

JUDGE PECK ENDORSES PREDICTIVE CODING

Conor R. Crowley, Esq., CIPP/US¹

On February 8, 2012, the parties in *Da Silva Moore, et al., v. Publicis Groupe, et al.* appeared before Magistrate Judge Andrew Peck in the Southern District of New York to seek a ruling from Judge Peck on a variety of discovery issues. Among those issues was the protocol to be employed by the defendant in using predictive coding. It is important to note that nowhere in the transcript of the February 8 hearing does Judge Peck mandate the use of predictive coding. Furthermore, Judge Peck does not devise, and impose on the parties, a protocol for the use of predictive coding. However, he does acknowledge the efficacy of the use of predictive coding for making relevance determinations, whilst acknowledging that perfection is neither achievable nor required:

"THE COURT: [Predictive coding] certainly works better than most of the alternatives, if not all of the alternatives. So the idea is not to make this perfect, it's not going to be perfect. The idea is to make it significantly better than the alternative without nearly as much cost."²

In his response to a request for clarification filed by Plaintiffs' counsel on February 20, Judge Peck indicated on February 21 that he would be issuing a written opinion. The parties filed their Parties' Proposed Protocol Relating to the Production of Electronically Stored Information ("Parties' Protocol") on February 22. Judge Peck filed his written Opinion and Order ("Peck Opinion") addressing the parties' proposed protocol and the defensibility of the use of predictive coding on February 24.

In addition to addressing the specifics of the Parties' Protocol, Judge Peck provided clear guidance on the use of predictive coding in civil discovery. Importantly, his opinion "recognizes that computer-assisted review is an acceptable way to search for relevant ESI in appropriate cases."³ Judge Peck provided four related lessons for the future: (1) courts will not be able to "approve a party's proposal as to when review and production can stop until the computer-assisted review software has been trained and the results are quality control verified"; (2) phased discovery, with the most likely relevant sources being reviewed and produced first, is an

¹ Conor R. Crowley is the founder of Crowley Law Office, advising corporations and laws on electronic discovery, information governance and data privacy. He is the Vice Chair of The Sedona Conference Working Group on Best Practices for Electronic Document Retention and Production (WG1), and a member of The Sedona Conference Working Group on International Electronic Information Management, Discovery and Disclosure (WG6), the International Association of Privacy Professionals, the Advisory Board for Georgetown University Law Center's Advanced E-Discovery Institute and the Board of Advisors for BNA's Digital Discovery & e-Evidence. He is a Senior Editor of a number of The Sedona Conference's publications including The Sedona Conference Commentary on Legal Holds and The Sedona Principles (Second Edition): Best Practices Recommendations & Principles for Addressing Electronic Document Production. He can be reached at ccrowley@crowleylawoffice.com.

² Transcript of February 8, 2012 Hearing at 76:8-12.

³ Peck Opinion at 2.

appropriate way to control discovery costs; (3) cooperation, where the producing party shares information about the identities of key custodians and the search methodology to be employed, is likely to make courts more apt to agree to a proposed discovery protocol; and, (4) "[i]t also is important for the vendors and/or knowledgeable counsel to be able to explain complicated ediscovery concepts in ways that make it easily understandable to judges who may not be tech-savvy."⁴

Judge Peck also addressed the situation where the producing party wants to use predictive coding but the requesting party objects. He stated that the question to be asked in such a situation was what alternative approach the requesting party proposed. He went on to discuss the unreliable nature, and unjustified expense in large volume cases, of traditional linear manual review.⁵ Judge Peck also discussed problems with using keywords to find relevant documents noting that they are often picked with insufficient thought or discussion, frequently resulting in over-inclusive search results that contain large numbers of irrelevant documents, and thus have been shown to not be very effective.⁶ Judge Peck concludes that:

"Computer-assisted review appears to be better than the available alternatives, and thus should be used in appropriate cases. While this Court recognizes that computer-assisted review is not perfect, the Federal Rules of Civil Procedure do not require perfection."⁷

Objections to Judge Peck's February 8 rulings allowed Judge Peck to address three common objections to the use of predictive coding: (1) Federal Rule of Civil Procedure ("F.R.C.P.") 26(g)(1)(A) requires counsel to certify that their client's document production is "complete" and "correct" as of the time it was made and the use of predictive coding to identify relevant, non-privileged documents for production does not allow counsel to so certify;⁸ (2) the use of predictive coding to identify relevant documents implicates Federal Rule of Evidence ("F.R.E.") 702 and the *Daubert* standard;⁹ and, (3) there is no way to be certain at the beginning of the process that the results achieved by the use of predictive coding are accurate, or that the methodology employed is reliable.¹⁰

Judge Peck correctly noted that plaintiffs' reliance on F.R.C.P. 26(g)(1)(A) was erroneous because that rule applies to disclosure rather than discovery.¹¹ The applicable rule, F.R.C.P. 26(g)(1)(B) does not require certification that a discovery response is "complete" but instead relies on the proportionality considerations of F.R.C.P. 26(b)(2)(C) which incorporates a rule of

⁴ Id. at 23-25.

⁵ Id. at 18-19, citing 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), and 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 48.

⁶ Id. at 19-20, citing, *inter alia*, David L. Blair & M.E. Maron, An Evaluation of Retrieval Effectiveness for a Full-Text Document Retrieval System, 28 Comm. ACM 289 (1985).

⁷ Id. at 21, relying on Fed. R. Civ. P. 1, and Fed. R. Civ. P. 26(b)(2)(C).

⁸ Id. at 13.

⁹ Id. at 14-15.

¹⁰ Id. at 15-17.

¹¹ Id. at 13.

reasonableness.¹² With respect to F.R.E. 702 and the *Daubert* standard, Judge Peck noted that these apply to the admissibility of evidence at trial and are "not applicable to how documents are searched for and found in discovery."¹³ Finally, with respect to plaintiffs' objection as to accuracy and reliability, Judge Peck noted that these concerns were, at best, premature.¹⁴ An important factor in this decision was the fact that defendants had agreed to be totally transparent with respect to the methodology employed and with respect to both the specific documents that were found to be relevant and those that were found to be not relevant. Additionally, Judge Peck found that concerns about reliability were premature because reliability would be addressed, after a quality assessment had been performed, using a proportionality analysis and the Court lacked sufficient information at this time to conduct such an analysis.¹⁵ In addition to the factors discussed above, another important consideration in Judge Peck's decision, albeit not explicitly recognized as such, may have been the fact that the human review was conducted by senior attorneys rather than junior associates or paralegals.¹⁶

Judge Peck's ultimate conclusion provides the most explicit judicial support to date for the use of computer-assisted review:

"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. As with keywords or any other technological solution to ediscovery, counsel must design an appropriate process, including use of available technology, with appropriate quality control testing, to review and produce relevant ESI while adhering to Rule 1 and Rule 26(b)(2)(C) proportionality. Computer assisted review now can be considered judicially-approved for use in appropriate cases."¹⁷

Although this is a decision by one Magistrate Judge in a single courtroom in the Southern District of New York, and thus not binding on any other court, the esteem with which Judge Peck is held means that this opinion is likely to be very influential.

Ultimately, there are three key takeaways from Judge Peck's opinion: (1) predictive coding is a reasonable, and probably more effective, approach for identifying relevant ESI than currently available alternatives; (2) the appropriate technology must be combined with the right process to be defensible; and, (3) simplistic counter-arguments to the use of predictive, such as "black box", "not perfect" and "using predictive coding means meeting the *Daubert* standard", will no longer be persuasive.

¹² Id.

¹³ Id. at 15.

¹⁴ Id. at 16.

¹⁵ Id.

¹⁶ Id. at 10.

¹⁷ Id. at 25-26 (emphasis added).