

Podo Legal's "Discovery Spotlight"

02/24/11: To Predict Or To Prioritize? Predictive coding and prioritization technology look like the next big things in document review technology. Predictive coding refers to automated processes that are applied to a dataset to make decisions about documents, whether about responsiveness or about relevance to key case issues. Prioritization technology employs trained relevance ratings to generate numeric scores for documents, with the review order flowing from the highest-scoring documents to the lowest.

Neither approach is brand spanking new. (Nor is prioritization new to search technology in general; you may have heard of a little company called Google.) Both have been available as add-on features from several litigation technology providers for a few years now. Recently, however, there seems to be a renewed push to market predictive coding and prioritization technology.

Given the ever-increasing size of datasets to be searched and reviewed, it's easy to see the appeal. The main strength of predictive coding is to save human review time (and thus money) by automating the initial, and perhaps only, determinations on documents. Prioritization facilitates the early identification of the most important documents and issues; and enables the review of these documents by higher-level (and higher-cost) reviewers, leaving less crucial documents to be reviewed at lower cost.

Defensibility concerns linger. Is delving the depths of datasets digitally as good as lifting the proverbial lid off the good old bankers box? Do we have enough of the right people to train the predictive coding and prioritization systems with keywords? How receptive are members of the judiciary? What will opposing counsel say in a discovery spat?

We're not *telling* you; we're *asking* you. Please share any experience you've had - good, bad or ugly - with predictive coding or prioritization technology on a real live document review. We're all ears.