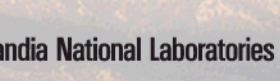
SAND2012-7268P

LET YOUR USERS DO THE RANKING

Ojas Parekh
1465 - Discrete Math and Complex Systems

DATACon 2011

Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin company, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.



Enterprise Analytics at Sandia

Improve the ability of the workforce to find the information they need to perform their jobs, wherever it may reside

From: XXXXXXXX
Sent: Thursday, September 16, 2010 1:40 PM
To: LibSearch
Subject: Help with SearchPoint

My search for "WebCars" came up with this as the first hit:
Sandia Data Storage Service (SDSS): Private vs. Shared
…there is absolutely nothing about WebCars there … ***Why can't SearchPoint ever seem to find anything useful or relevant?*** …

Regards,
XXXXXXXXX

Please let us know if you need assistance with any of the following:
~~-Adding your website to SearchPoint~~
~~-Removing your website from SearchPoint~~
~~-Getting a search box for your website~~
~~-Adding an acronym to SearchPoint~~
~~-Adding a word to SearchPoint's spell checker~~
Replace SearchPoint with something that works

- Complaints from users of SearchPoint, Sandia's enterprise search solution
- Late 2009: Enterprise Analytics project formed and funded to advance algorithmic technology in search applications
- 2010: Embarked on effort to develop and implement a model to improve ranking of search results based on user behavior

Using Users

- Enterprise search is hard; usual suspects don't work:
  "Some customers ask to what extent Google universal search technology relies on PageRank. The answer is: very little."
- Implicit feedback from user search patterns
- Radlinski & Joachims, KDD 2005
- Apply machine learning/optimization to passively collected user feedback

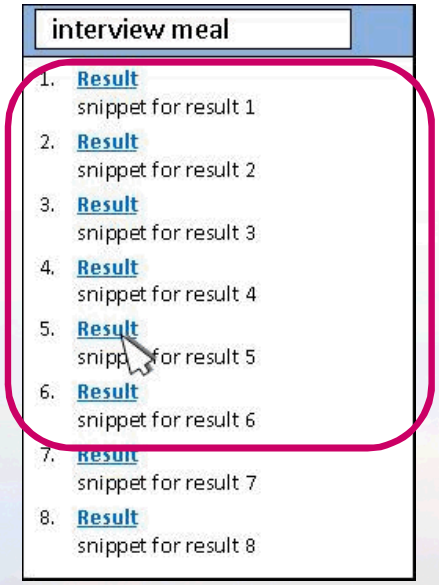
Aggregate Behavior Model

- Collect historical data: What are the queries? Which search results are clicked?
- Apply machine learning/optimization to discover preferred results
- Adjust the ranking of search results for queries based upon user preference
- New data (user behavior) incorporated daily
- Results improve:
  - **Immediately upon implementation, and**
  - **Naturally and automatically over time as a function of user actions**

How do Users Search?

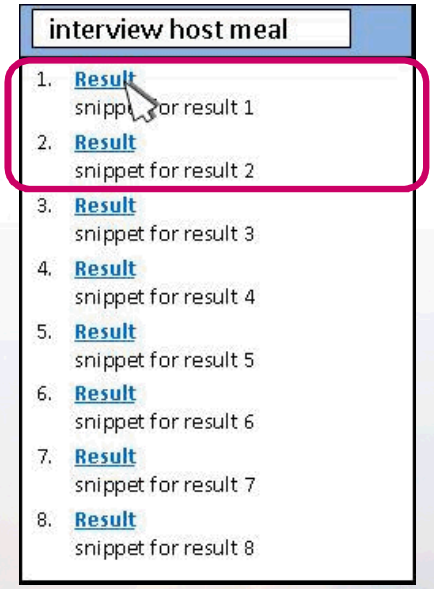
How do I get reimbursed for lunch when interviewing a job candidate?

- Initial query may not be the "right" query
- Results are used (repeatedly) to refine the query
- Generates a query chain



Query: interview meal
User clicked on: #5

Query: interview host
User clicked on: none

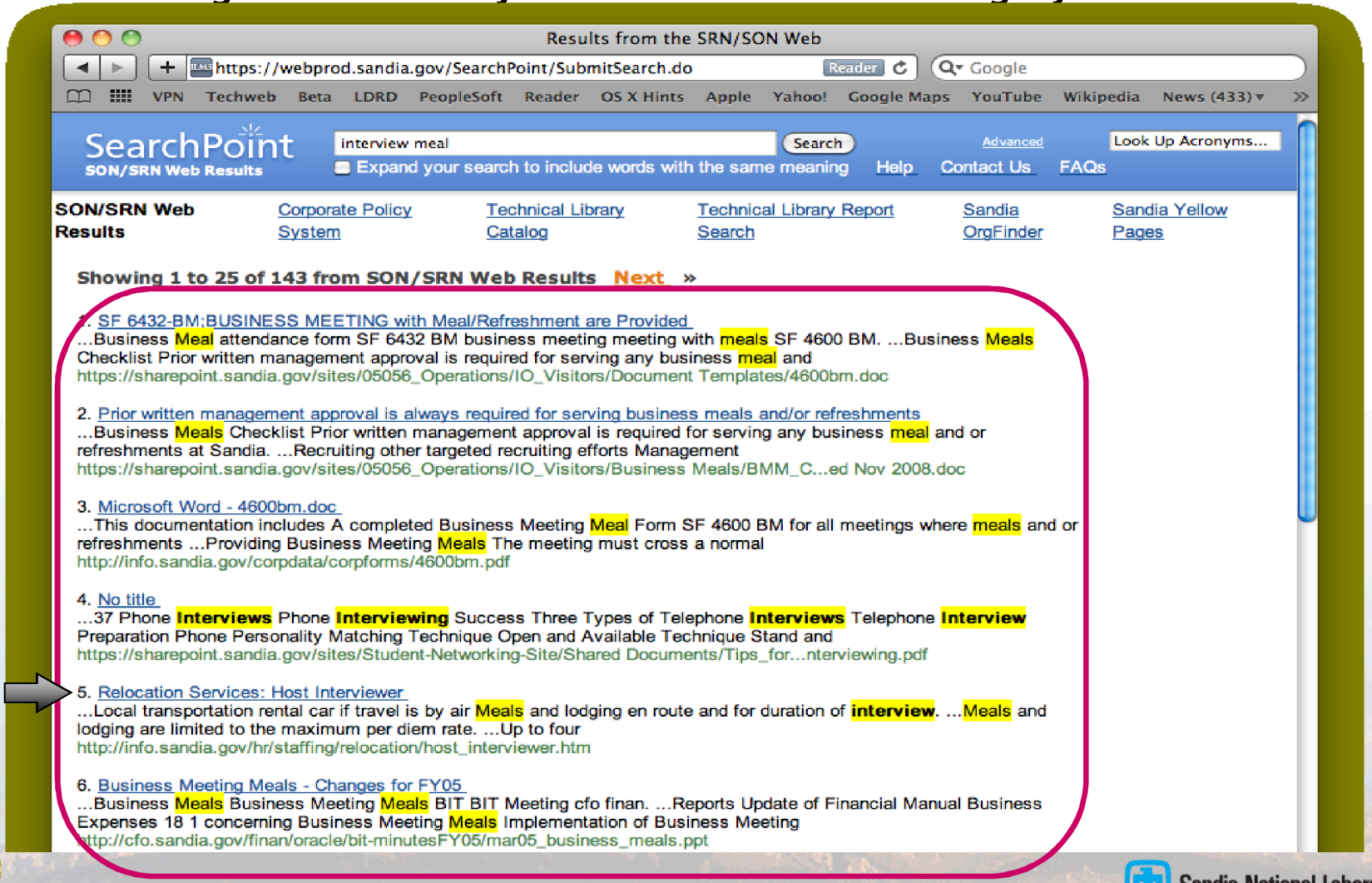


Query: interview host meal
User clicked on: #1

Outlined region: What user looked at (eye tracking studies)

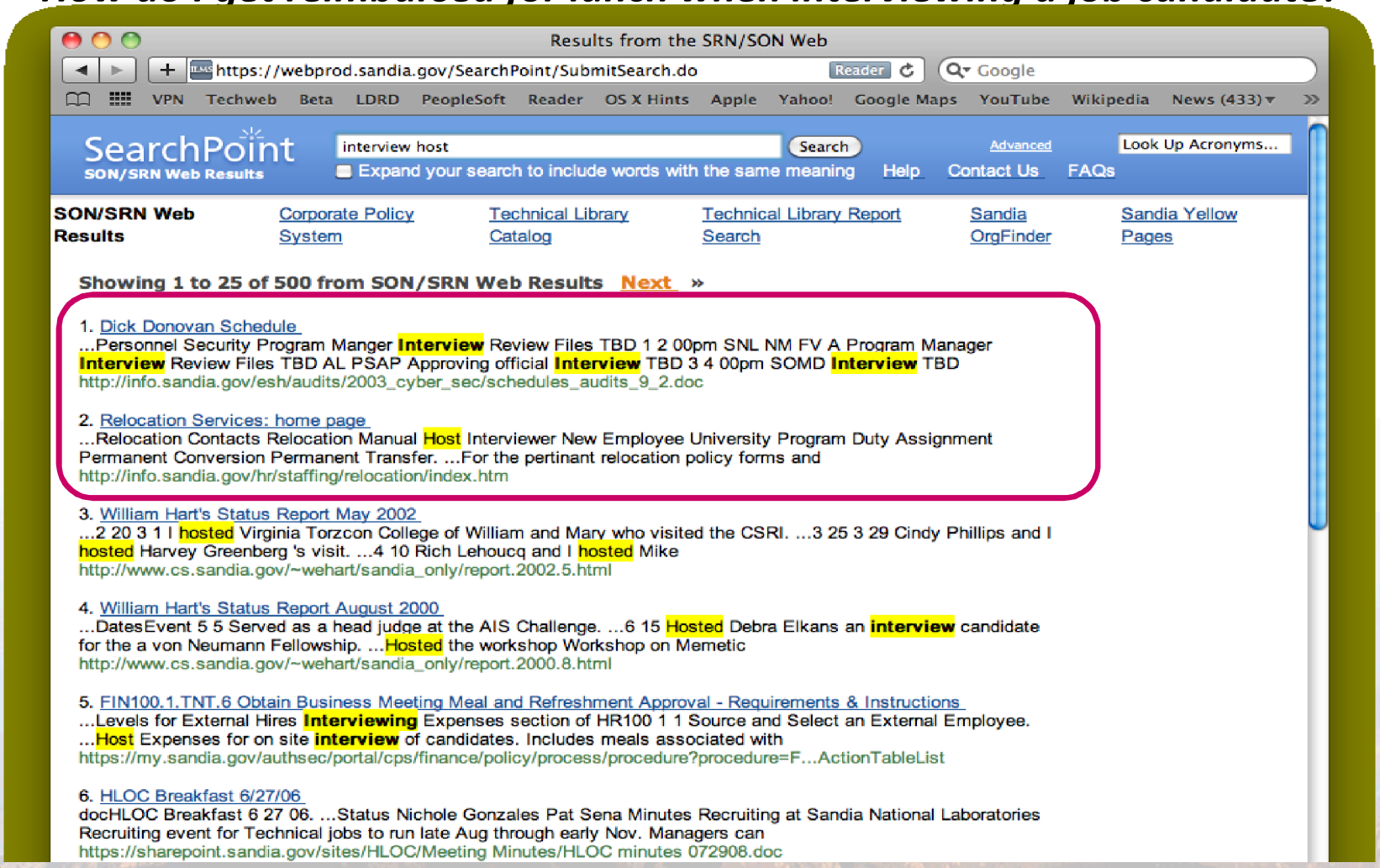
How do Users Search?

How do I get reimbursed for lunch when interviewing a job candidate?



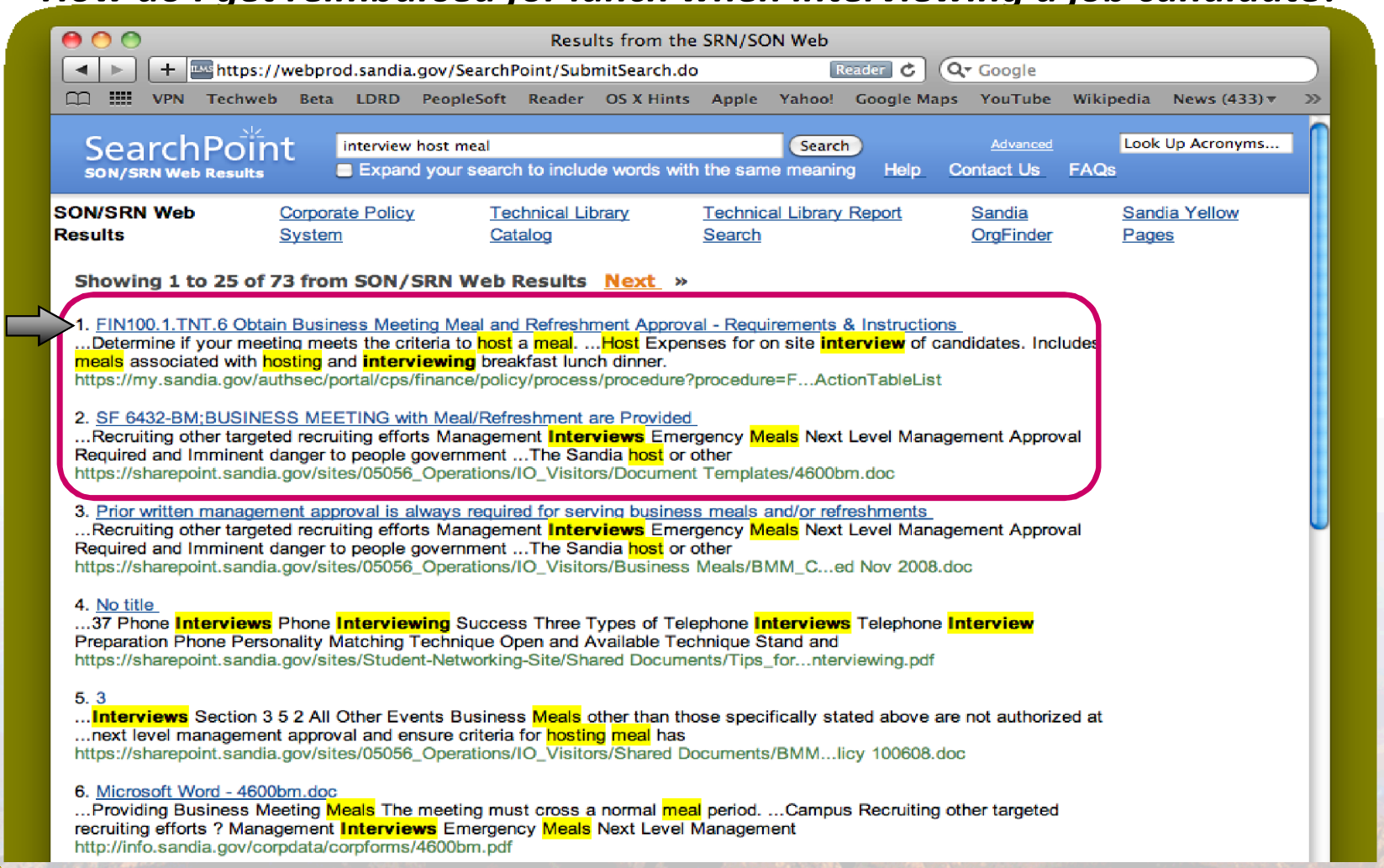
How do Users Search?

How do I get reimbursed for lunch when interviewing a job candidate?

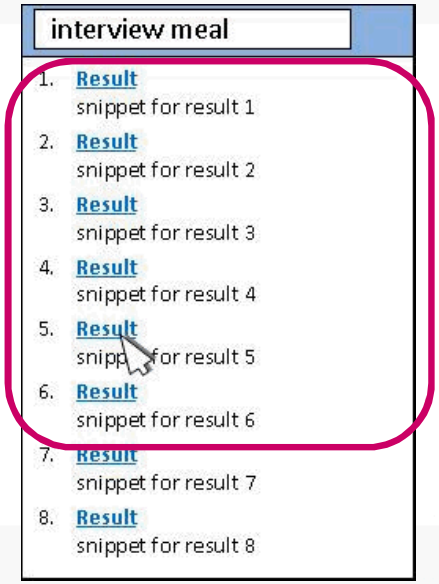


How do Users Search?

How do I get reimbursed for lunch when interviewing a job candidate?



Inferring User Preferences

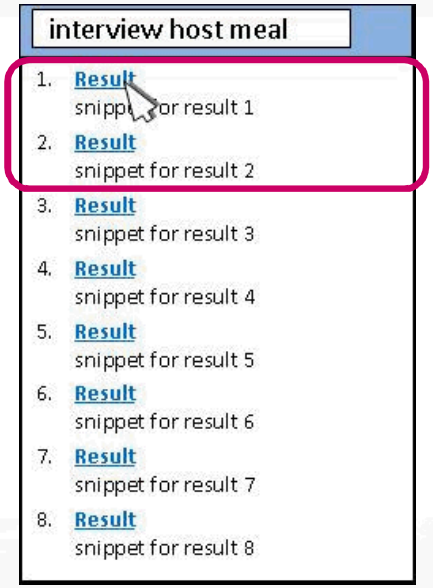


Query: interview meal
User clicked on: #5

Inference:
#5 preferred over #1-4 and #6

Query: interview host
User clicked on: none

Inference:
none



Query: interview host meal
User clicked on: #1

Inference:
#1 preferred over #2

Inference from Entire Query Chain:
#1 (from last query) preferred over #5 (from first query)

Reranking

- Harvest preferences from query chains
- For each query, seek reordering of docs
- Goal: satisfy “important” preferences
  - May be impossible to satisfy all
  - May be undesirable to satisfy all

Support Vector Machine Model

- Model preferences with underlying feature space
- Learn vector w, weighting of feature space
- For each preference $r_i >_q r_j$, impose constraint:

$$\varphi(r_i, q) \bullet w > \varphi(r_j, q) \bullet w$$

vector indicating relationship in feature space (e.g. common terms)

vector weights each feature

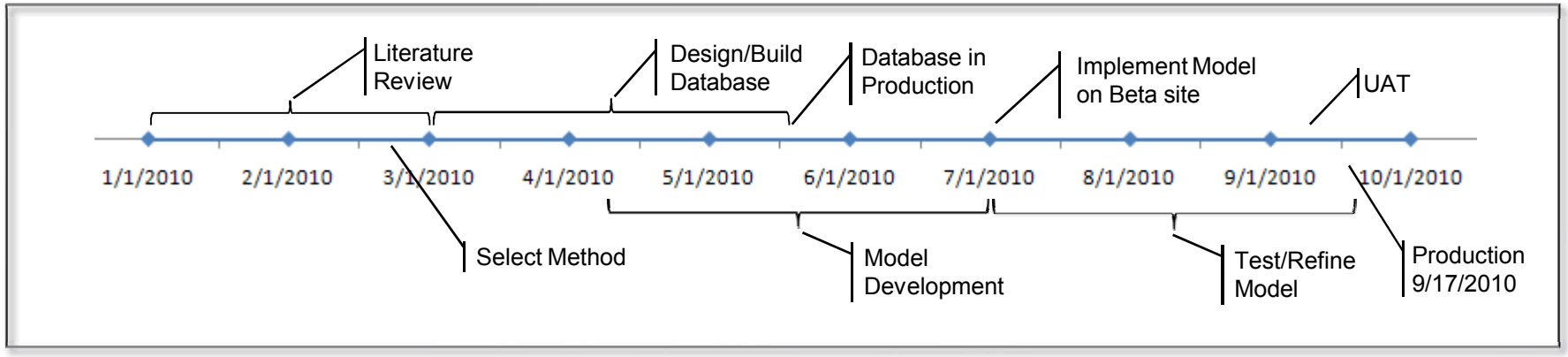
Search Log Data

Search ID
Search Time
Click Time
User ID
Query Text
Clicked URL
URL Position
Number of Results
URL #1
URL #2
•
•
•

- One row of data each time the user clicks on a result
- If the user clicks on 3 URLs after entering a query, there will be 3 rows with differing values in Click Time, Clicked URL, and URL Position; remaining columns will have the same values for each of the 3 rows
- If the user did not click on any of the returned results, there will be one row for which Clicked URL will be null and URL Position will be zero.

Develop & Deploy Model

Model Development Timeline



Re-Ranking Process (nightly):

- Identify query chains
- Harvest preferences
- Learn w using SVM solver
- For each query, rank its results by value $\varphi(r_i, q) \bullet w$

Search Process:

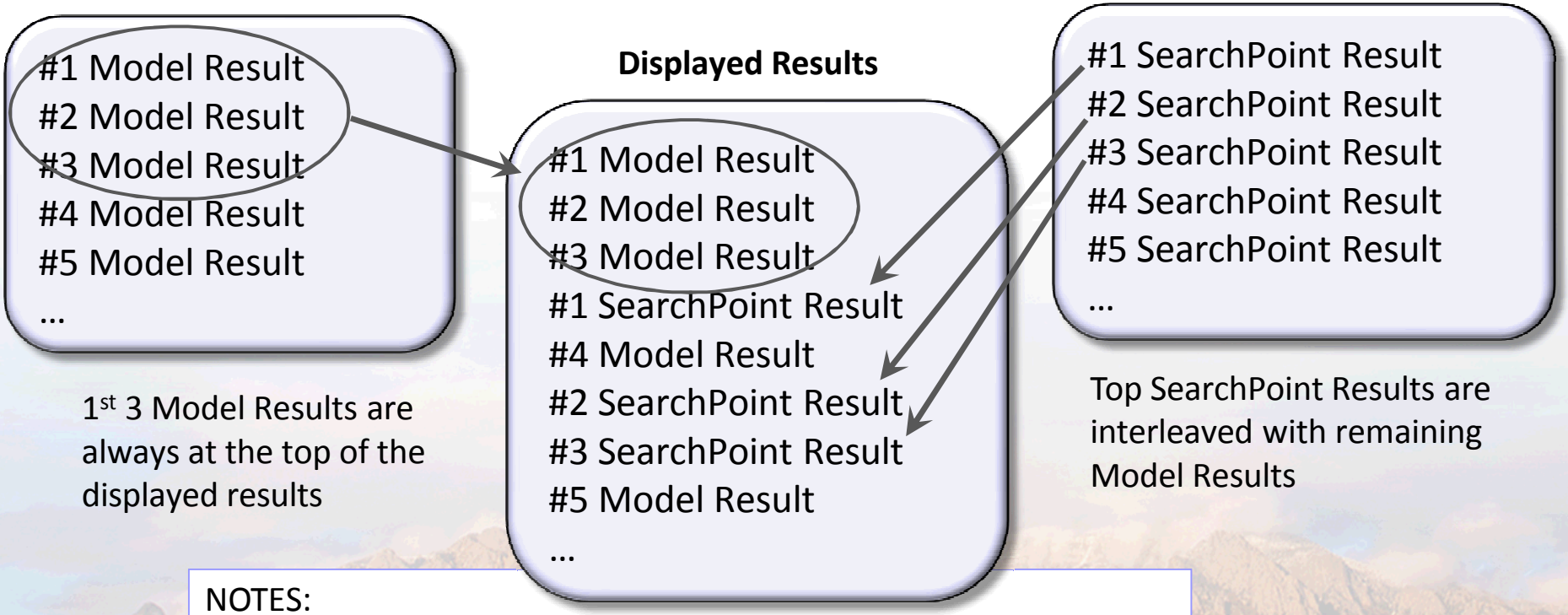
- User executes search query
- Do we have behavior associated with the query?
- Yes: present re-ranked results interleaved w/SearchPoint results
- No: present search engine results

What about "new" search results?
(e.g. new web pages)

FACT: Users rarely look past the 1st page of results

ISSUE: If the 1st page of results contains ONLY results from the model, users will not have the chance to click on new search results

SOLUTION: Discover timing and methods for injecting new results (new web pages, new topics) into the re-ranked results



NOTES:
- If Best Bets exist, they are at the top of the SearchPoint Results.
- Displayed Results are de-duped.

Aggregate Behavior Modeling Example

Query: retirement

Raw Search Engine Results (Retirement 101 is #9)

1. Reimbursement Voucher Form
   snippet for result 1
2. 401(k) Login
   snippet for result 2
3. All Sandia Summary Plan Descr
   snippet for result 3
4. 401(k) Savings Plans
   snippet for result 4
5. Payroll Services
   snippet for result 5
6. Retirement Help
   snippet for result 6
7. Pension Plan & Savings Plan M
   snippet for result 7

Raw Results Re-Ranked by Model (Retirement 101 is #1)

1. Retirement 101 Homepage
   snippet for result 1
2. Retirement Memory Book
   snippet for result 2
3. Sandia 401(k) Savings Plans
   snippet for result 3
4. Benefits Homepage
   snippet for result 4
5. Pension Security Plan (PSP)
   snippet for result 5
6. Retirement Income Plan (RI
   snippet for result 6
7. Sandia National Laboratorie
   snippet for result 7

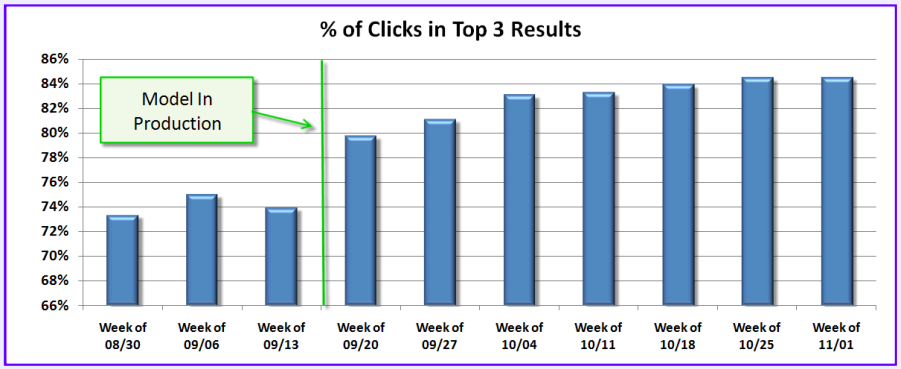
- Just prior to executing this query, Sandia had announced a change to the employee pension plan.
- These 2 results moved higher shortly after the announcement and remained highly ranked for several weeks.
- Now, many months after the announcement, interest has waned, and these 2 results are typically ranking around #20.

NOTE: The Model Results above are "raw results". They do not include injected SearchPoint results that are in final results as displayed to the user.

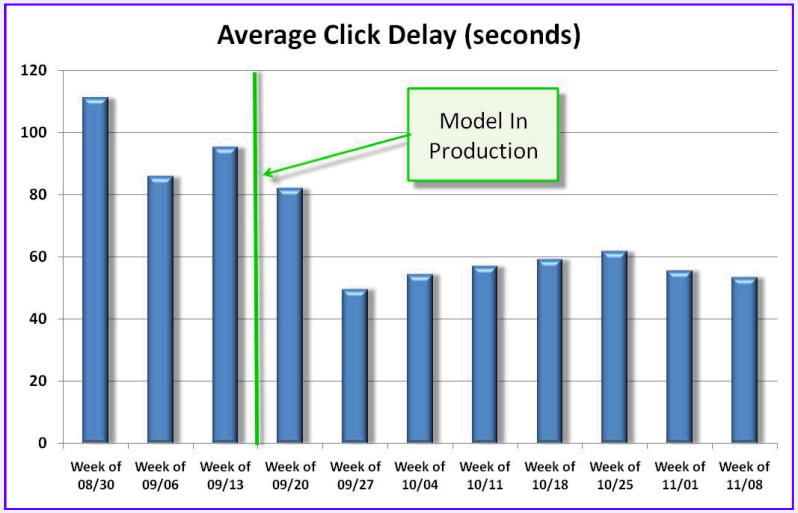
Aggregate Behavior Improvement Metrics

If the model's ranking of search results truly places the more relevant results higher in the result set, we expect:

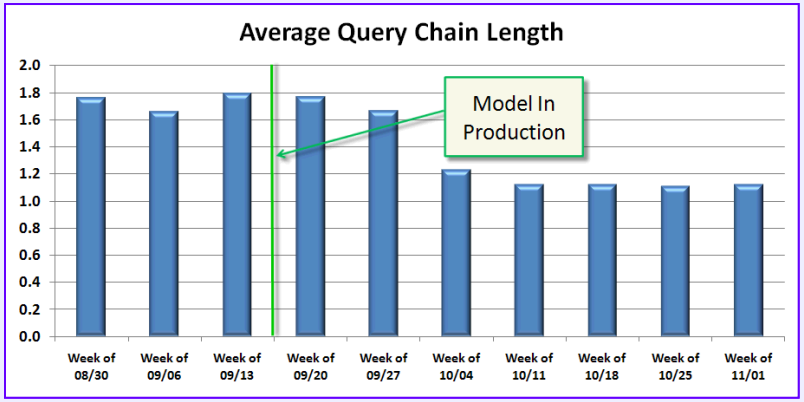
… users will click more frequently on results near the top of the results list



… the user will more quickly find, and click on, results that that the user sees as relevant



… the user will not find it necessary to modify the query terms



And … complaints are now at an all time low

The "grumpy email" at the beginning of this presentation was sent ***one day*** before the Aggregate Behavior Model went into production.

Two business days after the first email, the same gentleman sent this:

From: XXXXXX
Sent: Monday, September 20, 2010 12:20 PM
To: Spomer, Judith E
Cc: LibSearch
Subject: RE: Help with SearchPoint

Greetings all:

It works much better now. And thank you for putting up with my grumpy e-mail.

Best regards,

XXXXXX

What are we really solving?

Alice	Bob	Carlos
1	1	2
2	3	2
3	3	X
4	2	1

?
?
?
?

(Partial) Rank Aggregation

- Given: n items, m (partial) rankings over items
- Optional: a priority/weight for each ranking
- Produce a single consensus ranking optimizing objective
- Many reasonable objectives

Challenges

- Malicious users can skew results
- Do we have a good model for web search?
- Algorithms with a more precise performance guarantee?
- Algorithmic identification of poor ranking?

Questions?