(a) Objection to Magistrate Rulings were served on all counsel of record by operation of the Court's ECF system.

/s/ Sheri Pan

Sheri Pan