

# Robot Review

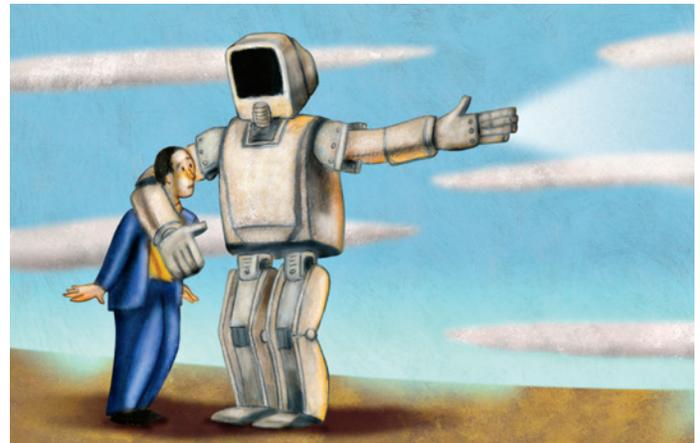
Will predictive coding win the trust of the courts? *By Farrah Pepper*

*Above the Law* declared predictive coding its “legal technology buzzword of 2011.” More recently, as covered on *EDD Update*, the buzz became a deafening roar when Recommind announced a patent for its predictive coding technology. Let’s take a dive into what “predictive coding” really means, explore some of its purported benefits, and ask the biggest question: Is it reasonable and therefore defensible to use?

We’ll start with what predictive coding is not. It is not “linear” document review, where human reviewers manually review and code documents that might have been subject to some simple applications of limiting and organizing factors, such as custodian, date ranges, and keywords. While linear review might offer value in matters with small amounts of electronically stored information—or other concerns that necessitate human eyes on every page—it can be unwieldy when facing huge volumes of data. Document review is widely believed to be the most costly component of electronic data discovery, because cases can now involve terabytes of data.

Predictive coding providers claim they can automate much of the review process, with human guidance. Documents can be prioritized into likely order of importance, typically based on a “learning set” of documents coded up front by a subject matter expert. That essentially creates a rebuttable presumption of relevance for other coding, that can be tested via sampling and revised if necessary. Then, the argument goes, the attorneys leading the case can dig into the substance a whole lot faster.

As discussed in “Crash or Soar” (*LTN*, Oct. 2010), perceived advantages include cost savings; efficiency; transparency (contrasting with linear review, which



typically provides only the conclusion, with no insight as to how a reviewer evaluated documents); replicability (because technology-assisted review can be applied almost identically over time while linear review is more variable); and enhanced confidentiality (fewer people see documents).

While the supposed benefits sound enticing, is it defensible? The touchstone of the developing EDD case law standards is reasonableness, not perfection. In

*Pension Comm. of the Univ. of Montreal Pension Plan v. Banc of Am. Sec., LLC*, 685 F. Supp. 2d 456, 461 (S.D.N.Y. 2010), U.S. District Court Judge Shira Scheindlin declared that “courts

cannot and do not expect that any party can meet a standard of perfection.”

Still, we are left to wonder if predictive coding meets the standard of “reasonableness” (including the certification of a “reasonable” inquiry to respond to discovery requests that attorneys make pursuant to Federal Rule of Civil Procedure 26(g)(1)(A),

*E-Discovery Showcase*

and the need for “reasonable” steps to prevent disclosure of privileged information to prevent waiver under Federal Rule of Evidence 502(b)).

But what if the computer gets it wrong? How do counsel justify the reasonableness of relying on an algorithm rather than the judgment of an experienced legal professional? Some commentators suggest the question be flipped to ask how counsel rely on the human brain, prone to daydreaming, and which, according to at least one study, gets the coding wrong as much as, if not more often than, the machines do. A recent law review article by Maura Grossman and Gordon Cormak found that technology-assisted searching can “yield better overall results, and better precision” than manual review. *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, 17 Rich. J.L. & Tech. 11 (2011). Such studies support the argument that relying on an algorithm, rather than tired human eyes clicking through documents, may make predictive coding defensible, particularly when applied to first pass review. But courts have been notably silent on the reasonableness question — or anything at all about predictive coding. As it is common for law to lag behind technology, perhaps such silence should not be surprising. The case of *Victor Stanley, Inc. v. Creative Pipe, Inc.*, 250 F.R.D. 251, 259 (D. Md. 2008) before U.S. Chief Magistrate Judge Paul Grimm, comes close. Although *Victor Stanley* does not mention predictive coding or machine learning search techniques, Grimm says “there is room for optimism that as search and information retrieval methodologies are studied and tested, this will result in identifying those that are most effective and least expensive to employ for a variety of ESI discovery tasks.” And in a recent law review article he said: “It is hoped that future courts will be receptive and accommodating to the use of these screening methods to prevent disclosure of privileged and protected information. While these methods are not perfect, there is growing evidence that they are as good, or far better than, ‘eyes on’ review of all digital information by an attorney or paralegal. There is every reason to believe that computer-based screening methods’ recall (completeness) and precision (accuracy) rates will continue to improve.” Paul Grimm, Lisa Yurwit Bergstrom,

and Matthew Kraeuter, “*Federal Rule of Evidence 502: Has It Lived Up to Its Potential?*” 17 Rich. J.L. & Tech. vol. XVII, issue 3 (2011).

While courts have yet to embrace predictive coding and other alternative search methodology in a widespread fashion, it may be just a matter of time before such methodology is as routinely discussed as contract attorneys and keywords are now. If the academic literature is any indication, technology-assisted review may be an idea whose time has come — and it is just a matter of lawyers and the courts catching up. The most fool-proof prediction about predictive coding is likely this: the buzz is just getting started.

---

**Farrah Pepper** is of counsel at Gibson, Dunn & Crutcher, based in New York. She is vice-chair of its electronic discovery and information law practice group, and a coauthor of ALM’s *EDDUpdate* blog. E-mail: [FPepper@gibsondunn.com](mailto:FPepper@gibsondunn.com).