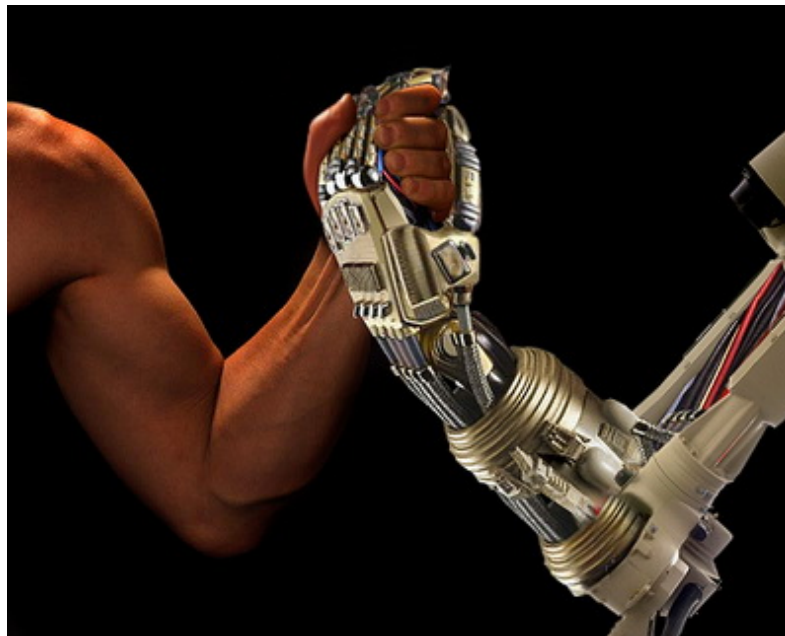




Man vs. Machine: The Promise/Challenge of Predictive Coding & Other Disruptive Technologies

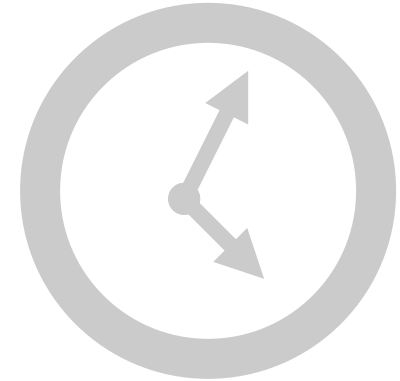


Speakers

- **Judge Andrew Peck**, U.S. Magistrate Judge for the Southern District of New York -- **LTN Award Winner (Champion of Technology)*
- **Ralph Losey**, Partner and National e-Discovery Counsel for Jackson Lewis, LLP
- **Maura Grossman**, Litigation Counsel at Wachtell, Lipton, Rosen & Katz
- **Dean Gonsowski**, eDiscovery Counsel at Symantec

Today's Agenda

- 1 The Human Review Process as the “Gold Standard”
- 2 Defining Technology Assisted Review (TAR)
- 3 How Defensible is TAR?
- 4 Are Certain Use Cases Better Suited for TAR?
- 5 Will the Role of Attorneys, Judges and Litigants Change?
- 6 The Recipe for Success



Human Review: Gold or Lead?

- TREC and Blair/Marion tell us that the human review process is “absurdly ineffective”
- “Lawyers are even worse than non-lawyers in making relevancy reviews.”
- “65% recall at 65% precision is the best retrieval effectiveness achievable given the inherent uncertainty in human judgments of relevance”
- **How do we get comfortable with the other 35%?**



“Technology Assisted Review” Defined

- TAR is an umbrella concept that involves keyword search, conceptual search, clustering, relevance ranking, sampling and predictive (aka computer-assisted) coding
- “By computer-assisted coding, I mean tools (different vendors use different names) that use sophisticated algorithms to enable the computer to determine relevance, based on interaction with (i.e., training by) a human reviewer.
- ...Unlike manual review, where the review is done by the most junior staff, computer-assisted coding involves a senior partner (or team) who review and code a ‘seed set’ of documents. The computer identifies properties of those documents that it uses to code other documents. As the senior reviewer continues to code more sample documents, the computer predicts the reviewer’s coding.”

Defensibility -- What is the Judicial Standard?

- Is it really Man v. Machine?
- Maybe the question should be: “are either methods acceptable?”
- “Most computer assisted search is already better than average keyword or manual search, so it should be accepted as reasonable under the law without confidence inflation.”



Use Case Distinctions

- “It is *way* too risky to turn it *all* over to the machines. They are not that good! The reports of their excellence have been grossly over-stated.”
- “All that counts in litigation are the *hot* documents, the highly relevant ones with strong probative value, not the documents which are just relevant, not to mention just responsive.”



Even if they were Equal...

- How can you rationalize spending more for a similar (flawed) output?
- Where does the role of proportionality and cost come into play?



What's the Recipe for Success?

1. Bottom line driven **proportional review**, where the projected costs of review are estimated at the beginning
2. High quality, **tech assisted review tools**, with predictive coding type software, and multiple expert review of key seed-set training documents using both subject matter experts and AI experts
3. Direct supervision and **feedback by counsel** - affirming under 26(g)
4. **Quality control**, including training, sampling, positive feedback loops, batching, etc.
5. Experienced and **motivated reviewers** who aren't afraid of the technology
6. New **methods to keep the reviewers engaged** and motivated
7. Project managers who know the tools and can help coach the team
8. **Strategic cooperation** between opposing counsel
9. Final **review of the production sets** before they go out the door

Is the Student Ready?

“When the student is ready, the master appears.”

-Buddhist Proverb



For more information...

e-discovery 2.0 thoughts about the evolution of e-discovery

Losing Weight, Developing an Information Governance Plan, and Other New Year's Resolutions

BY DEAN GONSOWSKI ON JANUARY 17TH, 2012



It's already a few weeks into the new year and it's easy to spot the big lines at the gym, folks working on fad diets and many swearing off any number of vices. Sadly perhaps, most popular resolutions don't even really change year after year. In the corporate world, though,

it's not good enough to simply recycle resolutions every year since there's a lot more at stake, often with employee's bonuses and jobs hanging in the balance.

It's not too late to make information governance part of the corporate 2012 resolution list. The reason is pretty simple – most companies need to get out of the reactive firefighting of eDiscovery given the risks of sloppy work, inadvertent productions and looming sanctions. Yet, so many are caught up in the fog of eDiscovery war that they've failed to see the nexus between the upstream, proactive good data management hygiene and the downstream eDiscovery chaos.

In many cases the root cause is the disconnect between differing functional groups (Legal, IT, Information Security, Records Management, etc.). This is where the emerging umbrella concept of Information Governance comes to play, serving as a way to tackle these information risks along a unified front. Gartner defines information governance as the:

"specification of decision rights, and an accountability framework to encourage desirable behavior in the valuation, creation, storage, use, archiving and deletion of information, ... [including] the processes, roles, standards, and metrics that ensure the effective and efficient use of information to enable an organization to achieve its goals."

Perhaps more simply put, what were once a number of distinct disciplines—records management, data privacy, information security and eDiscovery—were rapidly coming together in ways that are important to those concerned with mitigating and managing information risk. This new information governance landscape is comprised of a number of formerly discrete categories:

- **Regulatory Risks** – Whether an organization is in a heavily regulated vertical or not, there are a host of regulations that an organization must navigate to successfully stay in compliance. In the United States these include a range of disparate regimes, including the Sarbanes-Oxley Act, HIPPA, the

- Information Governance Plan, and Other New Year's Resolutions
- > Information Governance Gets Presidential Attention: Banking Bailout Cost \$4.76 Trillion, Technology Revamp Approaches \$240 Billion
- > Lessons Learned for 2012: Spotlighting the Top eDiscovery Cases from 2011
- > Q&A with William P. Butterfield on his Testimony Regarding the Costs and Burdens of eDiscovery Before the House Judiciary Committee's Subcommittee on the Constitution
- > New Utah Rule 26: A Blueprint for Proportionality in eDiscovery
- > Q&A with The Sedona Conference's John Rabej on Chief Justice Roberts, Proposed FRCP Amendments, and Congress' Interest in eDiscovery
- > Backup Tapes and Archives Bursting at the Seams? The Seven Year Itch Has Technology to Answer the Scratch

TAGS
analysis archive Clearwell collection costs data retention defensibility discoverable discovery

to address preservation and sanctions issues is necessary?

Yes
 No
 Not Sure

Vote

View Results

EMAIL US
Have a question?
Submit it

- BLOG ROL**
- > Archiving
 - > Clearwell
 - > Death By
 - > Dennis K
 - > E-Discover
 - > E-Discovery team
 - > add blog online
 - > ED RM
 - > Electronic Discovery Blog
 - > Enterprise Vault eDiscovery Blog
 - > In Re Discovery
 - > John Battelle's Search Blog

VISIT OUR BLOG

clearwellsystems.com/e-discovery-blog

to integrated information governance. And yet, this resolution doesn't need to encompass every possible element of information governance. Instead, it's best to put foundational pieces into place and then build the rest of the infrastructure in methodical and modular fashion.

Tags: analysis, archive, archiving software, budget, cloud, cloud computing, collection, costs, custodians, data retention, data sources, defensibility, discoverable, duty to preserve, e-discovery, e-discovery in the cloud, e-discovery software, EDD, ediscovery, ediscovery costs, ediscovery in the cloud, ediscovery workflow, EDRM, electronic data discovery, electronic discovery, Electronically Stored Information, email, email archive

Share Tweet Email Share Share Digg Like +1

Social Sharing