

Judge Questions Efficacy of Boolean Search and Invites Parties to Consider “More Innovative” Techniques

By: Thomas A. Lidbury & Nicholas S. Feltham

The Hon. Shira A. Scheindlin, of the United States District Court for the Southern District of New York, has ordered the federal government to try again to collect documents responsive to requests issued under the Freedom of Information Act (FOIA) in *Nat'l Day Laborer Org. Network v. U.S. Immigration and Customs Enforcement Agency*. In doing so, she also suggested that the government explore using more advanced technology than keyword searches.

Judge Scheindlin found that the federal government initially approached too few employees and allowed those who were approached to self-select ad hoc methods for search and review that were not up to “twenty-first century” standards. Faced with a record lacking in detail as to how various government employees searched their own electronic files, Judge Scheindlin held that “the government will not be able to establish the adequacy of its FOIA searches if it does not record and report the search terms that it used, how it combined them, and whether it searched the full text of documents.”

Several of the federal agencies relied on individual employees to search their own electronic files and to select their own methodology for doing so. In many cases, specifics of the chosen methods were not available. Judge Scheindlin finds it “impossible to evaluate the adequacy of an electronic search for records without knowing what search terms have been used.” Moreover, “it is not enough to know the search terms” but rather the “method in which they are combined and deployed is central to the inquiry” as well. Judge Scheindlin also concluded that “most custodians cannot be ‘trusted’ to run effective searches because designing legally sufficient electronic searches in the discovery of FOIA context is not part of their daily responsibilities.”

Judge Scheindlin goes further, pointing out potential weaknesses to the Boolean search methods that lawyers have considered standard practice for decades. She identified the risks of misspellings, poor selection of “logical connectives” – such as “and,” “or,” “w/10,” and the difficulty in anticipating all of the relevant words and phrases without

iterative searching and testing. Judge Scheindlin also quotes Judge John Facciola, who equated evaluating the adequacy of particular Boolean search methods with going “where angels fear to tread.” In light of these weaknesses to Boolean methods, Judge Scheindlin suggests that “beyond the use of keyword search, parties can (and frequently should) rely on latent semantic indexing, statistical probability models, and machine learning tools to find responsive documents.”

Judge Scheindlin stops short of ordering the government to use more advanced methods but she does order the parties to negotiate a more robust search protocol. And she encourages the parties to at least consider “predictive coding techniques and other more innovative ways to search.” Opinions such as this one, which paint Boolean search technology as an antiquated technology of a bygone century, suggest a possible future in which courts will affirmatively require the use of more advanced search technology.

Click [here](#) to view a copy of the district court’s Opinion and Order in *Nat’l Day Laborer Org. Network v. U. S. Immigration and Customs Enforcement Agency*, No. 10 Civ. 3488 (SAS), 2012 U.S. Dist. LEXIS 97863 (S.D.N.Y. July 13, 2012).

If you have any questions about this case, please contact the authors of the alert.

Thomas A. Lidbury
Partner & Leader of the Data
Management & Discovery Group
(312) 569-1356
Thomas.Lidbury@dbr.com

Nicholas S. Feltham
Commercial Litigation
Associate
(215) 988-2688
Nicholas.Feltham@dbr.com

For any other e-discovery questions, please contact team leader Thomas A. Lidbury (above) or other members of the firm’s [Data Management & Discovery group](#).