

---

# **The E-Discovery Games:** **A Closer Look at Technology Assisted Document Review**

David D. Lewis, Ph.D., Information Retrieval Consultant

Kara M. Kirkeby, Esq., Document Review Manager, Kroll Ontrack

# Dave Lewis, Ph.D.

- President, David D. Lewis Consulting
- Co-founder TREC Legal Track
- Testifying expert in *Kleen Products, LLC, et al. v. Packaging Corp. of America, et al*
- Fellow of the American Association for the Advancement of Science
- 75+ publications; 8 patents in:
  - e-discovery
  - information retrieval
  - machine learning
  - natural language processing
  - applied statistics
- Past research positions: University of Chicago, Bell Labs, AT&T Labs
- <http://www.DavidDLewis.com>



# Kara M. Kirkeby, Esq.

- Manager of Document Review Services for Kroll Ontrack
- Previously managed document reviews on complex matters for a large law firm
- Member: Minnesota State Bar Association (Civil Litigation Section), the Hennepin County Bar Association, the American Bar Association, Minnesota Women Lawyers (Communications Committee)
- Served as a judicial law clerk for Hon. Karen Klein, Magistrate judge of the U.S. District Court of North Dakota
- J.D., magna cum laude, Hamline University School of Law
- E-mail: [kkirkeby@krollontrack.com](mailto:kkirkeby@krollontrack.com)



# Discussion Overview

- What is Technology Assisted Review (TAR)?
- Document Evaluation
- Putting TAR into Practice
- Conclusion

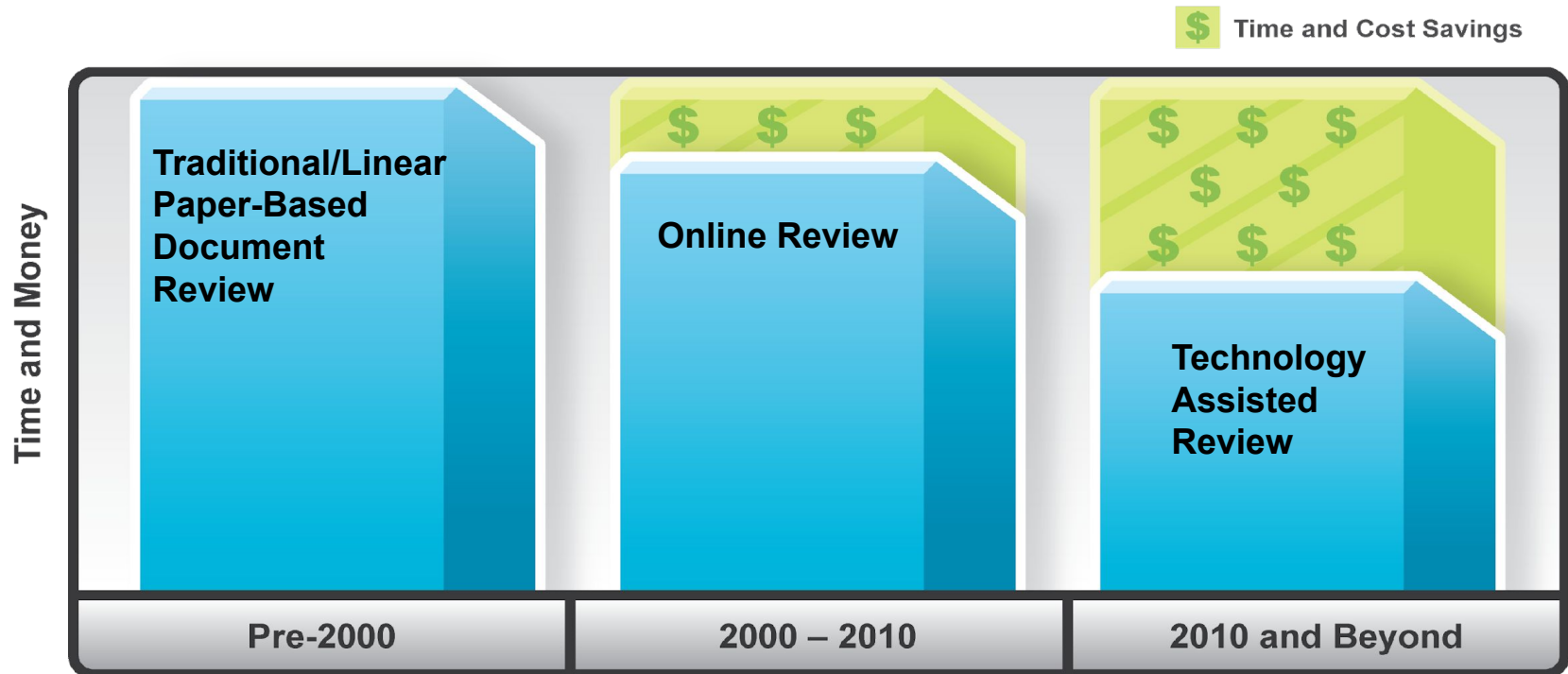


---

# What is Technology Assisted Review?

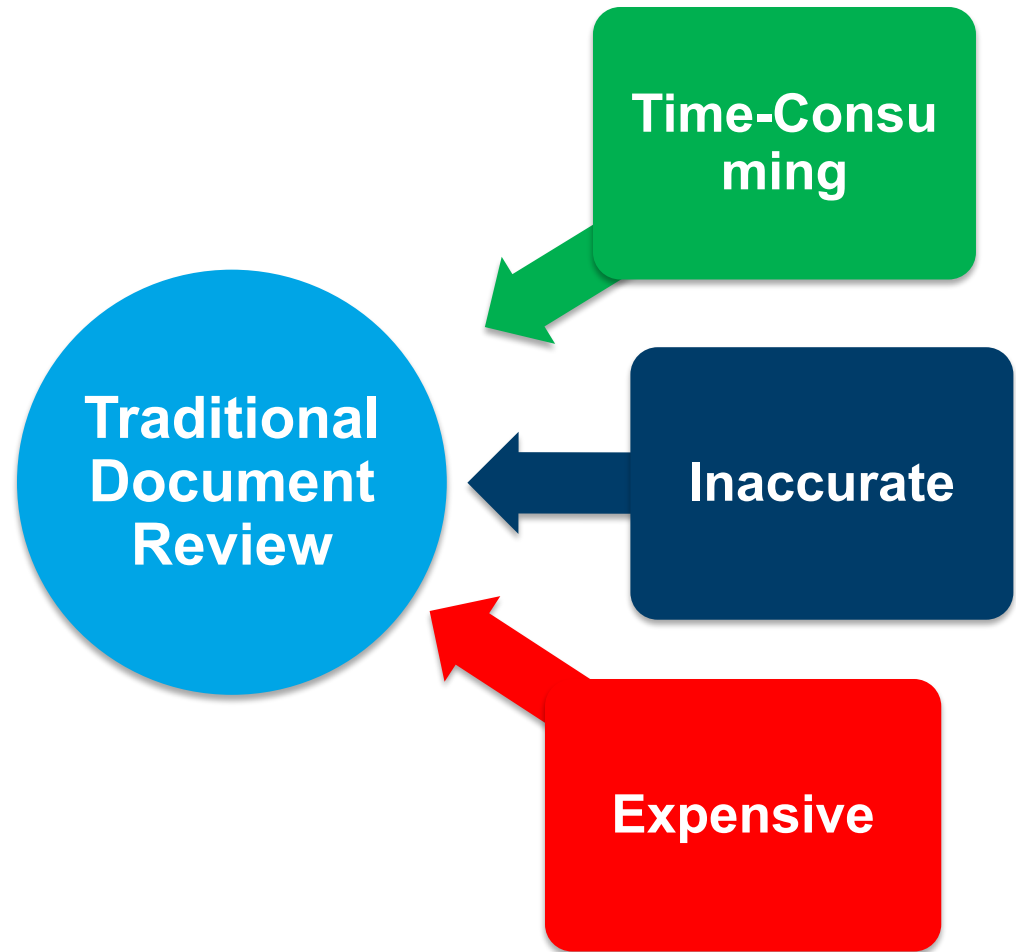
# Why Discuss Alternative Document Review Solutions?

Document review is routinely the most expensive part of the discovery process. Saving time and reducing costs will result in satisfied clients.



# Why Discuss Alternative Document Review Solutions?

- **Conducting a traditional linear document review is not particularly efficient anymore**
- Focus instead on a relevance driven review process involving lawyers and technology working together



# What Is Technology Assisted Review (TAR)?

## Three major technologies:

- ✓ Supervised learning from manual coding
- ✓ Sampling and statistical quality control
- ✓ Workflow to route documents, capture manual decisions, and tie it all together in a unified process

		True	
		X	Y
Predicted	X	<i>a</i>	<i>b</i>
	Y	<i>c</i>	<i>d</i>

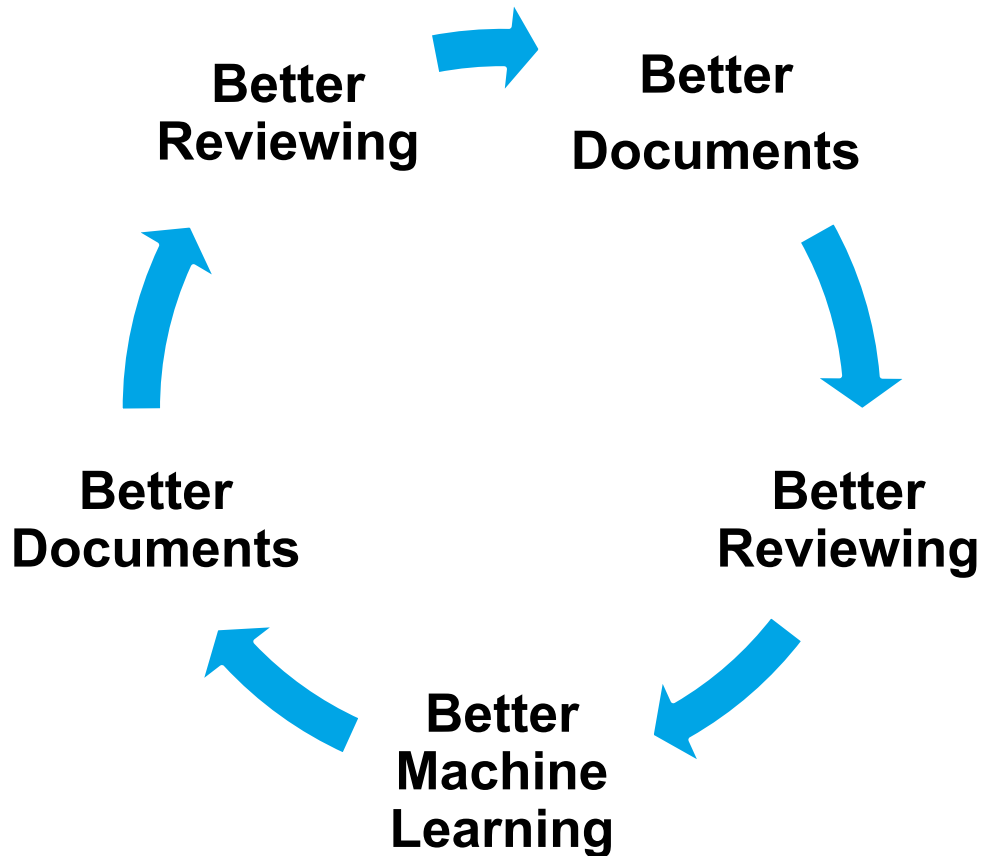
**recall: 85% +/- 4%**

**precision: 75% +/- 3%**



# Supervised Learning: The Backbone of TAR

By iterating supervised learning, you target documents most likely to be relevant or on topic, creating a virtuous cycle:



# Supervised Learning: The Backbone of TAR

- Software learns to imitate human actions
- For e-discovery, this means learning of *classifiers* by imitating human coding of documents
- Any content-based sorting into classes can be imitated
  - Responsive vs. Non-responsive
  - Privileged vs. Non-privileged
  - Topic A vs. Topic B vs. Topic C
- Widely used outside e-discovery:
  - Spam filtering
  - Computational advertising
  - Data mining

# Research & Development: TREC Legal Track

- *Text REtrieval Conference* (“TREC”), hosted by *National Institute of Standards and Technology* (“NIST”) since 1992
  - Evaluations open to academics and industry
- TREC Legal Track (since 2006) provides simulated review for responsiveness task
- Focus is on comparing technology assisted approaches
  - Not a human vs. machine bakeoff
  - Not a product benchmark
- However, results suggest advantages to technology assisted review

# Research & Development: TREC Legal Track

## 1. High effectiveness of TAR runs

- Best T-A runs in TREC 2009 examined 0.5% to 4.1% of collection while finding an estimated 76.7% of responsive documents with 84.7% precision

## 2. Low effectiveness of manual review

- Substantial effort needed by TREC organizers to clean up manual review to point it can be used as gold standard

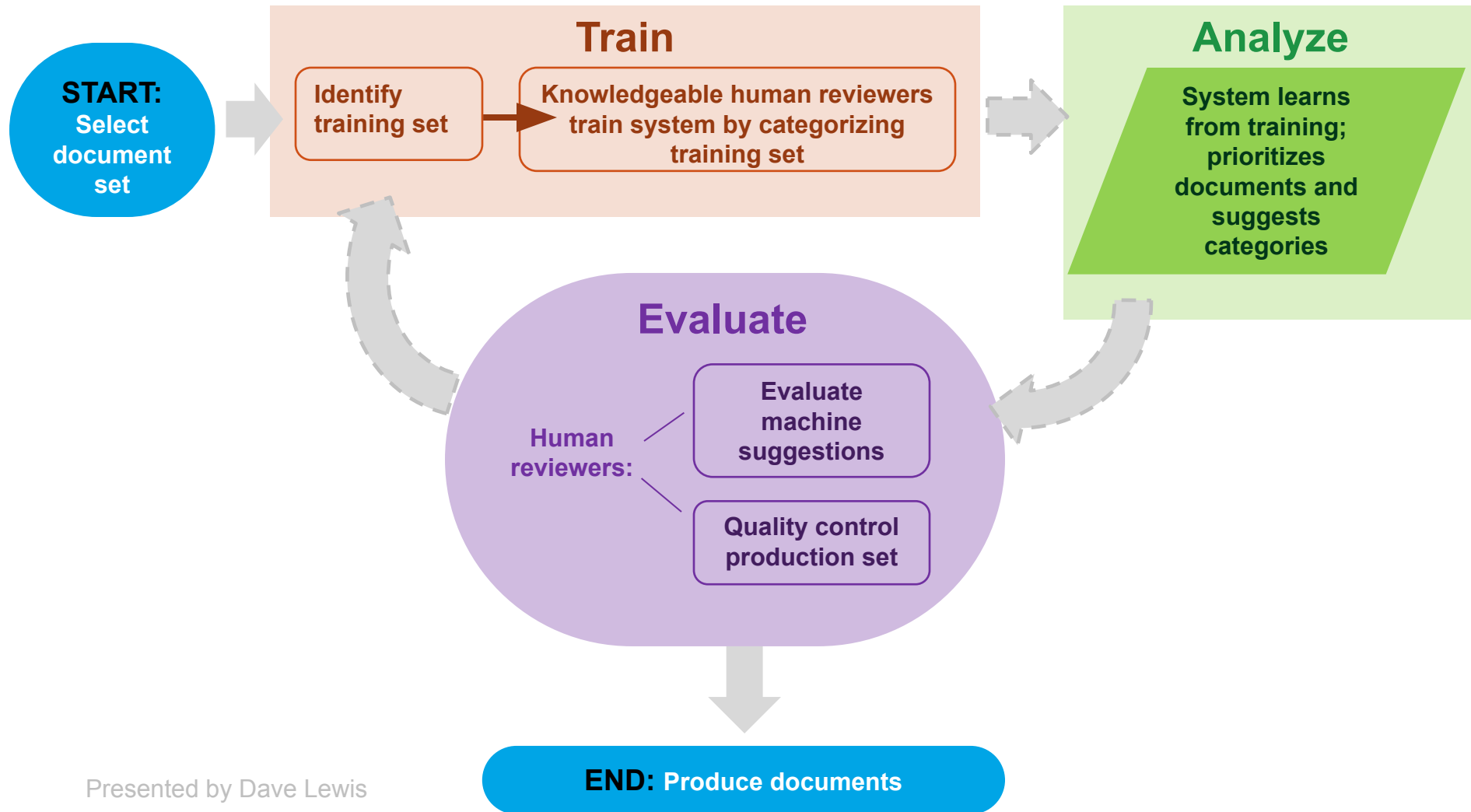
## 3. An argument can be made (Grossman & Cormack, 2011) that 2009 data shows **TAR results better than pre-cleanup manual review**

### Text REtrieval Conference (TREC)

*...to encourage research in information retrieval from large text collections.*

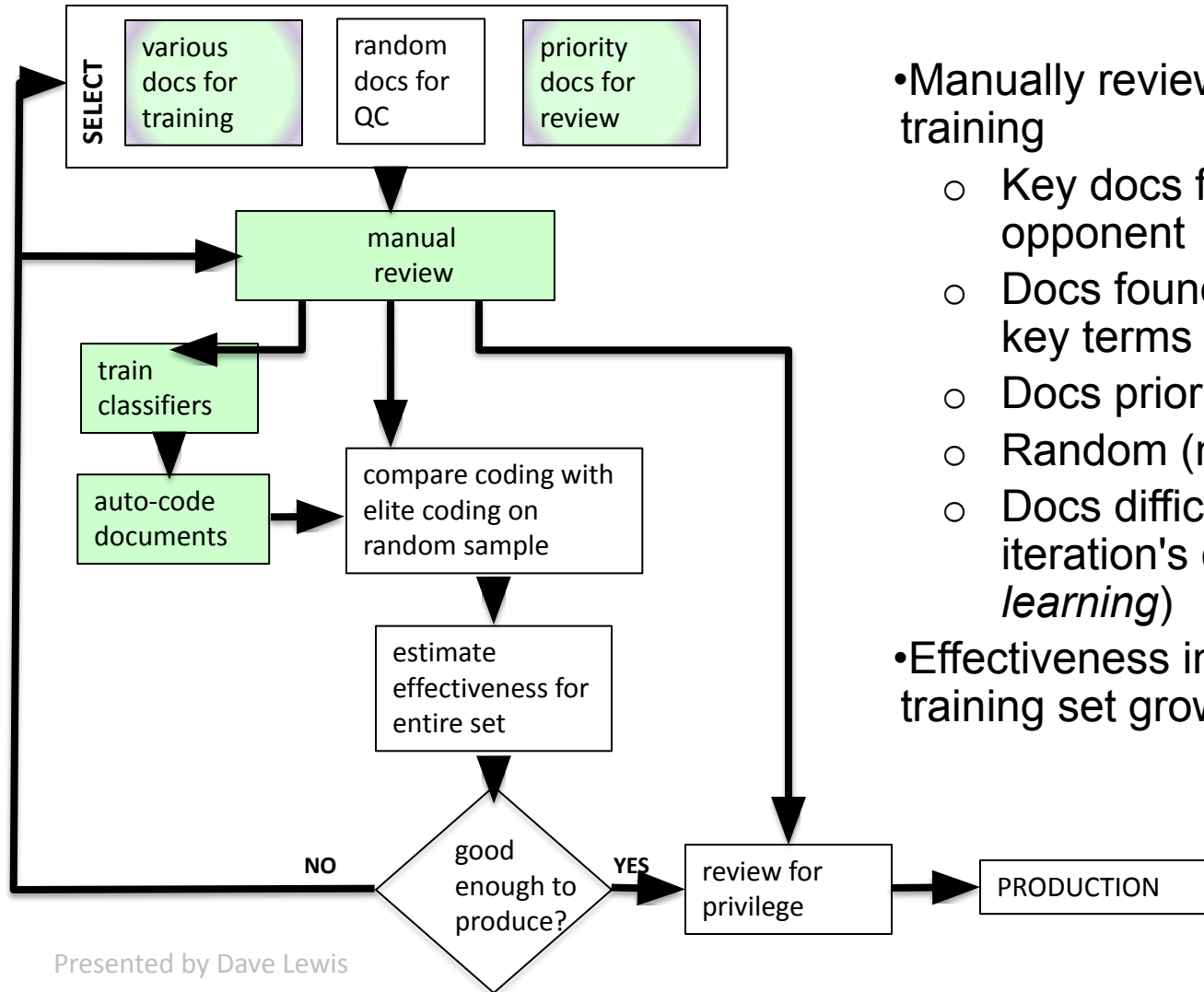


# What is Technology Assisted Review?



Presented by Dave Lewis

# Learning and Classification

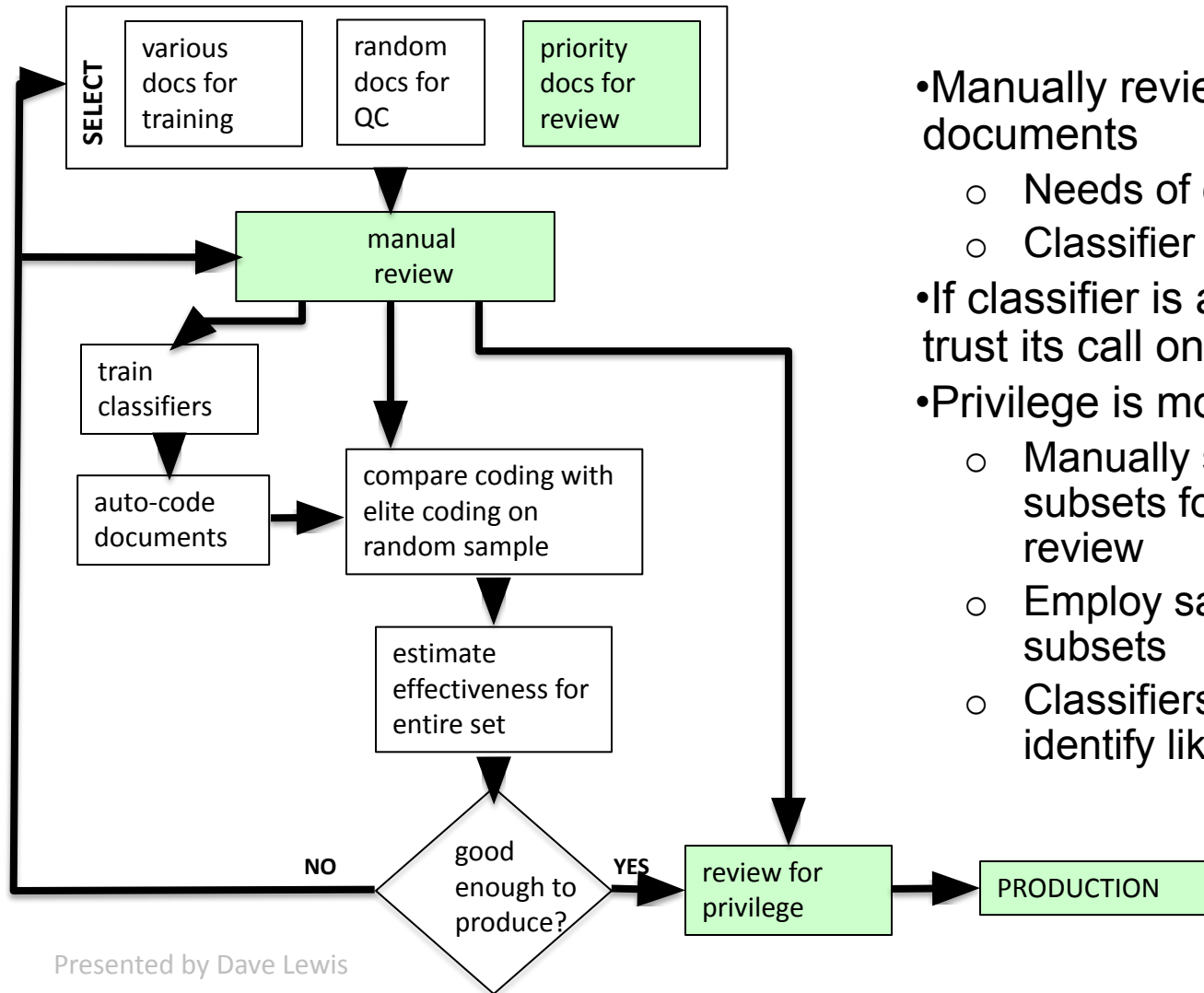


- Manually review documents for training

- Key docs from your side or opponent
- Docs found by searches on key terms
- Docs prioritized for review
- Random (non-QC) docs
- Docs difficult for previous iteration's classifier (*active learning*)

- Effectiveness increases as training set grows

# Production



- Manually review prioritized documents
  - Needs of case
  - Classifier predictions
- If classifier is accurate enough, trust its call on responsiveness?
- Privilege is more sensitive
  - Manually select some subsets for 100% privilege review
  - Employ sampling for other subsets
  - Classifiers can also help identify likely privileged docs

Presented by Dave Lewis

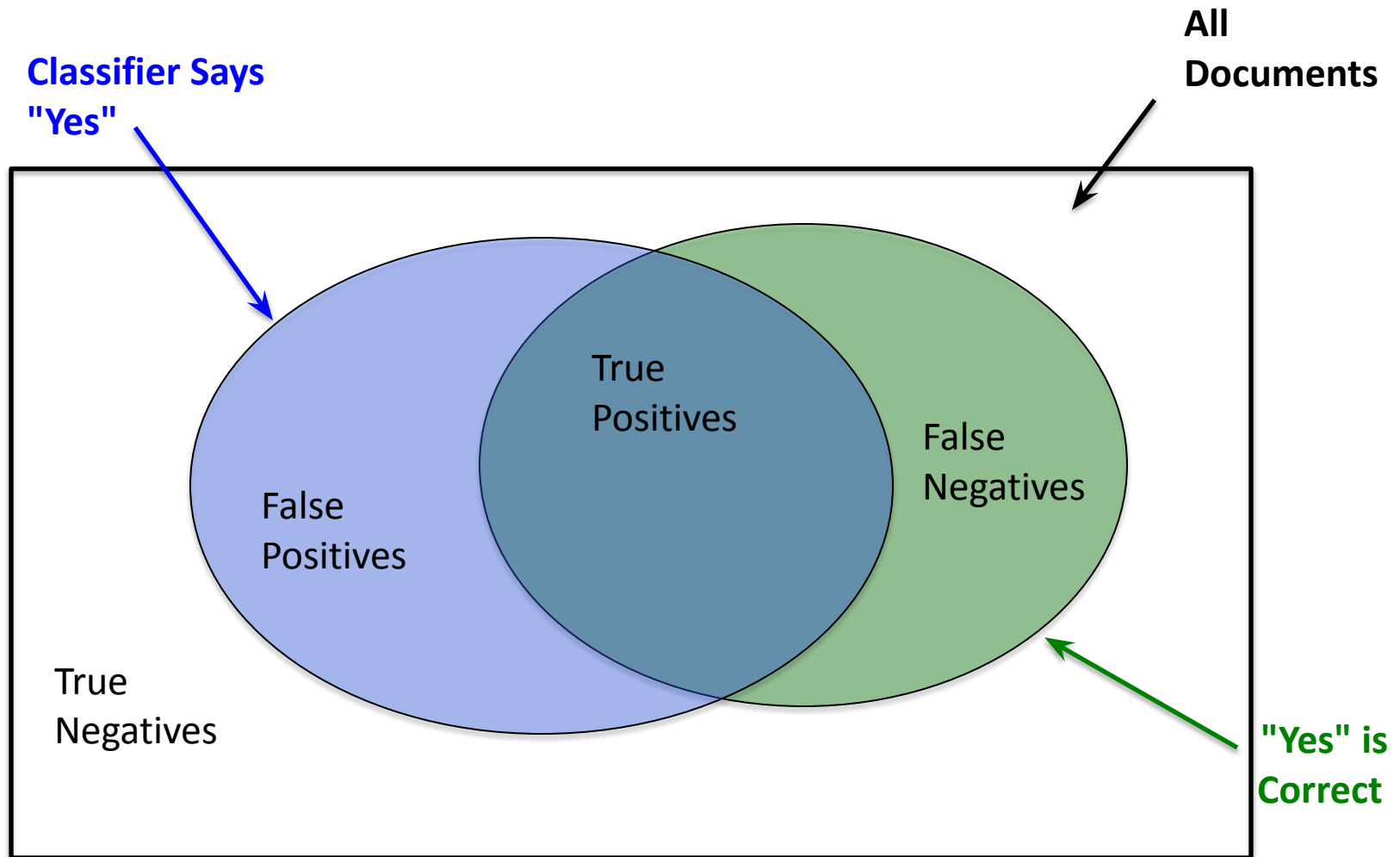
# Classification Effectiveness

		"Truth"	
		Yes	No
Prediction	Yes	<i>TP (true positives)</i>	<i>FP (false positives)</i>
	No	<i>FN (false negatives)</i>	<i>TN (true negatives)</i>

- Any binary classification can be summarized in a 2x2 table
  - Linear review, automated classifier, machine-assisted...
  - Responsive v. non-responsive, privileged v. non-privileged...
- Test on sample of  $n$  documents for which we know answer
  - $TP + FP + FN + TN = n$



# Classification Effectiveness



Presented by Dave Lewis

# Classification Effectiveness

		"Truth"	
		<i>Yes</i>	<i>No</i>
Prediction	<i>Yes</i>	<i>TP (true positives)</i>	<i>FP (false positives)</i>
	<i>No</i>	<i>FN (false negatives)</i>	<i>TN (true negatives)</i>

- Recall =  $TP / (TP+FN)$ 
  - Proportion of interesting stuff that the classifier actually found
- High recall of interest to both producing and receiving party

# Classification Effectiveness

		"Truth"	
		<i>Yes</i>	<i>No</i>
Pre dic tio n	<i>Yes</i>	<i>TP (true positives)</i>	<i>FP (false positives)</i>
	<i>No</i>	<i>FN (false negatives)</i>	<i>TN (true negatives)</i>

- Precision =  $TP / (TP+FP)$ 
  - Proportion of stuff found that was actually interesting
- High precision of particular interest to producing party: cost reduction!

# Research & Development: Blair & Maron

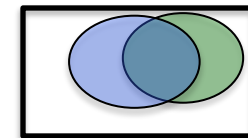
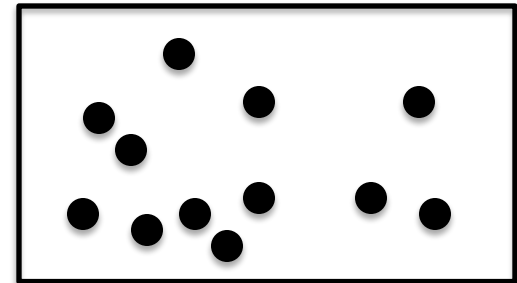
## Seminal 1985 study by Blair & Maron

- Review for documents relevant to 51 requests related to BART crash
- Boolean queries used to select documents for review
  - Process iterated until reviewer satisfied 75% of responsive documents found
- Sampling showed recall of ***less than 20%***
- B&M has been used to argue for everything from exhaustive manual review to strong AI
  - Real lesson is about need for sampling!

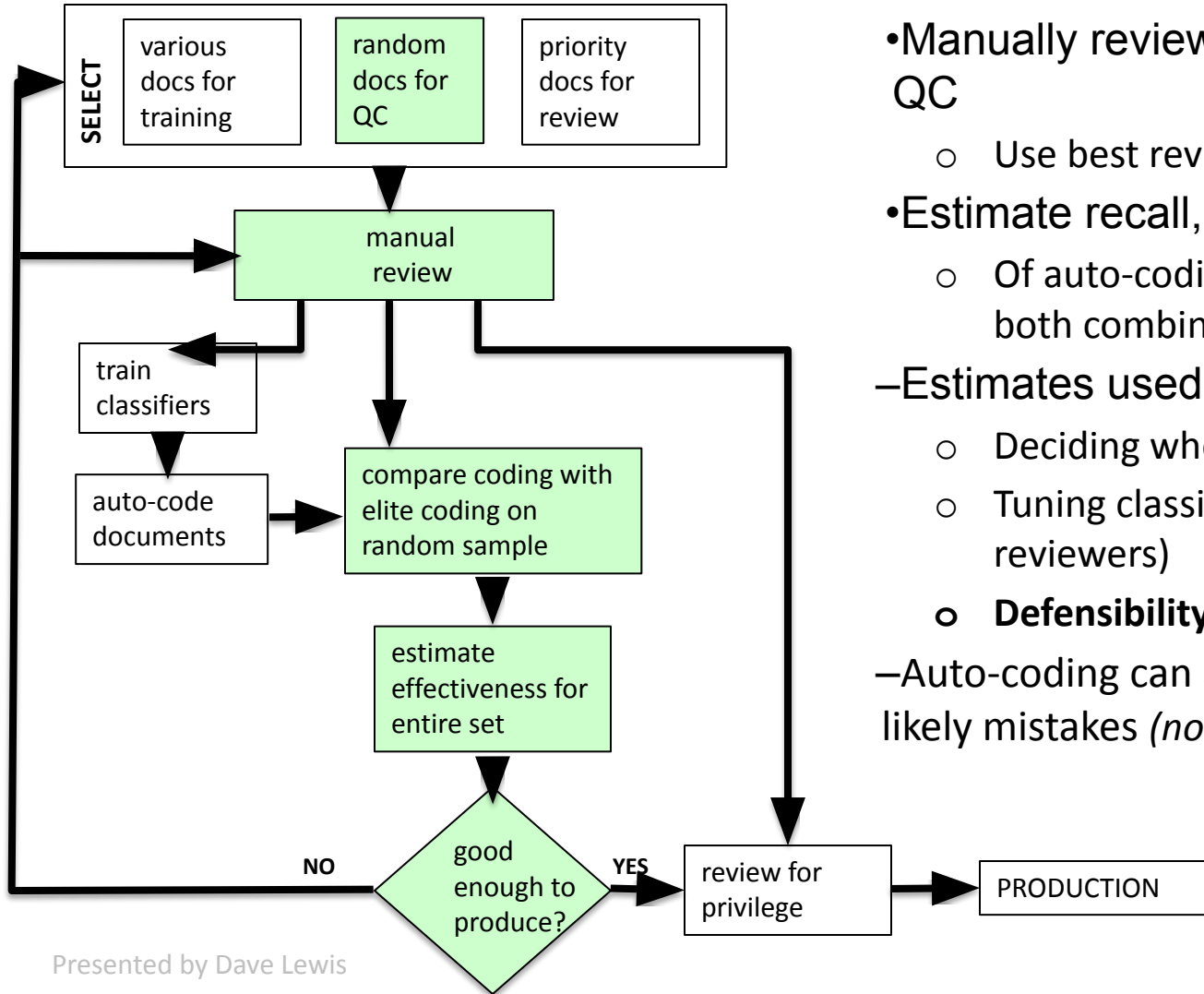


# Sampling and Quality Control

- Want to know effectiveness without manually reviewing everything. So:
  - Randomly sample the documents
  - Manually classify the sample
  - *Estimate* effectiveness on full set based on sample
- Type of estimates:
  - *Point estimate*, e.g. F1 is 0.74
  - *Interval estimate*, e.g. F1 in [0.67,0.83] with 95% confidence
- Sampling is well-understood
  - Common in expert testimony in range of disciplines



# Sampling and Quality Control



- Manually review random sample for QC
  - Use best reviewers here
- Estimate recall, precision, etc.
  - Of auto-coding, manual review, or both combined
- Estimates used in:
  - Deciding when finished
  - Tuning classifiers (and managing reviewers)
  - **Defensibility**
- Auto-coding can also be used to find likely mistakes (*not shown*)

Presented by Dave Lewis

---

# Putting TAR into Practice

# Barriers to Widespread Adoption



- **Industry-wide concern: *Is it defensible?***
- **Concern arises from misconceptions about how the technology works in practice**
  - » Belief that technology is devoid of any human interaction or oversight
  - » Confusing “smart” technologies with older technologies such as concept clustering or topic grouping
  - » Limited understanding of underlying “black box” technology
- **Largest barrier: Uncertainty over judicial acceptance of this approach**
  - » Limited commentary from the bench in the form of a court opinion
  - » Fear of being the judiciary’s “guinea pig”



# Developing TAR Case Law

- ***Da Silva Moore v. Publicis Groupe***

- » Class-action suit: parties agreed on a protocol signed by the court
- » Peck ordered more seeding reviews between the parties
- » “Counsel no longer have to worry about being the first ‘guinea pig’ for judicial acceptance of computer-assisted review ... [TAR] can now be considered judicially approved for use in appropriate cases.”

- **Approximately 2 weeks after Peck’s *Da Silva Moore* opinion**, District Court Judge Andrew L. Carter granted plaintiff opportunity to submit supplemental objections

- » Plaintiff later sought to recuse Judge Peck from the case

- **Stay tuned for more....**

# Developing TAR Case Law

- ***Kleen Products v. Packaging Corporation of America***
  - » Defendants had completed 99% of review, Plaintiffs argue that they should use Predictive Coding and start document review over
  - » Not clear whether Defendants did more than keyword search
- **Other notable points from *Kleen Products***
  - » Defendants assert they were testing their keyword search queries, not just guessing
    - Argue they did not use Predictive Coding because it did not exist yet
- **Stay tuned for more....**

## Technology Assisted Review: What It *Will Not* Do

- **Will not** replace or mimic the nuanced expert judgment of experienced attorneys with advanced knowledge of the case
- **Will not** eliminate the need to perform validation and QC steps to ensure accuracy
- **Will not** provide a magic button that will totally automate document review as we know it today

# Technology Assisted Review: What It *Can* Do

- **Reduce:**

- » Time required for document review and administration
- » Number of documents to review; if you choose an automated categorization or prioritization function
- » Reliance on contract reviewers or less experienced attorneys

- **Leverage** expertise of experienced attorneys

- **Increase** accuracy and consistency of category decisions (vs. unaided human review)

- **Identify** the most important documents more quickly

# TAR Accuracy

- TAR must be as accurate as a traditional review
- Studies show that computer-aided review is as effective as a manual review (if not more so)
- **Remember:** Court standard is *reasonableness*, not perfection:
  - “[T]he idea is not to make it perfect, it’s not going to be perfect. The idea is to make it significantly better than the alternative without as much cost.”

-U.S. Magistrate Judge Andrew Peck in *Da Silva Moore*

# What is Intelligent Review Technology (IRT) by Kroll Ontrack?

Augments the human-intensive document review process to conduct faster and cheaper discovery

Intelligent  
Categorization

Automated  
Workflow



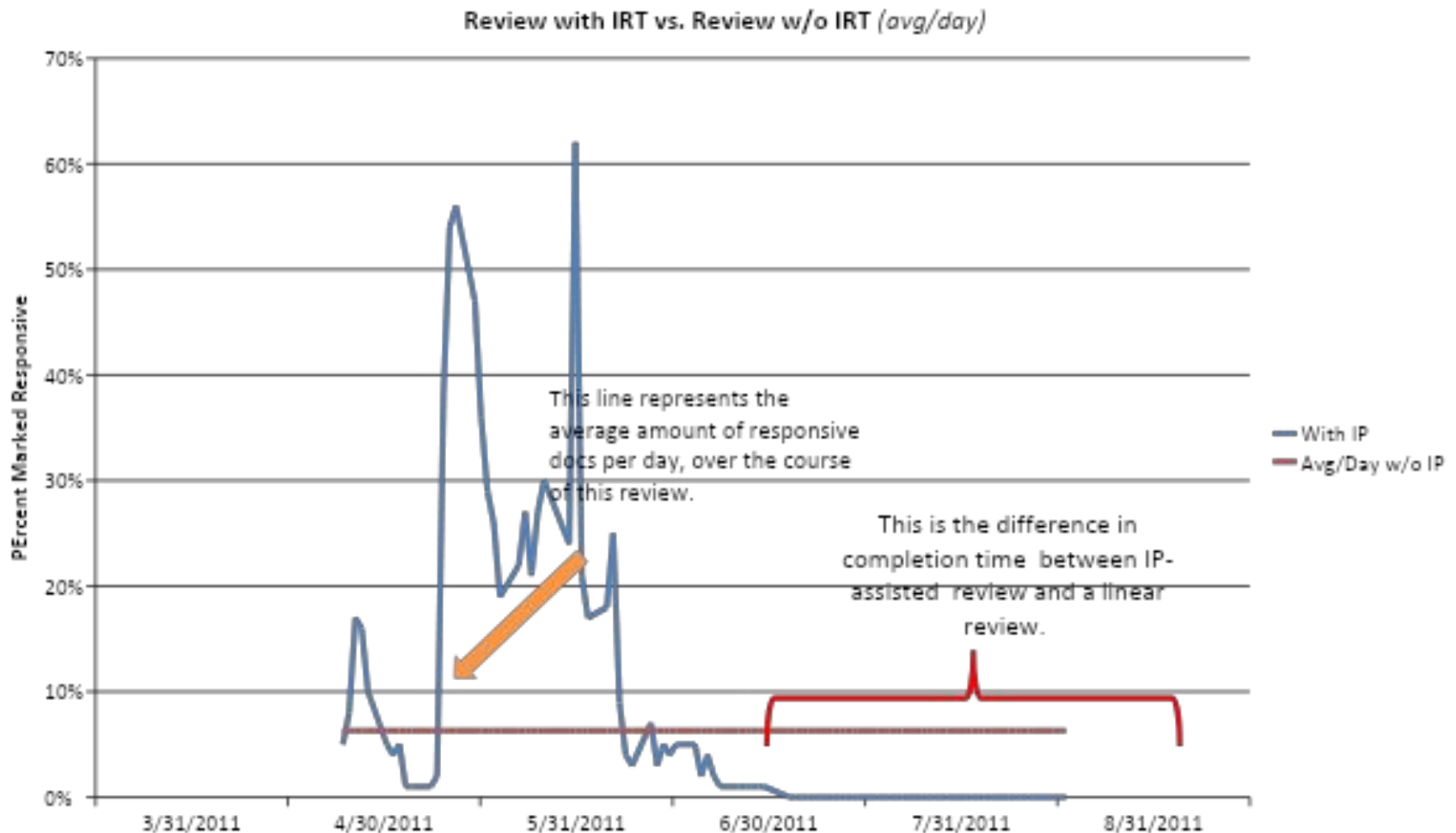
Reviewing  
Efficiently,  
Defensibly &  
Accurately

Intelligent  
Prioritization

## Successes in the Field: Kroll Ontrack's IRT

1. **Cut off review** after prioritization of documents showed marginal return of responsive documents for specific number of days
2. **Cut off review** of a custodian when, based on prioritization statistics that showed only non-responsive documents remained
3. Used suggested categorizations to **validate human categorizations**
4. Used suggested categorizations to **segregate documents as non-responsive** at >75% confidence level. After sampling that set, customer found less than .5% were actually responsive (and only marginally so). Review was cut off for that set of documents
5. Used suggested categorizations to **segregate categories suggested as privilege and responsive** at >80% confidence. Sampled, mass categorized
6. Use suggested categorizations to **mass categorize documents** and move them to the QC stage, by-passing first-level review
7. Used suggested categorizations to **find documents on a new issue** category when review was nearing completion

# Successes in the Field: Kroll Ontrack's IRT





---

# Conclusion

# Parting Thoughts

- **Automated review technology helps lawyers focus on resolution – not discovery – through available metrics**
  - » Complements human review, but **will not replace the need for skillful human analysis and advocacy**
- We are on the cusp of full-bore judicial discussion of Automated Review Technologies
  - » Closely monitor judicial opinions for breakthroughs
  - » Follow existing best practices for reasonableness and defensibility
- Not all Technology Assisted Review solutions are created equal
  - » Thoroughly vet the technology before adopting

# Q & A

