

ALBANY
AMSTERDAM
ATLANTA
AUSTIN
BOSTON
CHICAGO
DALLAS
DELAWARE
DENVER
FORT LAUDERDALE
HOUSTON
LAS VEGAS
LONDON*
LOS ANGELES
MEXICO CITY+
MIAMI
NEW JERSEY
NEW YORK
ORANGE COUNTY
ORLANDO
PALM BEACH COUNTY
PHILADELPHIA
PHOENIX
SACRAMENTO
SAN FRANCISCO
SHANGHAI
SILICON VALLEY
TALLAHASSEE
TAMPA
TYSONS CORNER
WASHINGTON, D.C.
WHITE PLAINS

*Strategic Alliances with
Independent Law Firms***

MILAN
ROME

Federal Judge Approves Predictive Coding Technology for e-Discovery

In an eagerly anticipated decision, Magistrate Judge Andrew Peck of the Southern District of New York issued the first judicial opinion formally approving the use of “predictive coding” technology, which offers the promise of substantially reducing costs associated with exhaustive, human review of electronically stored information (ESI).¹ Recognizing the Bar has been hesitant to embrace this technology, Judge Peck wrote: “What the Bar should take away from this Opinion is that computer-assisted review is an available tool and 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. Counsel no longer have to worry about being the “first” or “guinea pig” for judicial acceptance of computer assisted review ... [which] can now be considered judicially-approved for use in appropriate cases.”²

What is Predictive Coding?

The term “predictive coding” (also commonly referred to as “computer-assisted review” or “technology-assisted review”) refers to “tools ... that use sophisticated algorithms to enable [a] computer to determine relevance based on interaction with (*i.e.* training by) a human reviewer.”³ The process involves a senior attorney, with extensive knowledge of the case issues, reviewing and coding a “seed set” of documents sampled from the universe of collected data. This individual document review “trains” the computer to recognize other relevant documents so that it can “predict” the reviewer’s coding for the entire dataset. “When the system’s predictions and the reviewer’s coding sufficiently coincide, the system has learned enough to make confident predictions for the remaining documents.”⁴

Predictive coding technology properly implemented permits quickly narrowing the pool of potentially relevant electronic documents and obviating the need for time-consuming, expensive, manual review of hundreds of thousands, if not millions, of documents. Indeed, Judge Peck recognized that in a case involving large quantities of documents, a “linear manual review is simply too expensive.”⁵

Independent studies suggest that computer-assisted review is as effective, if not more effective, than conventional e-discovery tools, such as keyword and conceptual searches followed by manual, individual document review. Judge Peck noted that “while some lawyers still consider manual review to be the ‘gold standard,’ that is a myth, as statistics clearly show that computerized searches are at least as accurate, if not more so, than manual review.”⁶ Citing several judicial opinions criticizing “keyword searches” commonly used by e-discovery practitioners, Judge Peck observed that parties selecting keywords to cull down potentially responsive records resembles the child’s game “Go Fish,” noting such

searches are at once over-inclusive (locating numerous irrelevant documents) and under-inclusive (citing a study suggesting that keywords returned only 20 percent of relevant documents).⁷

The Predictive Coding Protocol Employed in *Da Silva Moore*

Da Silva Moore is a putative class action brought by female employees of “one of the world’s ‘big four’ advertising conglomerates,” alleging claims of gender and pregnancy discrimination. Defendant MSL collected approximately 3 million documents from custodians agreed upon by the parties. The parties agreed generally with Defendant’s plan to use predictive coding, but Plaintiffs objected to the specific, proposed methodology. Judge Peck ordered the parties to submit an ESI protocol, including predictive coding protocols, to which Plaintiffs objected.

The parties’ ESI protocol is noteworthy for its collaborative approach. The Court required that documents reviewed and issue coded by Defendant’s counsel as the “seed set” to “train” the system include a 2,399 document random sample plus other relevant documents identified by keyword searches suggested by both the Defendant and Plaintiffs. The parties agreed to run seven “iterative rounds,” whereby defense counsel would review a sample of 500 documents in each round to test whether the computer was accurately returning relevant documents. At the end of this process, Defendant would review another sample of documents coded as non-relevant as a quality control check. The protocol also required Defendant to provide its adversary all documents (less privileged documents), including relevance and issue tag codes, at each stage of the process.⁸

Judge Peck’s Ruling

Judge Peck found that predictive coding technology furthers the goal of Rule 1 of the Federal Rules of Civil Procedure to “secure the just, speedy, and inexpensive determination” of lawsuits, as well as the “proportionality doctrine” set forth in Rule 26(b)(2)(C).⁹ He concluded that predictive coding was appropriate in *Da Silva Moore* in light of: (1) the parties’ agreement to use predictive coding; (2) the vast amount of ESI to be reviewed; (3) the superiority of computer-assisted review to the alternatives; (4) the need for cost effectiveness and proportionality under Rule 26(b)(2)(C); and (5) the transparent process proposed by Defendant.¹⁰ The Court also emphasized the importance of the parties embracing transparency when approaching an ESI protocol and “highly recommend[ed] that counsel in future cases be willing to at least discuss, if not agree to, such transparency in the computer-assisted review process.”¹¹

Lessons for the Future

Judge Peck’s opinion – intended for a broader audience than the parties before him – provides litigants with four “lessons for the future” when considering implementing predictive coding technology:

1. With respect to predictive coding protocols, the process needs to be “quality control verified.”
2. e-Discovery consultants have a constructive role to play, including appearing at discovery hearings to address the Court when these issues present (jokingly referred to as “bring your geek to court day”).¹²
3. Consider “staging” discovery, by starting with the most relevant sources of data, as a useful way to control discovery costs.

4. Parties requesting documents with existing knowledge of the producing party's systems should cooperate with their adversary to identify potential sources of ESI and inform the adversary's records review and collection. The Court encouraged both sides to engage in "strategic proactive disclosure of information" in order to achieve a more efficient discovery process.¹³

As the first official judicial acceptance of predictive coding technology, Judge Peck's ruling in *Da Silva Moore* paves the way for clients to embrace and adopt this technology. Whether this technology is appropriate for discovery in any given case depends on a variety of factors, including the volume of documents at issue and the amount in controversy in the dispute. It was important to Judge Peck that the parties agreed to use predictive coding technology, noting that if one party objects to predictive coding, it is a "slightly more difficult question" for a court to approve using this technology. But with *Da Silva Moore* on the books, parties with large datasets at issue have formidable support to encourage Courts to authorize this type of cost-saving technology, even over objections. Predictive coding technology also shows cost-saving promise for purely internal document reviews, such as internal investigations, that do not require buy-in from adversaries.

¹ *Monique Da Silva Moore v. Publicis Groupe & MSL Group*, No. 11 Civ. 1279 (ALC)(AJP), Dkt. No. 96 (Slip Op.) (S.D.N.Y. Feb. 24, 2012) ("*Da Silva Moore*").

² *Da Silva Moore*, slip op. at 25-26.

³ *Id.* at 3.

⁴ *Id.* at 4.

⁵ *Id.* at 18.

⁶ *Id.*

⁷ *Id.* at 20.

⁸ *Id.* at 9-12.

⁹ *Id.* at 21-22.

¹⁰ *Id.* at 22.

¹¹ *Id.* at 23. Plaintiffs promptly appealed Judge Peck's order to the District Court. We anticipate that the District Court will defer to Judge Peck's expertise and affirm the Magistrate Judge's ruling.

¹² *Id.* at 23-24.

¹³ *Id.* at 24.

This *GT Alert* was prepared by Phil Cohen and Ryan Harsch. Questions about this information can be directed to:

- [Phil Cohen](mailto:cohenp@gtlaw.com) | 212.801.2145 | cohenp@gtlaw.com
- [Ryan Harsch](mailto:harschr@gtlaw.com) | 212.801.6420 | harschr@gtlaw.com
- Or your [Greenberg Traurig](#) attorney

Albany
518.689.1400

Las Vegas
702.792.3773

Philadelphia
215.988.7800

Amsterdam
+ 31 20 301 7300

London*
+44 (0)203 349 8700

Phoenix
602.445.8000

Atlanta
678.553.2100

Los Angeles
310.586.7700

Sacramento
916.442.1111

Austin
512.320.7200

Mexico City+
+52 55 5029.0000

San Francisco
415.655.1300

Boston
617.310.6000

Miami
305.579.0500

Shanghai
+86 21 6391 6633

Chicago
312.456.8400

New Jersey
973.360.7900

Silicon Valley
650.328.8500

Dallas
214.665.3600

New York
212.801.9200

Tallahassee
850.222.6891

Delaware
302.661.7000

Orange County
949.732.6500

Tampa
813.318.5700

Denver
303.572.6500

Orlando
407.420.1000

Tysons Corner
703.749.1300

Fort Lauderdale
954.765.0500

Palm Beach County North
561.650.7900

Washington, D.C.
202.331.3100

Houston
713.374.3500

Palm Beach County South
561.955.7600

White Plains
914.286.2900

*This Greenberg Traurig Alert is issued for informational purposes only and is not intended to be construed or used as general legal advice. Please contact the author(s) or your Greenberg Traurig contact if you have questions regarding the currency of this information. The hiring of a lawyer is an important decision. Before you decide, ask for written information about the lawyer's legal qualifications and experience. Greenberg Traurig is a service mark and trade name of Greenberg Traurig, LLP and Greenberg Traurig, P.A. ©2012 Greenberg Traurig, LLP. All rights reserved. *Operates as Greenberg Traurig Maher LLP. **Greenberg Traurig is not responsible for any legal or other services rendered by attorneys employed by the strategic alliance firms. +Greenberg Traurig's Mexico City office is operated by Greenberg Traurig, S.C., an affiliate of Greenberg Traurig, P.A. and Greenberg Traurig, LLP.*