METROPOLITAN
CORPORATE
COUNSEL®

MARCH 2017 WWW.METROCORPCOUNSEL.COM

7 Things You Need to Know About Technology-Assisted Investigations

The right approach can make a big difference in your success

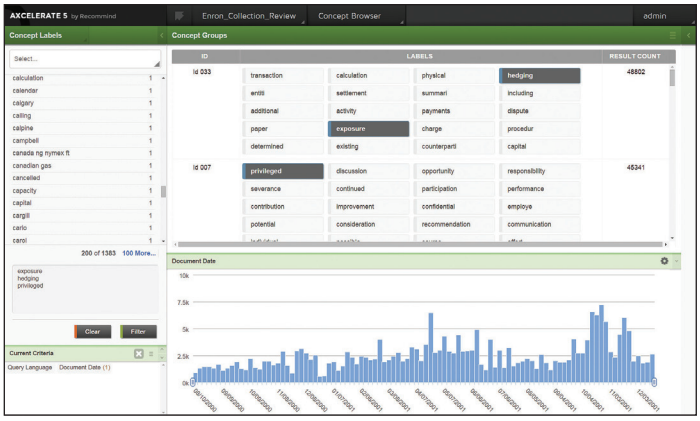
By Hal Marcus / OpenText Discovery

At Legaltech 2017 in New York City, OpenText's Adam Kuhn moderated Technology-Assisted Investigations for the Enterprise, a one-hour panel session with legal and compliance experts. They were John Davis from UBS, Laura Roman from the New York Stock Exchange, David Yerich of the UnitedHealth Group and Stephen Medlock of Mayer Brown LLP.

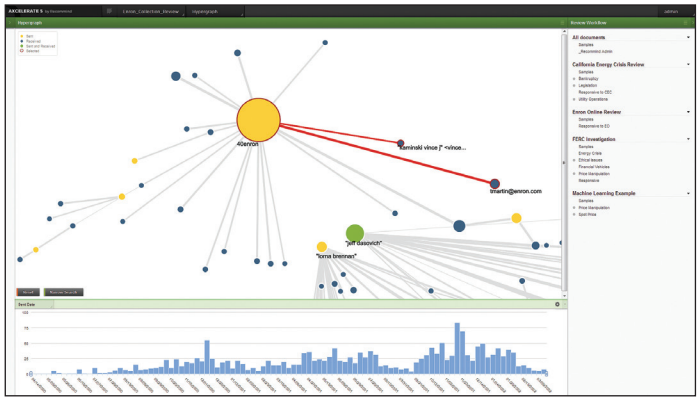
The session offered lessons about e-discovery technology and fact-finding strategies that yield the best results for corporate investigations. What follows are seven takeaways based on the panelists' collective insights, which they shared with the understanding that their comments are not attributable to them or their companies.

METROPOLITAN CORPORATE COUNSEL®

OPENTEXT™



Concept groups help uncover code words and obfuscations; histograms display activity levels over time.



Hypergraph communication maps visualize connections between individual communicators and domains.

1. You can't always tell the people you're investigating that you're investigating them.

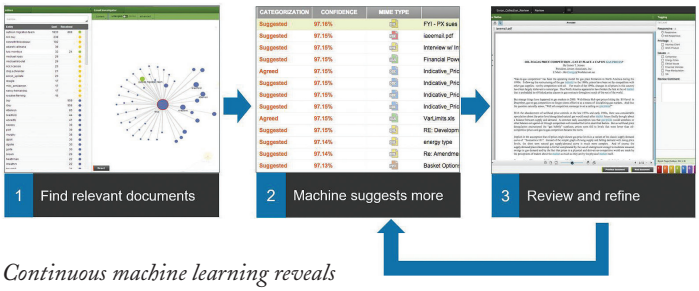
In some circles it's known as the Hawthorne Effect: The act of observing people changes their behavior. So, unlike litigation discovery, an internal investigation may need to be clandestine. Not being able to interact with the custodians of key data sources challenges your ability to collect data thoroughly and do initial fact-finding via interviews. Instead you have to use all the tools at your disposal to track down the facts you need to assess your organization's exposure.

2. Start with the knowns, move on to the known unknowns – then maybe you'll find the unknown unknowns.

Every investigation begins with what you think you know: time frames, geographies, players and allegations. Start with searches and filters related to those points and identify relevant content, then cast a wider net. Look for the facts you know you need to find; try to fill in those blanks. If you do your job well, the things you haven't even thought of (answers to questions you didn't know to ask) will make themselves apparent as you go.

3. Bad actors' code words will confound keyword search, but when you find them through analytics, they're key to your investigation.

No one ever used the word "bribe" in an email; if they did, keyword search would be all you need. Conceptual analysis looks beyond the literal, grouping documents by the weighted, statistical co-occurrence of their terms in totality. As a result, concept groups can help uncover code words and obfuscations in communications. Finding these terms not only enables you to broaden your efforts, it unveils precisely the kind of evasive behavior that should draw an investigator's attention. [See screen capture left top.]



Continuous machine learning reveals more relevant documents conceptually similar to the ones you've already found.

4. Look for anomalies: sudden drop-offs, people avoiding each other, email sent to personal addresses.

Liability-causing activity will usually lead to changes in behavior. When players stop communicating, stop showing up or take an email exchange off the corporate domain to personal Gmail accounts, these are indicators that you're on the right track. Histograms and communication maps can make these behaviors easy to spot – displaying activity levels, email volumes, unexpected connections, missing (possibly deleted) messages and changes in the email domains used. [See screen capture left bottom.]

5. When analyzing communications between custodians, don't forget the phone logs.

This point highlights that in an investigation, structured and semistructured data (such as phone logs) can play a key parallel role to unstructured data (emails, chats, slides and Word documents). Visualizing both can be highly instructive. It's also a good reminder that communication takes many forms, some of them more "old school" and analog than others. The lack of any emails demonstrating liability does not necessarily ensure that your organization's risk is contained.

6. Investigations are iterative, not linear – you have to make multiple passes on the data.

Unlike litigation discovery, which is at least partially about going document by document and checking boxes to meet external obligations, internal investigations are self-driven, with little structure. You need to test your initial theories against the data; revise those theories accordingly; and then test them again. This necessitates revisiting the same universe of data repeatedly guided by the new facts you've uncovered. Search and analytics tools that facilitate this kind of iterative analysis are your best allies.

7. Predictive coding has to be flexible: no 'stabilization,' no formal training.

Predictive coding technology has come a long way. Once seen as a rigid classification system for multimillion-document litigation projects, it has emerged as an indispensable tool for internal investigations, M&A due diligence, data security breach response and more. What makes the difference is *continuous machine learning*, in which the system never "stabilizes" or terminates its training. Instead, it keeps learning as you learn, refining its document models to ever more accurately reveal relevant evidence as your investigation progresses. [See diagram at top.]

Bonus lesson:

If you draft an exemplar document for machine learning, be sure that you don't produce it.

Yes, it drew an audience laugh, but the kind of nervous laugh that reminds us that such things do happen. To guide the predictive coding engine, you may wish to concoct a few documents resembling the kind of evidence that you're seeking. While this technique can be effective, your discovery platform should contain safeguards to keep you from inadvertently producing any such fake documents to a regulator. Because that would be *bad*. (Important safety tip; thanks, Egon!)

Hal Marcus *is an e-Discovery attorney and the director of product marketing for OpenText Discovery (formerly Recommind). In that role, he educates corporate and law firm counsel on technology strategies for litigation, investigations, compliance and information governance. He can be reached at hms@opentext.com*