



Pure & Applied Research: The Good, the Bad, and the Lucky

Peter Jackson
Chief Scientist
The Thomson Corporation

Thomson R&D's Mission Statement

Text: “R&D performs applied research and rapid development to help drive sustainable growth in Thomson companies through technological innovations that meet customer needs”

Subtext:

Generate new revenues!



R&D's *Secret* Mission Statement

Text: “Work on cool stuff with cool people, go to conferences, publish papers, file patents, acquire lots of bragging rights, and celebrate our successes”

Subtext:

Have fun!



A Tale of Two Projects

- History Assistant (1995)
 - Identify judge's ruling in an appellate case, any negative commentary, and the previous cases impacted
 - Solved one hard problem and beat up two others
 - Published 4 papers: 3 conferences (CIKM, ICAIL) and 1 journal article (AIJ)
 - Spent about \$1M over 3 yrs
 - Never operationalized
 - Final assessment
- ResultsPlus (2002)
 - Recommend documents from other databases to people searching case law
 - Reduced problem to one involving search & classification
 - Paper submitted to SIGIR and rejected
 - Spent about \$1M in 6 mo
 - Fully productionized
 - Made over \$2M in first year
 - Now a \$40M revenue stream

The History Problem

- Appellate courts hand down 500 new decisions a day
- These decisions impact previously decided, published cases
 - Prior case only cited about 10% of the time
- Lawyers need to know if a published case has been overruled, remanded, etc.
- Legally qualified editors read new cases, identify ‘history’, i.e.,
 - What is the judge’s ruling in this case?
 - What does the judge say about cases used as precedents?
 - What cases are directly or indirectly affected?

Rojas v. TK Communications, Inc.

87 F.3d 745

C.A.5 (Tex.), 1996.

July 11, 1996. (Approx. 7 pages)

Appeal from the United States District Court for the Western District of Texas.

FACTS

In 1991, Camille Rojas was employed as a disc jockey by TK Communications Inc. ('TK'), which operated KXTN radio station in San Antonio, Texas. During her tenure at the station, Rojas alleges that she was sexually harassed by her supervisor.

PROCEEDINGS

On October 30, 1995, the district court granted TK's motion to dismiss and Tichenor's motion for summary judgment. The court first ruled that Rojas must arbitrate her claims against TK in accordance with the arbitration clause in her employment agreement ...

DISCUSSION

The district court's dismissal of Rojas' claims and grant of summary judgment are subject to de novo review. *Burns-Toole v. Byrne*, 11 F.3d 1270 (5th Cir.1994) (internal citation omitted). A district court's grant of summary judgment is proper when 'there is no genuine issue as to any material fact' and 'the moving party is entitled to judgment as a matter of law.' Fed.R.Civ.P. 56(c)

CONCLUSION

For the foregoing reasons, the judgment of the district court is **AFFIRMED**.

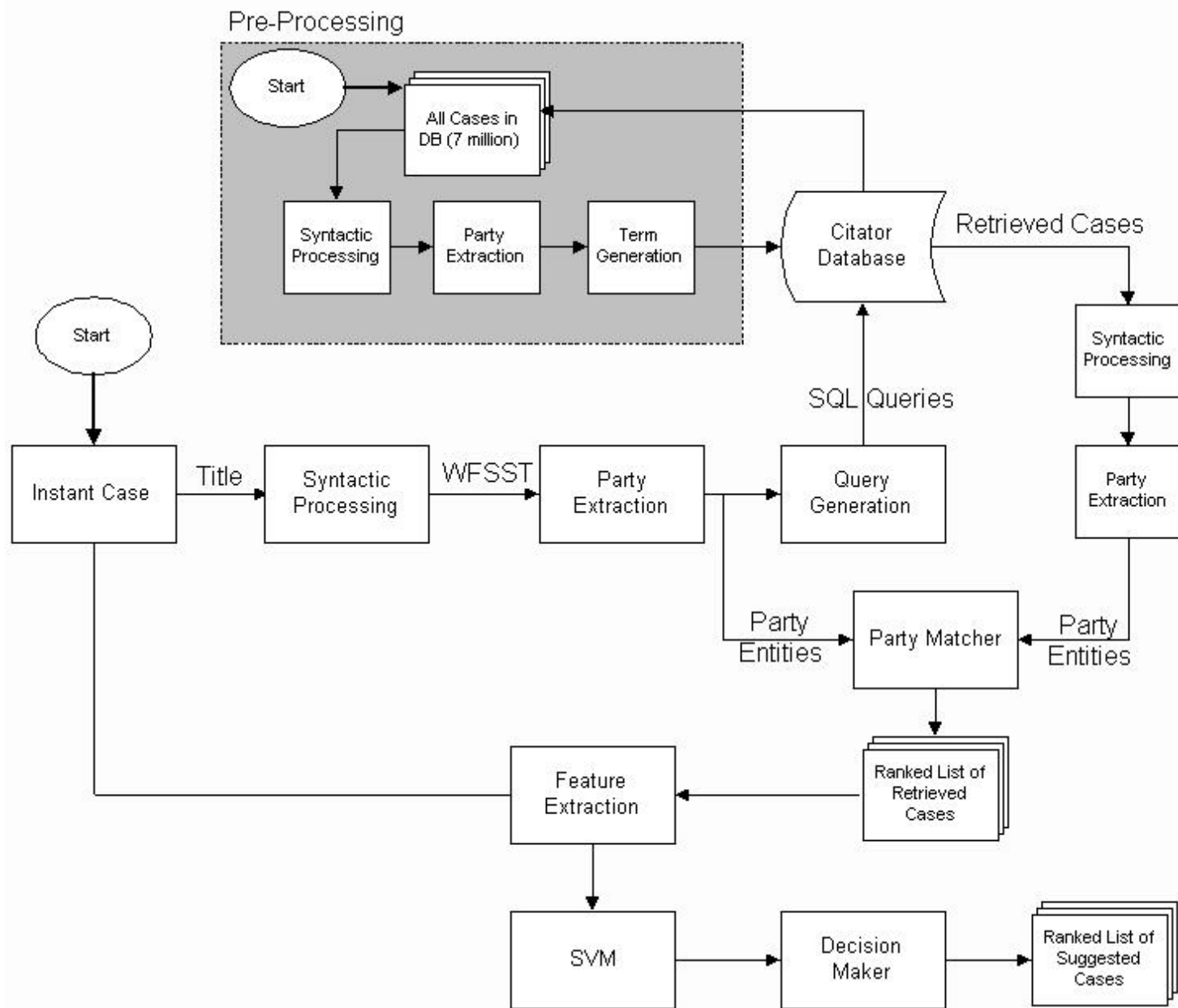
History Assistant: Obstacles to Success

- Changing requirements
 - Began as “help editors annotate one time reload of data”
 - Just help find the dispositive language
 - Moved to “reduce number of editors needed to annotate data”
 - Must find *prior case* automatically at human level of accuracy
- Data availability
 - No corpus when project began
 - Knowledge engineering effort, no statistics
 - Hard to get representative sample of jurisdictions
- Reluctance to innovate
 - Editorial was skeptical, but eventually collaborated
 - Development did not want to “go there”



The History Solution

- Created “diagram” of case by extracting parties, judge, court, date, etc.
 - Useful in looking for prior cases impacted by decision
 - “This case comes before the court from the District Court ...”
- FASTUS-style partial parsing with semantic grammar, for finding dispositive language
 - Used a CYK/chart parser instead of an FSM (7,000 rules!)
 - “I reverse the ruling and remand to the District Ct for a new trial”
- More complex parsing & citation analysis needed for identifying “negative indirect” history
 - Need for anaphora & cataphora resolution
 - “The reasoning in Smith mixes apples and bricks”



History Assistant: Outcome

- State of the art results
 - HA had 78% recall at 54% precision on direct history task
 - C.f. FASTUS: 55% recall at 44% precision, GE: 62% at 53%
 - 78% recall and 89% precision on the treatment history task
 - Anaphora resolution module got 62% correct
 - Prior case retrieval (PCR): 98% recall at 55% precision
 - Recall is at human level, but editorial oversight needed
- Editorial accepted the PCR module for production
 - Never implemented by Development as part of reengineering
- A twist in the tale: We got lucky!
 - Party name matcher used in Westlaw Litigator to link cases & briefs
 - Revenue attributed to linking in 2006: \$25M

The ResultsPlus Problem

- But first ... the problem formulation problem
 - “People aren’t using online analytical materials enough”
 - Real problem: fishing expeditions are replacing traditional research
 - “How can users search all of Westlaw (000s of databases)?”
 - Real problem: how to recover from poor searches with weak results
- Then comes ... the solution formulation problem
 - “We need a computer program that will read the user’s mind”
 - Good luck with that!
- Primary problems obscured by secondary problems
 - Poor formulations, premature solutions often abound

The ResultsPlus Solution

- Recommend analytical materials next to case law searches
 - Use keywords, case classifications, citation patterns, etc.
- Analytical materials already indexed by proprietary document classification process (CaRE)
 - Primarily keyword and word pair (co-occurrence) index
 - Secondarily uses any other taxonomic information available
- Due to limited screen real estate, ranking suggestions is crucial
 - Use click-thru data as one of many features in ranking SVM
 - Run A/B experiments online to guide development

Edit Search | Locate in Result

Search: intoxication /p liability

RESULT- 245 Documents

Result Options

1. **McGilveary v. Baron,**

--- N.Y.S.2d ---, 2004 WL 260992, 2004 N.Y. Slip Op. 00861, N.Y.A.D. 4 Dept., Feb 10, 2004

... 223 Intoxicating Liquors 223XI Civil Damage Laws 223k307 Evidence 223k310 k. Weight and Sufficiency. For purposes of dram shop liability, visible intoxication may be established by circumstantial evidence, including expert and eyewitness testimony. McKinney's Alcoholic Beverage Control Law § 65,...

... 223k310 k. Weight and Sufficiency. Proof of a high blood alcohol count alone generally does not establish the "visible" intoxication required for dram shop liability. McKinney's Alcoholic Beverage Control Law § 65, subd. 2. [3] 228 Judgment 228V On Motion...

2. **Waldon v. Little Flower Children's Service,**

308 A.D.2d 320, 764 N.Y.S.2d 49, 2003 N.Y. Slip Op. 16532, N.Y.A.D. 1 Dept., Sep 04, 2003

...asserted (see D'Amico v. Christie, 71 N.Y.2d 76, 85, 524 N.Y.S.2d 1, 518 N.E.2d 896). There is no evidence of intoxication (see id.) nor any evidence, notwithstanding her mental illness, that the mother was likely to be violent, and, aside from...

...defendant Little Flower in regard to the provision of foster care, and hence, as a matter of law, no liability may be imposed against it for plaintiff's injuries. N.Y.A.D. 1 Dept., 2003. Waldon v. Little Flower Children's Service ...

3. **McGlynn v. St. Andrew the Apostle Church,**

304 A.D.2d 372, 761 N.Y.S.2d 151, 2003 N.Y. Slip Op. 12961, N.Y.A.D. 1 Dept., Apr 10, 2003

...caused by illegal sale of intoxicating liquor; (5) there was no cause of action for injury or damage caused by intoxication of underage person against defendant which did not sell alcohol to alleged assailants; (6) church was not liable for injury...

...which did not sell alcohol to alleged assailants; (6) church was not liable for injury or damage caused by intoxication of underage person; (7) parents who merely knew of underage drinking and did nothing else to encourage it were not...

...of underage drinking and did nothing else to encourage it were not liable for injury or damage caused by intoxication of underage person; (8) those who procured and/or furnished beer for party could be liable for

ResultsPlussm

About

Am.Jur. Proof of Facts

1. Tavern Keeper's Liability Under Dramshop Act

Am.Jur. Trials

2. Dram Shop Litigation

Am.Jur. Trials

3. Liquor Provider Liability

Am.Jur.2d

4. Intoxicating Liquors Grounds or Essentials of Action, Injuries or Damages, Injury to, or death of, purchaser

Am.Jur.2d

5. Intoxicating Liquors Grounds or Essentials of Action, In General, Sale or furnishing of liquor

Am.Jur.2d

6. Intoxicating Liquors Liability under Civil Damage or Dram Shop Act, In General, Generally

Key Number

modify

7. INTOXICATING LIQUORS Persons liable > In general

Key Number

modify

ResultsPlus Algorithms

- Training corpus
 - Analytical text, enriched by case summaries
 - Indexing uses term co-occurrences
- Results fusion
 - Thresholds on CaRE scores for each publication to maintain precision and create global pool of candidate recommendations
 - Initially used ranking SVM based on CTR per publication, etc.
- Personalized ranking
 - Expected CTRs across multiple dimensions
 - User preferences for pub types, jurisdictions by pub type
 - Back off from specific user to law firm to practice area to general user; similarly back off pubs to pub types, etc
 - Statistics continuously updated as new data arrives

ResultsPlus: Use of Ranking SVMs

- Seeks to minimize inverted document pairs
 - Each document represented by a feature vector
 - Learns function such that sorting by value minimizes inversions
- Pair-wise training model
 - Input is labeled pairs of invertible document vectors
 - Actually we subtract one vector from the other for SVM
 - If first document is preferred to second, label is 1, else 0
- Generating the training sample
 - Labels derived from historical clickthrough data
 - Models tested on held out validation set
 - Documents taken from multiple queries & randomized across results
 - Try to avoid bias towards top positions being clicked on

ResultsPlus: Obstacles to Success

- Reluctance to Change
 - “It’s like ads”, “quality will never be good”, etc.
- Reluctance to Take Risk
 - High potential payoff, but some worries that users will be annoyed
- Arguments over methodology
 - Rating recommendations as A, B, C, F
 - Inter-rater agreement problem
 - Relevant v. highly relevant an unstable measure

ResultsPlus: Outcome

- Original experiment
 - Suggest up to 3 articles and 3 key numbers per query
 - 80% of results had at least one A, only 4% all Fs
 - *Current* numbers are > 99% and < 1% respectively
 - ResultsPlus now offers over 300 suggestion types
 - Drawn from a pool of over 1M documents
- Revenue optimization
 - Reranking added \$2.5M in revenues in 2006
- We got lucky again! A \$40M+ business
 - Other companies want their own ResultsPlus
 - Reranking now seen as useful for other products
 - Paper finally published in AAAI Workshop



What Does It All Mean?

- Problem formulation is just as important in applied research as in pure research
- Applied research is just as risky as pure research
- Good experimental methodology is just as essential
 - Evaluation is always hard, but must be done
 - A/B testing with real customers is a huge plus
- Good software engineering is even more essential
- Being aligned with the business is the key to it all
- Leadership is critical: be decisive, be steadfast
 - And be lucky!

The R&D ResultsPlus Team



Chuck Elberti, Peter Jackson, Bill Keenan, Khalid Al-Kofahi, Dave Seaman, Arun Vachher, John Duprey