Predictive Coding: Emerging E-Discovery Tool
Leveraging E-Discovery Computer-Assisted Review to Reduce Time and Expense of Discovery

WEDNESDAY, JUNE 26, 2013
1pm Eastern  |  12pm Central  |  11am Mountain  |  10am Pacific

Today’s faculty features:
Conor R. Crowley, Principal, Crowley Law Office, McLean, Va.
Maura R. Grossman, Of Counsel, Wachtell Lipton Rosen & Katz, New York

The audio portion of the conference may be accessed via the telephone or by using your computer’s speakers. Please refer to the instructions emailed to registrants for additional information. If you have any questions, please contact Customer Service at 1-800-926-7926 ext. 10.
Tips for Optimal Quality

**Sound Quality**
If you are listening via your computer speakers, please note that the quality of your sound will vary depending on the speed and quality of your internet connection.

If the sound quality is not satisfactory and you are listening via your computer speakers, you may listen via the phone: dial **1-866-869-6667** and enter your PIN when prompted. Otherwise, please send us a chat or e-mail **sound@straffordpub.com** immediately so we can address the problem.

If you dialed in and have any difficulties during the call, press *0 for assistance.

**Viewing Quality**
To maximize your screen, press the F11 key on your keyboard. To exit full screen, press the F11 key again.
Continuing Education Credits

For CLE purposes, please let us know how many people are listening at your location by completing each of the following steps:

• In the chat box, type (1) your company name and (2) the number of attendees at your location
• Click the SEND button beside the box

If you have purchased Strafford CLE processing services, you must confirm your participation by completing and submitting an Official Record of Attendance (CLE Form).

You may obtain your CLE form by going to the program page and selecting the appropriate form in the PROGRAM MATERIALS box at the top right corner.

If you'd like to purchase CLE credit processing, it is available for a fee. For additional information about CLE credit processing, go to our website or call us at 1-800-926-7926 ext. 35.
Program Materials

If you have not printed the conference materials for this program, please complete the following steps:

• Click on the + sign next to “Conference Materials” in the middle of the left-hand column on your screen.
• Click on the tab labeled “Handouts” that appears, and there you will see a PDF of the slides for today’s program.
• Double click on the PDF and a separate page will open.
• Print the slides by clicking on the printer icon.
Technology-Assisted Review – Emerging eDiscovery Tool: Leveraging TAR to Reduce the Time and Expense of eDiscovery

Conor R. Crowley
Crowley Law Office
703.659.5030
crowley@crowleylawoffice.com

Maura R. Grossman
Wachtell, Lipton, Rosen & Katz
212.403.1391
mrgrossman@wlrk.com

Strafford Publications Webinar
June 26, 2013
Today’s Agenda

• Shortcomings of the “Traditional Approach” to Document Review

• The Many Faces of Technology-Assisted Review (“TAR”): What It Is and How It Works

• How We Know TAR Works Better Than the Alternatives

• The Efficiency of TAR: Time and Cost Savings

• The FRCP 26(g) Certification Requirement: What is a “Reasonable” TAR Process?

• Ongoing TAR Battles in the Courts

• Open TAR Issues and Take-Away Messages
What’s Wrong With the “Traditional Approach”?

• **Keyword Culling**
  – Culls out many relevant documents (“under-inclusion”)
  – Captures many irrelevant documents (“over-inclusion”)

• **Manual Review**
  – Misses lots of relevant documents (“false negatives”)
  – Identifies lots of irrelevant documents (“false positives”)

• **Keyword Culling + Manual Review Combined**
  – Even poorer result than keyword culling or manual review alone
    • Documents lost = relevant documents culled out + relevant documents missed during review
  – High review costs
    • Captured irrelevant documents must be reviewed, resulting in wasted time and effort
The Problem With Search Terms: The Blair & Maron Study (1985)

• In a landmark 1985 study, Blair & Maron found a disconnect between lawyers’ perceptions of their ability to find relevant documents, and their actual ability to do so.

• In a case with 40,000 documents (350,000 pages), expert paralegal searchers estimated that their manual searches had identified 75% of relevant documents, when in fact the results showed that only about 20% were found.

• This finding was replicated at TREC 2006 through 2008.

The Myth of the ★Perfect★ Manual Review

• **The Myth:**
  – That “eyeballs-on” review of each and every document in a massive collection of ESI will identify **essentially all** responsive (or privileged) documents; and
  – That computers are **less reliable than humans** in identifying responsive (or privileged) documents.

• **The Facts:**
  – **Humans miss a substantial number** of responsive (or privileged) documents;
  – **Computers** – aided by humans – **find at least as many** responsive (or privileged) documents as humans alone; and
  – **Computers** – aided by humans – **make fewer errors** on responsiveness (or privilege) than humans alone, and are **far more efficient** than humans.
Measures of Information Retrieval

• Recall =
  - \[ \frac{\text{# of responsive documents retrieved}}{\text{Total # of responsive documents in the entire document collection}} \]
  - (“How many of the responsive documents did I find?”)
    → A measure of **Completeness**

• Precision =
  - \[ \frac{\text{# of responsive documents retrieved}}{\text{Total # of documents retrieved}} \]
  - (“How much of what I retrieved was on-point?”)
    → A measure of **Accuracy**

• \( F_1 \) = The harmonic mean of Recall and Precision.
  - (An average weighted slightly towards the lesser of the two)
Confusion Matrix

<table>
<thead>
<tr>
<th></th>
<th>Coded Relevant</th>
<th>Coded Non-Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truly Relevant</td>
<td>True Positives (TP)</td>
<td>False Negatives (FN)</td>
</tr>
<tr>
<td>Truly Non-Relevant</td>
<td>False Positives (FP)</td>
<td>True Negatives (TN)</td>
</tr>
</tbody>
</table>

\[
\text{Recall} = \frac{TP}{TP + FN} \times 100\%
\]

\[
\text{Precision} = \frac{TP}{TP + FP} \times 100\%
\]
The Recall-Precision Trade-Off

![Graph showing the recall-precision trade-off with a star indicating perfection.](image-url)
## Effectiveness of Manual Review

<table>
<thead>
<tr>
<th>Study</th>
<th>Review</th>
<th>Recall</th>
<th>Precision</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voorhees</td>
<td>Secondary</td>
<td>52.8%</td>
<td>81.3%</td>
<td>64.0%</td>
</tr>
<tr>
<td>Voorhees</td>
<td>Tertiary</td>
<td>61.8%</td>
<td>81.9%</td>
<td>70.4%</td>
</tr>
<tr>
<td>Roitblat et al.</td>
<td>Team A</td>
<td>77.1%</td>
<td>60.9%</td>
<td>68.0%</td>
</tr>
<tr>
<td>Roitblat et al.</td>
<td>Team B</td>
<td>83.6%</td>
<td>55.5%</td>
<td>66.7%</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>Average</strong></td>
<td><strong>68.8%</strong></td>
<td><strong>69.9%</strong></td>
<td><strong>67.2%</strong></td>
</tr>
</tbody>
</table>
Defining “Technology-Assisted Review”

A process for prioritizing or coding an entire collection of documents using computer technologies that harness human judgments of one or more subject matter expert(s) on a small subset of the documents, and then extrapolate those judgments to the remaining documents in the collection.
Defining “Technology-Assisted Review” (Cont’d)

- Advanced analytic technologies and early case assessment (“ECA”) tools may aid TAR, but are **not** themselves TAR. Examples include:
  - Email threading, near de-duplication, etc.
  - Boolean, proximity, fuzzy, and concept search
  - Social network analysis and other visualization tools
  - Unsupervised Machine Learning (a.k.a. “clustering”)
Defining “Technology-Assisted Review” (Cont’d)

- There are a variety of TAR technologies:
  - **Machine Learning** (a.k.a. “Predictive Coding”)
    - Requires construction of a “training set” of relevant and non-relevant documents to train the system.
    - There are substantial differences between vendors in how the “training set” is constructed.
  
  - **Knowledge Engineering** (a.k.a. “Rule-Based” Approaches)
    - Approaches that seek to incorporate human knowledge into a computer system, e.g., decision trees.
    - Artificial Intelligence (“AI”) methods that seek to construct linguistic and other models that replicate the way humans think about complex problems.
Variants on Machine Learning

- **Passive Supervised Learning** = a set of documents is selected from the collection and coded by a human reviewer to form a “**training set**,” which is used as input to the learning system. The system then **classifies** each document in the collection as relevant or not, or **ranks** each document in the collection as to how likely it is to be relevant, given its similarity to the relevant (and dissimilarity to the non-relevant) documents in the training set.

  - Tends to be a more **static** process, therefore, the quality of the training set is critical.
Variants on Machine Learning (Cont’d)

- **Active Learning** = an initial training set – the “seed set” – is chosen, coded by a human reviewer, and used as input to the learning system. The learning system then identifies additional documents from the collection, which are coded by the human reviewer and added to the “training set.” This process is repeated several times.

  - Tends to be a more **iterative** and **interactive** (i.e., self-correcting) process, therefore, the quality of the initial seed set is less critical.
Options for Selecting the Training Set for Passive Supervised or Active Machine Learning

- **Random Selection**: Documents are selected at random from the collection, coded as relevant or not, and fed to the learning system.
  
  - Potential problem with this method: *May miss certain types of relevant documents that are rare in the collection and, therefore, unrepresented in a random sample; unbelievably slow.*

- **Judgmental Selection**: Documents are selected by an expert, using prior knowledge, *ad hoc* search, and advanced analytic tools, and fed to the learning system.
  
  - Potential problem with this method: *May miss certain types of relevant documents that are overlooked by the expert, or not identified by the search and analytic tools.*
Recall Versus Number of Documents Reviewed for an Actual Matter (N=294,000)
Active Learning Alternatives

• **Uncertainty Sampling**: After learning from the examples in the initial seed set, the learning method identifies documents about which relevance is **least certain**; that is, those which it finds “too close to call.” These documents are then coded by a human reviewer and added to the training set. *(Cf., an ophthalmologist's refraction test.)*

• **Relevance Feedback**: After learning from the examples in the seed and training sets, the learning method identifies those documents that it finds are **most likely to be relevant**. These documents are then coded by a human reviewer and added to the training set, growing the training set iteratively.
The Text REtrieval Conference (“TREC”): Measuring the Effectiveness of Technology-Assisted Review

- International, interdisciplinary research project sponsored by the National Institute of Standards and Technology (“NIST”), which is part of the U.S. Department of Commerce.
- Designed to promote research into the science of information retrieval.
- First TREC was held in 1992; TREC Legal Track began in 2006.
- Designed to evaluate the effectiveness of search technologies in the context of e-discovery.
- Employed hypothetical complaints and requests for production drafted by members of The Sedona Conference®.
- For the first three years (2006-2008), the documents used were from the publicly available 7 million-document tobacco litigation Master Settlement Agreement database.
- For 2009-2011, publicly available Enron data sets were used.
- Participating teams of information retrieval scientists from around the world and U.S. litigation support service providers contributed computer runs attempting to identify responsive, highly responsive, or privileged documents in response to requests for production.
Effectiveness of Technology-Assisted Reviews at TREC 2009

<table>
<thead>
<tr>
<th>Team</th>
<th>Topic</th>
<th># Docs. Produced</th>
<th>Recall</th>
<th>Precision</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waterloo</td>
<td>201</td>
<td>2,154</td>
<td>77.8%</td>
<td>91.2%</td>
<td>84.0%</td>
</tr>
<tr>
<td>Waterloo</td>
<td>202</td>
<td>8,746</td>
<td>67.3%</td>
<td>88.4%</td>
<td>76.4%</td>
</tr>
<tr>
<td>Waterloo</td>
<td>203</td>
<td>2,719</td>
<td>86.5%</td>
<td>69.2%</td>
<td>76.9%</td>
</tr>
<tr>
<td>H5</td>
<td>204</td>
<td>2,994</td>
<td>76.2%</td>
<td>84.4%</td>
<td>80.1%</td>
</tr>
<tr>
<td>Waterloo</td>
<td>207</td>
<td>23,252</td>
<td>76.1%</td>
<td>90.7%</td>
<td>82.8%</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>7,973</td>
<td>76.7%</td>
<td>84.7%</td>
<td>80.0%</td>
</tr>
</tbody>
</table>
Manual Versus Technology-Assisted Reviews

<table>
<thead>
<tr>
<th>Study</th>
<th>Review</th>
<th>Recall</th>
<th>Precision</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voorhees</td>
<td>Secondary</td>
<td>52.8%</td>
<td>81.3%</td>
<td>64.0%</td>
</tr>
<tr>
<td>Voorhees</td>
<td>Tertiary</td>
<td>61.8%</td>
<td>81.9%</td>
<td>70.4%</td>
</tr>
<tr>
<td>Roitblat et al.</td>
<td>Team A</td>
<td>77.1%</td>
<td>60.9%</td>
<td>68.0%</td>
</tr>
<tr>
<td>Roitblat et al.</td>
<td>Team B</td>
<td>83.6%</td>
<td>55.5%</td>
<td>66.7%</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>68.8%</td>
<td>69.9%</td>
<td>67.2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Team</th>
<th>Topic</th>
<th>Recall</th>
<th>Precision</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waterloo</td>
<td>201</td>
<td>77.8%</td>
<td>91.2%</td>
<td>84.0%</td>
</tr>
<tr>
<td>Waterloo</td>
<td>202</td>
<td>67.3%</td>
<td>88.4%</td>
<td>76.4%</td>
</tr>
<tr>
<td>Waterloo</td>
<td>203</td>
<td>86.5%</td>
<td>69.2%</td>
<td>76.9%</td>
</tr>
<tr>
<td>H5</td>
<td>204</td>
<td>76.2%</td>
<td>84.4%</td>
<td>80.1%</td>
</tr>
<tr>
<td>Waterloo</td>
<td>207</td>
<td>76.1%</td>
<td>90.7%</td>
<td>82.8%</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>76.7%</td>
<td>84.7%</td>
<td>80.0%</td>
</tr>
</tbody>
</table>
But!

- Roitblat et al., Voorhees, and the TREC 2009 Interactive Task all used **different datasets**, **different topics**, and **different gold standards**, so we cannot directly compare them.

- While technology-assisted review **appeared to be** at least as good as manual review, we had to control for these differences.
# TREC 2009 Topics Used in Grossman-Cormack Study

<table>
<thead>
<tr>
<th>Topic</th>
<th>Request for Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>201</td>
<td>All documents or communications that describe, discuss, refer to, report on, or relate to the Company’s engagement in structured commodity transactions known as “prepay transactions.”</td>
</tr>
<tr>
<td>202</td>
<td>All documents or communications that describe, discuss, refer to, report on, or relate to the Company's engagement in transactions that the Company characterized as compliant with FAS 140 (or its predecessor FAS 125).</td>
</tr>
<tr>
<td>203</td>
<td>All documents or communications that describe, discuss, refer to, report on, or relate to whether the Company had met, or could, would, or might meet its financial forecasts, models, projections, or plans at any time after January 1, 1999.</td>
</tr>
<tr>
<td>204</td>
<td>All documents or communications that describe, discuss, refer to, report on, or relate to any intentions, plans, efforts, or activities involving the alteration, destruction, retention, lack of retention, deletion, or shredding of documents or other evidence, whether in hard-copy or electronic form.</td>
</tr>
<tr>
<td>207</td>
<td>All documents or communications that describe, discuss, refer to, report on, or relate to fantasy football, gambling on football, and related activities, including but not limited to, football teams, football players, football games, football statistics, and football performance.</td>
</tr>
</tbody>
</table>
**Effectiveness of TREC 2009 Manual Versus Technology-Assisted Reviews**

<table>
<thead>
<tr>
<th>Topic</th>
<th>Team</th>
<th>Recall</th>
<th>Precision</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>201</td>
<td>Waterloo TREC (Law Students)</td>
<td>(†) 77.8%</td>
<td>(*) 91.2%</td>
<td>(*) 84.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75.6%</td>
<td>5.0%</td>
<td>9.5%</td>
</tr>
<tr>
<td>202</td>
<td>Waterloo TREC (Law Students)</td>
<td>67.3%</td>
<td>(*) 88.4%</td>
<td>(*) 76.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(†) 79.9%</td>
<td>26.7%</td>
<td>40.0%</td>
</tr>
<tr>
<td>203</td>
<td>Waterloo TREC (Professionals)</td>
<td>(*) 86.5%</td>
<td>(*) 69.2%</td>
<td>(*) 76.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25.2%</td>
<td>12.5%</td>
<td>16.7%</td>
</tr>
<tr>
<td>204</td>
<td>H5 TREC (Professionals)</td>
<td>(*) 76.2%</td>
<td>(*) 84.4%</td>
<td>(*) 80.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>36.9%</td>
<td>25.5%</td>
<td>30.2%</td>
</tr>
<tr>
<td>207</td>
<td>Waterloo TREC (Professionals)</td>
<td>76.1%</td>
<td>(†) 90.7%</td>
<td>82.8%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(†) 79.0%</td>
<td>89.0%</td>
<td>(†) 83.7%</td>
</tr>
<tr>
<td>Avg.</td>
<td>H5 / Waterloo TREC</td>
<td>(†) 76.7%</td>
<td>(*) 84.7%</td>
<td>(*) 80.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>59.3%</td>
<td>31.7%</td>
<td>36.0%</td>
</tr>
</tbody>
</table>

Results marked (*) are superior and overwhelmingly significant ($P < 0.0001$)

Results marked (†) are superior but not statistically significant ($P > 0.1$)
Keywords Versus Manual Versus TAR

![Graph showing precision versus recall for different methods: Blair & Maron Keyword Search, TREC 2009 Technology-Assisted Review, TREC 2009 Manual Review.]

- **Keywords**: Blair & Maron Keyword Search
- **Manual Review**: TREC 2009 Manual Review
- **TAR**: TREC 2009 Technology-Assisted Review
• **Document Request:**
  
  – All documents or communications that describe, discuss, refer to, report on, or relate to any intentions, plans, efforts, or activities involving the alteration, destruction, retention, lack of retention, deletion, or shredding of documents or other evidence, whether in hard-copy or electronic form.
Inarguable Error for Topic 204 – Should Have Been Marked Responsive

Date: Tuesday, January 22, 2002 11:31:39 GMT
Subject:

I’m in. I’ll be shredding ‘till 11am so I should have plenty of time to make it.
Inarguable Error for Topic 204 – Should Have Been Marked Non-Responsive

From: Mark Taylor
Sent: Wed May 09, 2001 19:13:00 GMT
To: Jeffrey Hodge
Subject: No More Confirms Agreement
Attachments: CONSENT AND AMENDMENT AGREEMENT.doc

Here is a first draft of an amendment agreement. Please feel free to rip it to shreds.
Let’s keep Shari on the distribution list permanently so that she can be my back up if I’m out sick, etc. She can just delete the lists when I am in the office.

Leslie

Is this forever, or just for the week?

Beginning this week, Shari will reply to EOL credit approval lists. I will be out after Tuesday of next week through Friday, Sept. 8. Will you add Shari to the distribution list effective as soon as possible so that she receives the list tomorrow.

Thanks,
Leslie
Topic 207 (TREC 2009)

- **Document Request:**
  - All documents or communications that describe, discuss, refer to, report on, or relate to fantasy football, gambling on football, and related activities, including but not limited to, football teams, football players, football games, football statistics, and football performance.
Inarguable Error for Topic 207 – Should Have Been Marked Responsive

From: Bass, Eric  
Sent: Thursday, January 17, 2002 11:19 AM  
To: Lenhart, Matthew  
Subject: FFL Dues

You owe $80 for fantasy football. When can you pay?
Inarguable Error for Topic 207 – Should Have Been Marked Non-Responsive

<table>
<thead>
<tr>
<th>From:</th>
<th>Barry Tycholiz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sent:</td>
<td>Fri Sept 14, 2001 12:21:34 GMT</td>
</tr>
<tr>
<td>To:</td>
<td>Jessica Presas</td>
</tr>
<tr>
<td>Subject:</td>
<td>Baseball Tickets</td>
</tr>
<tr>
<td>Importance:</td>
<td>Normal</td>
</tr>
<tr>
<td>Priority:</td>
<td>Normal</td>
</tr>
<tr>
<td>Sensitivity:</td>
<td>None</td>
</tr>
</tbody>
</table>

Jessica, did we place the order for the playoff tickets... BT
Arguable Error for Topic 207 – Should Have Been Marked ???

Subject: RE: How good is Temptation Island 2
They have some cute guy lawyers this year—but I bet you probably watch that manly Monday night Football.
What Causes Attorneys to Miscode Documents?

- There is a perception that because human reviewers disagree on the responsiveness of many documents, the concept of responsiveness must be “fuzzy” and ill-defined.

- If this hypothesis is correct, reviewer disagreement should be the result of “fringe” or “arguable” documents, rather than the result of clear-cut errors.

- Approximately 90% of first-pass miscoding was due to clear-cut human error.

- Approximately 5% of second-pass miscoding was due to clear-cut human error.

- Only 5% of miscoding was due to “arguable” or “fringe” documents.
Demonstration of Topic 204 from TREC 2009

http://cormack.uwaterloo.ca/abademo
Efficiency of Technology-Assisted Versus Exhaustive Manual Reviews

<table>
<thead>
<tr>
<th>Topic</th>
<th>Review Effort (# Docs.)</th>
<th>Effectiveness ($F_1$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>201</td>
<td>6,145</td>
<td>836,165</td>
</tr>
<tr>
<td>202</td>
<td>12,646</td>
<td>836,165</td>
</tr>
<tr>
<td>203</td>
<td>4,369</td>
<td>836,165</td>
</tr>
<tr>
<td>204</td>
<td>20,000</td>
<td>836,165</td>
</tr>
<tr>
<td>207</td>
<td>34,446</td>
<td>836,165</td>
</tr>
<tr>
<td>Average</td>
<td>15,521</td>
<td>836,165</td>
</tr>
</tbody>
</table>

- Exhaustive manual review involves coding **100%** of the documents, while technology-assisted review involves coding of between **0.5%** (Topic 203) and **4.1%** (Topic 207) of the documents (**1.9%, on average, per topic**).
- Therefore, **on average,** technology-assisted review was **50 times** more efficient than exhaustive manual review.
An Example of an Actual Document Production Using the “Traditional Approach”

- Response to a regulatory subpoena.
- Regulator supplied the search term list.
- Search terms retrieved 568 GB of data, or 2,195,712 records.
- “Traditional approach” would have required 30 contract reviewers to manually review the documents, 10 hours a day for 21 weeks.
- At the typical industry rate of $1.10 per document, the first-pass review for this matter would have cost $2,415,283.
The Same Matter Using “Technology-Assisted Review”

- Using technology-assisted review, we quickly reduce the 2,195,712 records to the 24,401 that were potentially responsive to the regulatory request (or 1.1% of the post-search term collection).
- This effort took two people (one information retrieval expert and one e-discovery expert) less than one week.
- The cost of this first-pass review was 1.1¢ per document, or approximately $25,000 (one hundred times cheaper than the “traditional approach”).
## Actual 2012 TAR Results

<table>
<thead>
<tr>
<th>Matter</th>
<th># Docs Collected</th>
<th># Docs Examined</th>
<th>% Savings Over Manual Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matter A</td>
<td>6,823,361</td>
<td>13,074</td>
<td>99.8%</td>
</tr>
<tr>
<td>Matter B</td>
<td>6,105,790</td>
<td>64,466</td>
<td>99.0%</td>
</tr>
<tr>
<td>Matter C</td>
<td>1,723,299</td>
<td>2,511</td>
<td><strong>99.9%</strong></td>
</tr>
<tr>
<td>Matter D</td>
<td>622,153</td>
<td>3,691</td>
<td>99.4%</td>
</tr>
<tr>
<td>Matter E</td>
<td>343,459</td>
<td>58,871</td>
<td><strong>85.4%</strong></td>
</tr>
<tr>
<td>Matter F</td>
<td>315,718</td>
<td>5,284</td>
<td>98.4%</td>
</tr>
<tr>
<td>Matter G</td>
<td>177,081</td>
<td>2,064</td>
<td>98.8%</td>
</tr>
<tr>
<td>Matter H</td>
<td>118,749</td>
<td>3,618</td>
<td>97.0%</td>
</tr>
<tr>
<td>Matter I</td>
<td>67,165</td>
<td>5,395</td>
<td>92.5%</td>
</tr>
</tbody>
</table>
The Certification Requirement Under Federal Rule of Civil Procedure 26(g)

- (1) **Signature Required; Effect of Signature.** Every disclosure under Rule 26(a)(1) or (a)(3) and **every discovery** request, **response** or objection must be signed by at least one attorney of record in the attorney’s own name. . . . **By signing, an attorney** or party **certifies** that to the best of the person’s knowledge, information, and belief formed **after a reasonable inquiry**:
  - (A) with respect to a disclosure, it is complete and correct as of the time it is made; and
  - (B) **with respect to a discovery** request, **response**, or objection, **it is**:
    - (i) **consistent with these rules**. . . .
    - (ii) not interposed for any improper purpose, such as to harass, cause unnecessary delay, or needlessly increase the cost of litigation; and
    - (iii) **neither unreasonable nor unduly burdensome or expensive** considering the needs of the case, prior discovery in the case, the amount in controversy, and the importance of the issues at stake in the action.
Certification Requirement Under Federal Rule 26(g) (Cont’d)

... 

• (3) *Sanction for Improper Certification.* If a certification violates this rule without substantial justification, the court, on motion or on its own, must impose an appropriate sanction on the signer, the party on whose behalf the signer was acting, or both. The sanction may include an order to pay the reasonable expenses, including attorney’s fees, caused by the violation.

• **Bottom Line:** Rule 26(g) requires parties to conduct a “reasonable” search.
What Constitutes a “Reasonable” TAR Search?

- **Using a TAR technology and process known *a priori* to work**
  - Independent testing (e.g., through TREC)
  - Vendor testing (trust, but verify)
  - **Do your own testing!**

- **Getting your adversary and the Court comfortable with what you are doing**
  - Cooperation and translucency about your search process

- **Using built-in vendor measures**
  - Trust, but verify
  - The TREC 2010 and 2011 Learning Task results show that participants are very poor at estimating their own performance.

- **Using external post hoc sampling to measure your recall**
Da Silva Moore v. Publicis Groupe


Hon. Andrew J. Peck

“This judicial opinion now recognizes that computer-assisted review is an acceptable way to search for relevant ESI in appropriate cases.”

- “Linear manual review is simply too expensive where, as here, there are over three million emails to review. Moreover, while some lawyers still consider manual review to be the ‘gold standard,’ that is a myth as statistics clearly show that computerized searches are at least as accurate, if not more so, than manual review.”

- “[K]eyword searches usually are not very effective.”

- “Computer-assisted review is an available tool and should be seriously considered in large-volume data cases where it may save the producing party (or both parties) significant amounts of legal fees in document review. Counsel no longer have to worry about being the ‘first’ or ‘guinea pig’ for judicial acceptance of computer-assisted review.... Computer-assisted review now can be considered judicially-approved for use in appropriate cases.”
Procedural Posture of *Da Silva Moore*

- Putative class and collective actions alleging gender discrimination under Title VII (and similar New York laws), pregnancy discrimination under Title VII and related violations of the Family and Medical Leave Act, as well as violations of the Equal Pay Act and Fair Labor Standards Act (and the similar New York labor law).

- 3 million documents for review.

- Defendant Publicis Groupe sought to use an Active Learning tool.

- Unclear whether the tool uses uncertainty sampling, relevance feedback, some hybrid of both, or something else entirely.

- Plaintiffs reluctantly agreed to the use of “predictive coding,” but disputed many aspects of defendant’s proposed search protocol.

- Magistrate Judge Peck approved defendant’s search protocol with minor amendments.

- District Judge Carter affirmed Magistrate Judge Peck’s order concerning the search protocol.

- Disputes over all aspects of the “predictive coding” protocol abound.

- Discovery has stalled pending the resolution of unrelated matters; the parties are in discussions over the validation protocol.

- Dispute pertaining to the collapse of three hangars at Dulles Airport.
- Two million documents were collected.
- Defendants proposed to use a variant of Passive Supervised Learning
  - Training set is randomly selected
  - Size of the training set is determined by running the tool and using an internal measure to estimate recall
  - Training set is increased iteratively by adding 100 random documents at a time until the tool internally estimates high recall (not through Active Learning)
  - Defendants agreed to release the non-privileged, non-sensitive training set documents to plaintiffs
  - Defendants promised to demonstrate 75% recall at a 95% level of confidence
- Court approved defendants’ protocol, over plaintiff’s objection, without prejudice
  - Plaintiff may still challenge the adequacy of the production after the fact
- Production was accepted without dispute.
  - Reported recall was 81% and reported precision was 80%.
• Antitrust case alleging anti-competitive price-fixing behaviors by seven cardboard box manufacturers.
• Defendants used “traditional” search methods:
  – Complex Boolean search strings
  – Validation process estimated that 5% of the documents missed by the searches were responsive (i.e., 27 responsive documents were identified in a sample of 660 purportedly non-responsive documents).
  – Plaintiffs argued (correctly) that the validation process did not measure recall and thus, did not provide any information about what percentage of the responsive documents were identified by the searches.
Kleen Products LLC v. Packaging Corp. of Am., 2012 WL 4498465 (N.D. Ill. Sept. 28, 2012) (Cont’d)

- Plaintiff is sought to compel the use of TAR (“Content Based Advanced Analytics”) with a valid recall estimate.

- After two full days of evidentiary hearings, and more than 11 meet and confers and conferences with the Court, plaintiff withdrew its motion to compel the use of TAR for the document production in question.
In Re: Actos (Pioglitazone) Products Liability Litigation, MDL No. 6:11-md-2299 (W.D. La) (Doherty, J.)

• Plaintiffs alleged that Actos, a prescription drug for the treatment of Type II Diabetes, increased the risk of bladder cancer in users.


• Parties reached an agreement on a “search methodology proof of concept to evaluate the potential utility of advanced analytics as a document identification mechanism for the review and production” of ESI.

• Protocol proposed the use of an Active Learning tool.

• Trial experiment was to be applied to the email of four key custodians.

• Laid out a complex protocol involving three experts from each side and extensive quality control measures.
**EORHB v. HOA Holdings**


**Vice Chancellor J. Travis Laster**

Complex multimillion dollar commercial indemnity dispute involving the sale of Hooters

Tr. at 66-67:

- “This seems to me to be an ideal non-expedited case in which the parties would benefit from using predictive coding. I would like you all, if you do not want to use predictive coding, to show cause why this is not a case where predictive coding is the way to go.”

- “I would like you all to talk about a single discovery provider that could be used to warehouse both sides’ documents to be your single vendor. Pick one of these wonderful discovery super powers that is able to maintain the integrity of both side’s documents and insure that no one can access the other side’s information. If you cannot agree on a suitable discovery vendor, you can submit names to me and I will pick one for you.”

- “The problem is that these types of indemnification claims can generate a huge amount of documents. That’s why I would really encourage you all, instead of burning lost of hours with people reviewing, it seems to me this is the type of non-expedited case where we could all benefit from some new technology use.”
EORHB v. HOA Holdings Redux


• On May 6, the Court entered a new Order stating that:
  • Defendants could retain their chosen vendor and utilize technology-assisted review;
  • Plaintiffs could conduct their document review using traditional review methods; and
  • The parties would not be required to retain a single vendor to be used by both sides.

- Plaintiffs filed a lawsuit against defendants alleging 11 causes of action, and seeking over $1 billion in damages.
- The Court determined that plaintiffs brought frivolous claims and engaged in litigation misconduct by operating in bad faith in pursuing objectively baseless claims that had no reasonable chance of success.
- The Court awarded in excess of $11.6 million in attorneys fees and costs.

• “The third aspect of Defendants' requested fees is $2,829,349.10 attributable to computer assisted, algorithm-driven document review. Defendants provide the following explanation for the resulting fees: ‘Over the course of this litigation, Defendants collected almost 12,000,000 records – mostly in the form of Electronically Stored Information (ESI).... Rather than manually reviewing the huge volume of resultant records, Defendants paid H5 to employ its proprietary technology to sort these records into responsive and non-responsive documents.’ After the algorithm determined whether documents were responsive or unresponsive to discovery requests, Black Letter [contract] attorneys reviewed the responsive documents for confidentiality, privilege, and relevance issues. For this reason, the review performed by H5 and Black Letter accomplished different objectives with the H5 electronic process minimizing the overall work for Black Letter. Again, the Court finds Cooley’s decision to undertake a more efficient and less time-consuming method of document review to be reasonable under the circumstances. In this case, the nature of Plaintiffs' claims resulted in significant discovery and document production, and Cooley seemingly reduced the overall fees and attorney hours required by performing electronic document review at the outset. Thus, the Court finds the requested amount of $2,829,349.10 to be reasonable.”
Defendants used keywords to cull a collection of **19.5 million documents** down to **3.9 million documents**, which were then deduplicated to yield a population of **2.5 million documents** for manual review.

TAR was then applied to the 2.5 million documents containing the keyword “hits” after deduplication.

Plaintiffs challenged the use of keywords before TAR and demanded a do-over using TAR on **all 19.5 million documents**, including the **16 million left behind by the keywords**.

- The Court concluded, on the basis of the statistics reported by defendant, that the validation results suggested “a comparatively modest number of documents would be found,” and, therefore, such a do-over would be disproportionate under Rule 26(b)(2)(c), since the cost of the production to date was $1.07M, and was anticipated to fall between $2M and $3.25M in total.

- But! Defendant had reported precision instead of recall to the Court. Had the evaluation been done correctly, the Court would have learned that the recall on the search terms was on the order of 58%. Since TAR is known to achieve a recall of about 80%, total recall on the end-to-end search and review process was 58% X 80%, or only about 46%.

- Query whether the decision might have come out differently had the Court known that 54% of the relevant documents had been left behind.
Open TAR Issues

- **Is it Reasonable to Continue to Use the “Traditional Approach” When There are Readily Available Alternatives That Have Been Shown to be More Effective and More Efficient?**

- **What Tools and Methods Work Best for What Tasks?**
  - How are parties to know whether the algorithm will work (or did work) as advertised or expected?
    - For example, some tools work better than others on spreadsheets, and OCR’d, short, or foreign-language documents.

- **What is the Best Process to Employ?**
  - Should one train the algorithm using random or judgmental samples?
  - Should one test as they go along, or just at the end of the process?
    - **There is a dearth of empirical evidence to guide parties!**
TREC 2010 Learning Task Results

Topic 401 Recall vs. Documents Reviewed

Recall (%) vs. Thousands of Documents Reviewed
Determining the Effectiveness of TAR Surgery

Considering a surgical procedure?

• How well does the procedure address cases like yours?
• Is your surgeon skilled at the procedure? Does he have state-of-the-art equipment?
• How much will it cost?
• Are there better and/or cheaper alternatives?
• How will you determine if your surgery is a success?

Inform yourself!

• Clinical trial results and scientific literature
• Credentials and track record of the surgeon and the hospital facility
• $Ask
• Compare clinical trial results and cost of the alternatives
• Post-op testing and monitoring
Open TAR Issues (Cont’d)

- **What Level of Transparency is Required?**
  - Must seed or training sets be disclosed?
  - Must documents coded as non-responsive be disclosed?

- **Is There a Bright Line for Acceptable Recall and Margin of Error?**
  - Will TAR be held to a higher standard of validation than other search and review methods?
  - What time and effort (i.e., cost) are justified for the purposes of validation?
    - **If more time and effort is spent on validation, less is spent on actually doing a good search!**
  - How to ensure that the cost of endless disputes does not obliterate the savings achieved by using TAR.
Take-Away Messages

- The “traditional” method (keyword search followed by manual review) is costly, time-consuming, and tends to give poor results; it will become increasingly difficult to defend.

- Measurement and validation is critical regardless of the method applied.

- Choose your vendors and your protocols carefully:
  - Ask lots and lots of questions;
  - Trust, but verify.
Take-Away Messages (Cont’d)

- **Challenge commonly held (and promoted) assumptions that lack independent empirical support:**
  - Seed or training sets must be random
  - Seed or training sets must be reviewed by partners
  - You must review thousands of documents as a prerequisite to employing TAR
  - TAR is more susceptible to reviewer error than the “traditional approach”
  - You should cull with keywords prior to employing TAR
  - Small changes in the seed or training sets can cause large changes in the outcome
  - TAR does not work for short documents, spreadsheets, foreign language documents, or OCR’d documents
  - TAR finds the “easy” documents at the expense of “hot” documents
  - TAR is too expensive or cumbersome for small matters
  - TAR requires a highly skilled operator
  - Skilled searchers can compose complex Boolean searches that are just as good as TAR
  - If you add new custodians you must completely re-train the system
  - If you add new documents to the collection, documents that were previously tagged highly relevant can become non-relevant
  - …
References and Additional Resources


- Conor R. Crowley, *Defending the Use of Analytical Software in Civil Discovery*, Digital Discovery & e-Evidence, 10 DDEE 16, Sept. 16, 2010


References and Additional Resources (Cont’d)

- TREC Legal Track – http://trec-legal.umiacs.umd.edu/

