Southern District of New York Addresses Predictive Coding During E-Discovery

February 23, 2012

Da Silva Moore v. Publicis Groupe, No. 11 Civ. 1279 (S.D.N.Y. Feb. 8, 2012).

On February 8, 2012, Magistrate Judge Andrew J. Peck of the U.S. District Court for the Southern District of New York heard the parties in *Da Silva Moore v. Publicis Groupe*, No. 11 Civ. 1279 (S.D.N.Y. Feb. 8, 2012) and ordered them to adopt a protocol for e-discovery that includes the use of predictive coding.

In the case, the plaintiff, Monique Da Silva Moore, filed a Title VII gender discrimination action against Publicis Groupe, an advertising conglomerate, on the behalf of herself and other women alleged to have suffered discriminatory job reassignments, demotions, and/or terminations. The case entered discovery and the plaintiff demanded documents relevant to determine if the defendant had violated Title VII. After the defendant produced almost 2.5 million documents, the plaintiff sought to use software that could predict relevant documents.

Generally, in order to take advantage of predictive coding during discovery, an attorney manually reviews a sample of all the documents produced during discovery. While reviewing the documents, the attorney classifies or tags certain documents as relevant to the case. The software then reviews the subset and trains itself to recognize characteristics of the document that make it relevant. Once the attorney has reviewed the entire subset and trained the software to recognize those documents, the software scans all the documents produced and pulls only ones it recognizes as relevant, based on the criteria it learned from the documents manually reviewed.

At the February 8 hearing, Judge Peck discussed the review of 2.5 million documents with the parties' attorneys and e-discovery consultants. The consultants provided guidance as to the proper subset size to yield adequate results. In doing so, they determined that the subset should consist of about 2,399 documents out of the 2.5 million. They found that by reviewing that many documents, it should yield a 95 percent confidence level and a 2 percent margin of error. The attorneys can expect to review between 15,000 to 20,000 documents in order to predicatively pull only the relevant documents from all the documents produced.

Judge Peck then instructed the parties to submit their draft discovery protocols by February 16th, and noted that he "may wind up issuing an opinion on some of what we did today." Thus, we can expect an opinion from Judge Peck addressing predictive coding very soon, which will be the first to address such an issue.