Continuous Active Learning (CAL) Proves Superior, Again

By John Burchfield, DSi
September 23, 2014

A few months ago Gordon V. Cormack and Maura R. Grossman published a study, "Evaluation of Machine-Learning Protocols for Technology-Assisted Review in Electronic Discovery," which found that continuous active learning (CAL) is the best Technology Assisted Review (TAR) method. We wrote about the benefits of CAL and the findings of the study in July.

TAR 2.0: Continuous Active Learning (CAL)

TAR 2.0 involve two interactive tools: a **KEYWORD SEARCH SYSTEM** and a **LEARNING ALGORITHM**. Because the technology *continues to learn* with all added documents, TAR 2.0 has many benefits over older TAR platforms.

▶ Works great with rolling data
▶ Subject-matter experts are only needed for quality control
▶ More efficiently reduces costs
▶ Potential for even greater document reduction as review continues

So, why are we talking about CAL again?

It's not just because we think the study proved excellent points about TAR (which we do) and it's not just because we respect Cormack and Grossman (which we also do). The study has resurfaced in eDiscovery news with a recent guest post by Cormack and Grossman on Ralph Losey's e-Discovery Team blog. Cormack and Grossman wrote the post in response to some disparaging (and inaccurate) criticisms of their study that have come from other vendors in recent weeks. Their post is a thorough response to these criticisms.

The value of the Cormack-Grossman paper is that it includes significant research and was peer-reviewed before publication. In our opinion, this makes it an unbiased account of the best methods for TAR.

In their blog post, Cormack and Grossman compare the controlled comparison of TAR methods to EPA fuel economy ratings. It's impossible to predict with certainty the fuel consumption of a specific trip, but you would be hard pressed to find a trip where a Toyota Prius consumed more fuel than a Hummer H1. The blog post says:

Manufacturers generally do not like controlled comparisons, because there are so few winners and so many also-rans. So it is with automobiles, and so it is with eDiscovery software. On the other hand, controlled comparisons help consumers and the courts to determine which TAR tools are reliable.

Just as consumers deserve to know the average gas mileage for a car they are buying, lawyers and corporations deserve to know the TAR tools that determine the best recall. Cormack and Grossman's research identified more than 100 instances where CAL outperformed other TAR methods. That's solid evidence.

We encourage you to read their whole post here. CAL is the way of the future of TAR, and we're excited to be at the forefront of it.

DSicovery ®

eDiscovery About People™

About DSi
Serving law firms and corporate legal departments worldwide, DSi is a litigation support services company that provides advanced eDiscovery and digital forensics services. Through five core business processes—DSicollect, DSintake, DSinsight, DSireview, DSisupport—DSi's highly trained staff will help you harness today's most forward technology to gain a competitive advantage. DSi is headquartered in Nashville, Tenn. with offices in Knoxville, Tenn., Cincinnati, Ohio, Charlotte, N.C., Minneapolis, Minn.,
Philadelphia, Penn., Atlanta, Ga. and Washington D.C. For more information, please visit DSi at www.dsicovery.com or follow us on Twitter at: @DSicovery.