The Case for Technology Assisted Review and Statistical Sampling in Discovery

Position Paper for DESI VI Workshop, June 8, 2015, ICAIL Conference, San Diego, CA

Christopher H Paskach
The Claro Group, LLC
350 South Grand Ave., Suite 2350
Los Angeles, CA
+1 (213) 228-5080
cpaskach@theclarogroup.com

F. Eli Nelson
McKenna Long & Aldridge, LLP
1900 K Street NW
Washington, DC 20006
+1 (202) 496-7303
enelson@mckennalong.com

Matthew Schwab
The Claro Group, LLC
321 N Clark Street, Suite 1200
Chicago, Illinois 60654
+1 (312) 546-3422
MSchwab@theclarogroup.com

Abstract:

Technology Assisted Review (TAR) and Statistical Sampling can significantly reduce risk and improve productivity in eDiscovery processes. TAR reduces risk by mathematically ranking documents in descending order of likely relevance. Productivity is increased by prioritizing review on the most relevant documents, while review of low ranked records may even be avoided entirely.

Relevance ranking is an ordinal measure, meaning higher scores are more likely to be relevant, but a high score is not a guarantee that a document is relevant, and a low score is no guarantee that a document is not relevant. Statistical sampling quantifies the degree that relevance ranking does not reflect true relevance.   With such measures, attorney stakeholders can make informed decisions about the reliability and accuracy of the review process, thus quantifying actual risk of error and using that measurement to maximize the value of expensive manual review. Law firms that adopt these techniques are demonstrably faster, more informed and productive than firms who rely solely on attorney reviewers who eschew TAR or statistical sampling.

Introduction

For the past decade companies from manufacturers to financial services have turned to data analytics to stay competitive. An article in Forbes states: "*Analytics and the insights they deliver are changing competitive dynamics daily by delivering greater acuity and focus.*"[1] At the same time, the volumes of unstructured data created and stored by these companies such as emails, web pages, text messages and user files has grown so rapidly that when litigation arises or an investigation occurs, the volume of data needing review quickly overwhelms traditional attorney review processes.

Historically, discovery was largely paper-based. Processes designed to deal with modest volumes of paper were repurposed to handle Electronically Stored Information (ESI). As data volumes have exploded,[2] companies have focused on reducing unit costs without revisiting assumptions about the underlying process.   The result is that companies that routinely approved bills for document reviews have demanded more efficient approaches to reduce the cost of "eDiscovery." The initial response has largely been to use lower-cost lawyers, some of them off-shore, to reduce the largest component of

---

[1] Louis Columbus, Forbes, June 24, 2014, Roundup Of Analytics, Big Data & Business Intelligence Forecasts And Market Estimates, 2014.

[2] Jason R. Baron, Law in the Age of Exabytes: Some Further Thoughts on 'Information Inflation' and Current Issues in E-Discovery Search, XVII RICH. J.L. & TECH. 9 (2011), http://jolt.richmond.edu/v17i3/article9.pdf.

eDiscovery expense – document review costs.[3] Incremental reduction of hourly rates without changing the review process ignores the disruptive effect of analytics-based review. Companies now understand how technology and advanced analytics can be leveraged to dramatically lower the effort, error rate, and cost of analyzing massive volumes of data using fundamentally different approaches.

Brief history of TAR in Discovery

Before the early 2000's companies and their counsel typically ignored ESI in document production requests, and only produced emails to the extent they had already been printed to paper. Innovative litigators were known to print emails, "Bates Stamp" them, scan them into image file formats, and then use OCR programs to extract text that was already available in the original electronic systems to begin with.

Today we would laugh at such processes, but in a profession based on "*stare decisis*," changing processes takes time. New technologies that could reliably load emails and attachments into a discovery review platform without first converting them to paper took years to supplant the outdated and wasteful process of converting ESI to paper or scanned images for production.

Business problems stemming from information overload have driven rapid innovation in how ESI is managed. Basic Boolean searches to locate relevant email records were once the apex of sophistication. Today, fast and highly automated text classification using sophisticated machine learning algorithms have liberated users from junk email through "spam" filtering. Classification tools have continued to evolve, and in an interesting reversal of purpose, now help identify *relevant* documents with increasing speed and consistency, rather than simply trimming irrelevant content.

Next generation text classification techniques have been applied to eDiscovery review with impressive results; studies have concluded that machine learning techniques can outperform manual document review by lawyers.[4] Investment in TAR can dramatically increase the productivity, accuracy, and cost-efficiency of legal knowledge workers. Not surprisingly, these improved outcomes are only achievable when the legal teams managing the effort change their processes to use TAR effectively.

Changing a legal process such as discovery review requires sufficient evidence that the change will actually improve overall outcomes, not simply reduce cost or turnaround time at the expense of quality. Measuring improvement has been a challenge. The National Institute of Standards and Technology (NIST) Text Retrieval Conference (TREC) Legal Track[5] studied the feasibility of using ranked relevancy text classification approaches in a simulated legal document review. The results were encouraging, but the superiority of machine learning approaches over human review was not well

---

[3] NN Pace and L Zakaras, Where the Money Goes - Understanding Litigant Expenditures for Producing Electronic Discovery; RAND Institute for Civil Justice, 2012.

[4] Maura R. Grossman & Gordon V. Cormack, Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review, XVII RICH. J.L. & TECH. 11 (2011), http://jolt.richmond.edu/v17i3/article11.pdf.

[5] http://trec-legal.umiacs.umd.edu/

established until recently[6]. Now, however, the scholarship is so convincing that a rash of recent legal opinions [beginning with Da Silva Moore] has placed a judicial imprimatur on the reasonableness of using TAR.

Application of Statistical Sampling to eDiscovery

While the use of statistical sampling in eDiscovery has not become a standard practice, it is gaining wider use because it can cost-effectively test whether the discovery process has been reasonably effective in finding relevant and responsive documents.

A January 2012 white paper, The Case for Statistical Sampling in e-Discovery[7], discussed how statistical sampling could be applied to the eDiscovery process to provide reasonable assurance that the process achieved an acceptable level of quality. The next month, in Da Silva Moore vs. Publicis Groupe[8], US Magistrate Judge Andrew Peck issued an opinion and order on predictive coding that discussed the use of both judgmental and statistical sampling to validate the results of predictive coding.

Later that year, in Kleen Products vs. Packaging Corporation of America[9], US Magistrate Judge Nan Nolan heard testimony on the sampling process used by defendants to test the results of using search terms rather than predictive coding to find relevant documents before ordering the parties to meet and confer and cooperate to find an approach acceptable to both. They eventually reached an agreement to rely on the keyword searches validated by sampling that had already been done rather than redo the work by applying predictive coding to retrieve documents for production.

More recently, in Global Aerospace v. Landow Aviation the court stated that predictive coding (aka TAR) including a statistically sound validation protocol was a reasonable means for locating and retrieving documents for production.[10]

Cost impact of TAR

There are not yet many published studies comparing the cost of unassisted document review vs. TAR validated by statistical sampling. We have, therefore, modeled a comparison of the cost advantage of a law firm using TAR for review, to the cost of traditional linear review with contract attorneys. The model below shows that the TAR approach would save approximately three-fourths of the time and thirty percent of the cost.

The model and assumptions regarding the rate of manual review are premised strictly on relevance classification, not a full production review. Variability in relevance densities, issue complexity, and

---

[6] Grossman & Cormack, Technology-Assisted Review …., http://jolt.richmond.edu/v17i3/article11.pdf.
[7] **CH Paskach, P Strauss and MJ Carter,** KPMG LLP, **The Case for Statistical Sampling in e-Discovery,. http://advisory.kpmg.us/content/dam/kpmg-advisory/PDFs/RiskConsulting/case-for-statistical-sampling-e-discovery.pdf**
[8] Da Silva Moore v. Publicis Group, No. 11-CV-1279, 2012 U.S. Dist. LEXIS 23350 (S.D.N.Y. Feb. 24, 2012
[9] Case: 1:10-cv-05711 Document #: 412 Filed: 09/28/12 Page 10 of 53 PageID #:10404
[10] Global Aerospace, Inc. v. Landow Aviation, L.P., No. CL 61040 (Vir. Cir. Ct. Apr. 23, 2012)

collateral obligations, such as privilege, present significant confounding variables that cannot be incorporated faithfully into a pure TAR evaluation. This is particularly true for work flows that call for full review of documents that score above a particular relevance threshold.

|	Linear	TAR	Difference	
Hours Comparison				
All Document Review After Keyword Filtering				
First Level Review (Contract)[1]	1,200	0	(1,200)	
First Level Review (Staff)	180	0	(180)	
TAR Training (Associate)[2]	0	26	26	
TAR Testing (Staff)[3]	0	93	93	
Statistical Analysis	0	70	70	
Total All Document Review	1,380	189	(1,191)	
Relevant Document Review				
Second Level Review (Associate)	240	234	(6)	
Total Hours	1,620	423	(1,197)	
Hours Summary				
Unassisted Contract	1,200	0	(1,200)	
Law Firm Staff	180	93	(87)	
Law Firm Associate	240	260	20	
Statistical Analysis	0	70	70	
Total Hours	1,620	423	(1,197)	-74%
Costs Summary				
Unassisted Contract	$ 78,000	$ -	$ (78,000)	
Law Firm Staff	31,500	16,275	$ (15,225)	
Law Firm Associate	84,000	91,000	$ 7,000	
Statistical Analysis	-	19,950	$ 19,950	
Total Review Fees	$ 193,500	$ 127,225	$ (66,275)	-34%
Cost of TAR Ranking	$ -	$ 9,000	$ 9,000	
Total Cost of Review	$ 193,500	$ 136,225	$ (57,275)	-30%

[1] Review 90,000 documents

[2] Review 1,950 documents

[3] Review 6,975 documents

Assumptions			
Number of Documents		Per Hour Rates	
After Keyword Filtering	90,000	Contract	$65
Reviewed per Hour	75	Law Firm Staff	$175
Relevant	18,000	Law Firm Associa	$350
TAR Per Document Cost	$0.10	Statistical Analys	$285

The pie-charts below show the breakdown of costs from the above model and are sized relative to the overall costs of each approach.

Linear
$193,500

TAR
$136,225
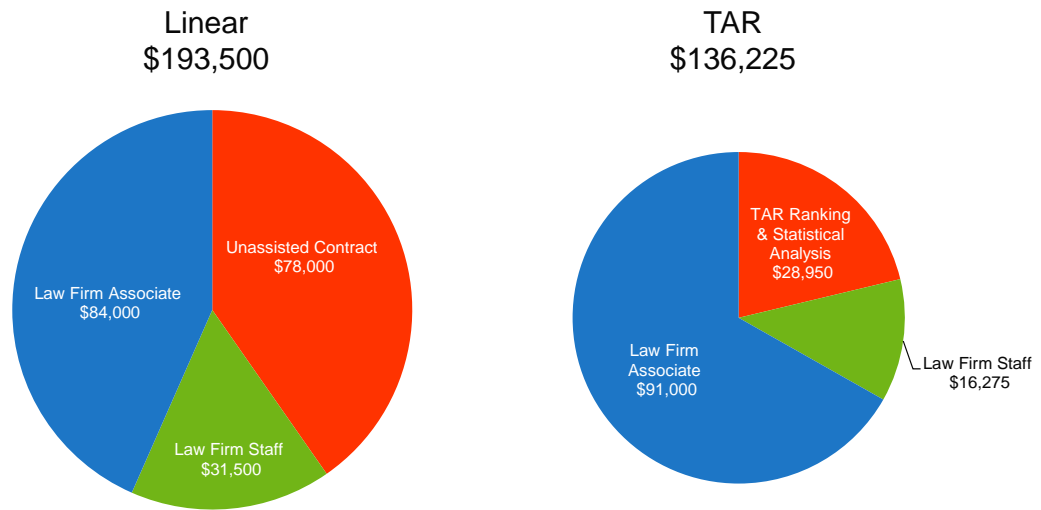
Risk impact addressed by stat sampling

Beyond the cost savings, another significant benefit of using TAR and sampling compared to traditional linear review is the validated consistency of the result. Traditional review approaches rely on having attorneys review the documents linearly, usually without statistical sampling to validate the results. Although statistical sampling could measure the reliability of unassisted linear review, the results would likely reveal just how inconsistent reviewers are, so such sampling is rarely done.

 In a paper published for the 2011 DESI IV conference comparing the results of seven separate review groups reviewing the same set of twenty-eight-thousand documents, the level of agreement on the responsive determination alone was only 9%, and on non-responsive was only 34% of the total document family count.[11] Another study published in 2012 by Cormack and Grossman found that inconsistent assessments of responsiveness are due in large part to human error. [12] Achieving review consistency without TAR is usually accomplished by the expensive and less effective approach of reviewing the documents multiple times, sometimes referred to as multi-level review. Applying TAR validated with statistical sampling yields far more consistent and reliable results than unassisted attorney review, with the added assurance that the process achieved a measurable level of success identifying the relevant and responsive documents. Unassisted review cannot achieve those results as cost effectively.

---

[11] T Barnett, S Godjevac, Faster, better, cheaper legal document review, pipe dream or reality? June 2011; Page 1, http://www.umiacs.umd.edu/~oard/desi4/papers/barnett3.pdf

[12] Gordon Cormack and Maura Grossman. Inconsistent Assessment of Responsiveness in E-Discovery: Difference of Opinion or Human Error?, 32 Pace L. Rev. 267 (2012)

Necessary Process Changes to Realize Benefits of TAR and Statistical Sampling

Before companies can achieve the efficiencies and improved results from TAR they have to change the review process to take advantage of TAR's ability to focus review on the most likely relevant documents. TAR typically uses a classifier to score documents, with higher scores more likely to be relevant. If the review team still reviews all of the documents as if no scoring had been done, the result will be to increase the cost of the review by the cost of applying the classifier.

To achieve cost savings, low scored documents cannot be reviewed in their entirety. That does not mean they will be ignored completely, but they can be sampled to rule out an unacceptably high frequency of actually relevant documents ranked below the relevance threshold.   In a typical review, the majority of documents will have low scores. Sampling is an effective way to determine the rate at which documents are incorrectly classified as low score or irrelevant. This approach provides a numerical benchmark at the level of confidence selected by the producing party so that the legal team can evaluate whether the error rate is reasonable given the proportionality constraints of discovery in a given matter. This is particularly valuable when, as is generally the case, incorrectly classified low score documents are only technically relevant, and do not contain novel, relevant content.

Knowing how to stratify the review process so that the highest ranked documents are reviewed first by attorneys with a good understanding of the case can save time since they will not have to wait for the traditional "first-level" review. The documents ranked in the middle can be reviewed by contract or staff attorneys in descending order until the rate of identifying relevant documents becomes so low that further review is uneconomic. Finally, the remaining documents with the lowest scores can be sampled to confirm that the process has not missed too many relevant documents.   Such sampling can also be supplemented by keyword searches or any other quality assurance process that the producing party desires.

Conclusion

Companies and law firms that take advantage of the rapid advances in TAR will be able to keep eDiscovery review costs down and reduce the investment in discovery by getting to the relevant facts faster. Those firms who stick with unassisted manual review processes will likely be left behind.

Disclaimer