

Buyer Beware:

How TREC Can Help You Evaluate Your E-Discovery Investments

H5

March 19, 2009



Introductions

Moderator

- **Douglas W. Oard**
Associate Dean for Research and Associate Professor, University of Maryland,
College of Information Studies
Founding Co-Coordinator, TREC Legal Track

Panelists

- **Jason R. Baron**
Director of Litigation, National Archives and Records Administration
Founding Co-Coordinator, TREC Legal Track
- **Maura R. Grossman**
Counsel, Wachtell, Lipton, Rosen & Katz
TREC Legal Track 2008 Topic Authority
- **Dan Brassil**
Manager of Linguistic Technology, H5
TREC 2008 Legal Track participant
- **Venkat Rangan**
CTO, Clearwell Systems
TREC 2008 Legal Track participant

Agenda

- Overview of TREC and the TREC Legal Track
- The importance of the TREC Legal Track
 - Recognition by The Sedona Conference[®] and the Courts
 - Value provided to the legal community
- TREC Legal Track initiatives
 - 2008 Legal Track
 - 2009 Legal Track
 - How to get involved
- Additional resources
- Q&A

Agenda

- Overview of TREC and the TREC Legal Track
- The importance of the TREC Legal Track
 - Recognition by The Sedona Conference[®] and the Courts
 - Value provided to the legal community
- TREC Legal Track Initiatives
 - 2008 Legal Track
 - 2009 Legal Track
 - How to get involved
- Additional resources
- Q&A

TREC

- Text REtrieval Conference of the National Institute of Standards & Technology (NIST)
- Allows evaluation of information retrieval (IR) methods and technologies
- Past participants: Microsoft, Oracle, Apple, etc.

TREC Legal Track

- Established in 2006
- Purpose: to evaluate e-discovery document review and search methodologies

The Interactive Task

Further revisions in 2008

- Purpose: to “model more completely and accurately the conditions and objectives of e-discovery in the real world”
- Introduction of Topic Authority
- Evaluation focused on single topics
 - Relatively large stratified sample
 - Adjudication of appeals by Topic Authority
 - Metrics: recall, precision, F1

TREC Metrics

- *Recall*: the percentage of the relevant documents in a collection that were successfully retrieved
- *Precision*: the percentage of retrieved documents that are in fact relevant
- *F1*: an average of recall and precision

Agenda

- Overview of TREC and the TREC Legal Track
- **The importance of the TREC Legal Track**
 - Recognition by The Sedona Conference® and the Courts
 - Value provided to the legal community
- TREC Legal Track Initiatives
 - 2008 Legal Track
 - 2009 Legal Track
 - How to get involved
- Additional resources
- Q&A

TREC Recognition

- By The Sedona Conference®
- By the Courts:
 - Judge Grimm – *Victor Stanley, Inc., v. Creative Pipe, Inc.*, 250 F.R.D. 251 (D. Md. 2008)
 - Judge Scheindlin – *Securities and Exchange Commission v. Collins & Aikman Corp.*, 2009 WL 94311 (S.D.N.Y. Jan. 13, 2009)

“... we encourage legal community stakeholders, particularly e-discovery service providers, law firms, and corporate counsel to participate in the TREC Legal Track, an international research effort aimed at objectively modeling the e-discovery review process for the purpose of evaluating the efficacy of a wide range of search methodologies.”

http://www.thesedonaconference.org/content/miscFiles/TREC_OPEN_Letter.pdf

(Emphasis added)

Judge Grimm on TREC Legal Track

“**[The Text REtrieval Conference (TREC) LegalTrack]** project can be expected to **identify both cost effective and reliable search and information retrieval methodologies and best practice recommendations**, which, if adhered to, certainly would support an argument that the party **employing them performed a reasonable ESI search ...**”

– U.S. Magistrate Judge Paul W. Grimm
Victor Stanley, Inc., v. Creative Pipe, Inc.

(Emphasis added)

Agenda

- Overview of TREC and the TREC Legal Track
- The importance of the TREC Legal Track
 - Recognition by The Sedona Conference[®] and the Courts
 - **Value provided to the legal community**
- TREC Legal Track Initiatives
 - 2008 Legal Track
 - 2009 Legal Track
 - How to get involved
- Additional resources
- Q&A

For Corporate Counsel and Law Firms

TREC can help practitioners:

- Recognize effective and defensible document review methods and IR techniques
- Objectively assess vendor claims from document review/IR providers
- Make more informed, educated buying decisions
- Navigate meet-and-confer negotiations and negotiations with regulators

For the Judiciary

TREC can serve as a reference point for bringing best practices before the bench.

For Regulatory Agencies

TREC can help identify:

- Methodologies that have attained the most complete and accurate productions
- Effective methodologies for sifting through large document populations

For E-Discovery Providers

TREC can provide:

- Transparent, scientifically valid feedback on e-discovery methods and technologies
- Independent measurement and validation of performance

For E-Discovery Providers (continued)

TREC can help providers develop defensible techniques.

- Technical element:
 - Evaluation of search methodologies to maximize recall and precision for a certain effort/cost
 - Validation of results using sampling
- Human/Process element:
 - Simulation of actual interactive search negotiations
 - Documentation and communication of results for increased transparency

Agenda

- Overview of TREC and the TREC Legal Track
- The importance of the TREC Legal Track
 - Recognition by The Sedona Conference[®] and the Courts
 - Value provided to the legal community
- **TREC Legal Track Initiatives**
 - 2008 Legal Track
 - 2009 Legal Track
 - How to get involved
- Additional resources
- Q&A

TREC 2008 Legal Track—Goals and Objectives

- Develop search technologies to meet the real-world e-discovery needs of lawyers
- Foster development of research communities
- Create benchmark evaluation resources
- Establish baseline results

2008 Interactive Task Participants

4 research teams submitted 7 runs

- Clearwell Systems
- H5
- University at Buffalo
- University of Pittsburgh

TREC 2008 Legal Track—Interactive Task Process

1. Model complaints, including “topics,” are released to public
2. Participants clarify topic definitions with TREC Topic Authorities
3. Participants review the documents for responsiveness
4. Participants deliver results to TREC
5. TREC document reviewers review samples of results
6. Participant results are compared to TREC results
7. Participants appeal any disputed responsiveness calls
8. Topic Authorities make final adjudication
9. Results are made public

Changes to the 2009 Project

- Two principal tasks (instead of three):
 - An interactive task
 - A batch task
- The Enron email corpus

How Can You Get Involved?

- Participate with your tool/methodology
- Encourage your e-discovery providers to participate
- Become a Topic Authority
- Become a Relevance Assessor
- Use the results as a resource

Learn More about TREC

- For more information visit the TREC Legal Track website <http://trec-legal.umiacs.umd.edu/>
- For TREC 2008 Legal Track results, download the overview document <http://trec.nist.gov/pubs/trec17/papers/LEGAL.OVERVIEW08.pdf>
- To be added to the TREC Mailing List, contact oard@umd.edu
- To participate by submitting computer runs in 2009, register at <http://trec.nist.gov/call09.html>
- To become a volunteer assessor in 2009, contact Jason Baron at jason.baron@nara.gov
- To discuss 2008 results, contact H5 and Clearwell at:
 - info@h5.com (www.h5.com)
 - info@clearwellsystems.com (www.clearwellsystems.com)