Predictive Coding is a Disruptive Innovation that will Change How Law is Practiced

Author: Bill Henderson
Source: Law Professors.typepad
Link: December 5, 2012

I have been reading about predictive coding for a few months now, and that is my conclusion. Predictive coding is the use of computer algorithms and machine learning to conduct the review of electronically stored information (ESI). For a useful primer, see Frederick Kopec, Predictive Coding in eDiscovery or Predictive Coding for Dummies (remarkably, there are two editions, one by Symantec and the other by Recommind, see Legal Tech Insider, A Tale of Two Predictive Coding Books).

From the client perspective, predictive coding is at least as good as first-level human review (typically junior attorneys screening for relevance and privilege) but dramatically less expensive. And note, whatever efficiency and accuracy benefits predictive coding has today, it will only improve in the months and years to come. It contrast, our processing capacity as humans is, well, static.

The big players in the space are Kroll Ontrack and Recommind. These are not insignificant companies. Kroll Ontrack started as a hard disk recovery service and evolved into the e-discovery and information management services. It now employs 1,500 workers in eleven U.S. and nineteen foreign locations around the world. In 2010, Kroll Ontrack had revenues of $250 million. A few layers up, it is owned by the Private Equity giant Providence Equity Partners. Recommind has approximately $15 million in annual revenues and approximately 100 employees spread over facilities in Massachusetts, California, London, Germany, and Australia. According to this June 2012 story at the CIO Agenda at Computer Business Review, Recommind is gearing up to go public.

Howard Sklar, Senior Corporate Counsel for Recommind, just posted an essay entitled, Legal Acceptance of Predictive Coding: A Journey in Three Parts. The parts are: (1) acceptance that predictive coding reasonable, (2) arguments that it is better and thus must be used in this case, (3) sua sponte judicial order that it be used. The fourth part, still to come argues Sklar, is a state bar ethics watchdog issuing a ruling that failure to use predictive coding is unethical.

Here is an excerpt from Sklar's post:

There’s a certain trajectory for technology adoption. Early adopters, mainstream acceptance, laggards. But, slow or fast, adoption occurs. The law is the same way, in its own fashion. But the legal acceptance of predictive coding has had a path that’s unorthodox. From the legal perspective, predictive coding has gone through three cycles, not entirely as expected.

In cycle one, companies began using predictive coding. The efficiencies are compelling. Better end results in less time at a cost savings. An ability to better find and understand the facts embedded—sometimes hidden—in your documents. These things are crucial in today’s corporate world. Law firms were slower, but generally followed their clients into predictive coding, and soon saw the benefits first hand.

Other vendors—usually the first to adopt new technology—were laggards. They fought the adoption of predictive coding as long as they could, mainly because they didn’t have the capability to do it themselves. Eighteen months ago, the most frequent question I would get at conferences was “has there been a court case approving the use of predictive coding?” In the “ridicule it and it will go away” marketing approach, they were hoping to scare corporations and law firms away from the benefits corporations could achieve.

Then came Da Silva Moore and Global Aerospace [which, against the objections of one of the litigants, ruled that predictive coding was a judicially reasonable method of conducting discovery.] ...
During this period, other vendors stopped criticizing predictive coding and started marketing it—sometimes with the capability, sometimes without. ...

After waiting for the first decision approving the use of predictive coding, we went to stage two faster than anyone had thought possible: not whether you can use predictive coding, but whether you must use it. This was the argument in the Kleen Products case. The defendants had completed their review, and the plaintiffs' argued that the review was defective because predictive coding wasn't used. Eventually, the parties cooperated to end that dispute, but the argument had been made. ...

We're now in stage three: a court has sua sponte ordered the use predictive coding. And not just any court, the Delaware Chancery Court, one of the most important corporate courts in the nation.

In the future, we'll enter stage four: the decision by a state bar's ethics watchdog that failure to use predictive coding is ethically questionable, if not unethical. After all, purposefully using a less-efficient, less accurate, more expensive option is problematic. I think that's probably 18 months away. But given how fast we've gone through the first three states, stage four may come next week.