



THE ROUND TABLE W/AVALON

by J.P. Midgley, President



Since the dawn of the e-discovery age in litigation, all of the articles, blogs, seminars and the like have focused on how to deal with the massive amount of data that is now critical to discovery in litigation. We are constantly hit with dozens - if not hundreds - of ads about how to keep e-discovery costs down.

Somewhere along the way the parties have realized that the biggest cost isn't the vendor costs for collecting and processing the data - it's the attorney review time associated with massive amounts of electronic emails and documents.

The real key to reducing e-discovery costs in today's litigious world is getting your attorneys access to the most relevant and pertinent information as quickly as possible while weeding out the "junk" data.

Most people are familiar with the most common type of data filtering such as date, file type and keyword filtering, and de-duplication of the data. In many cases these filtering techniques, along with a very well thought out and analyzed keywords list, are very effective. However, there are some newer tools in the market that can help drill deeper through this data to get to the heart of the matter.

Software vendors have developed what they are calling "early case assessment tools". The domain filter on these tools proves to be extremely valuable for a few reasons. First, it can immediately identify potentially privileged documents by comparing email domains to a registered list of over 40,000 law firms. Additionally, it can help identify SPAM email domains and remove them from your data set.

What you will likely start seeing a lot of is a new technology called "predictive coding". This is a far more advanced process that actually learns from your decisions. Given a set of a million documents, this software asks for no keywords. Instead, it feeds you batches of 40 documents at

a time and simply asks if the document is relevant or not. Once you have completed the 40 document batch, it analyzes the results and feeds you the next 40 documents to learn more about how you are making decisions. After about 1500 total documents have been reviewed, it can then make a decision of relevancy on the rest of your documents.

Technology continues to evolve and provide new tools that bring efficiencies to the area of data filtering, review and production. It's out there and changing fast, and it's our job to keep up.