

MAKING THE BEST OF POOR QUERIES

“My husband wants to know all about cancer”

David.Hawking@funnelback.com

II-SDV Conference, Nice, 15 April 2013

IT'S ALL ABOUT THE USER AND THEIR TASK

I-SOV Conference
www.i-sov.com/
 enter the site. I-SOV 2013, Nice, Côte d'Azur, France.

Search, Data Mining, and Visualization - I-SOV Conference
www.i-sov.com/sov/sov_home_page.html
 The I-SOV meeting takes place in Nice in April 2013 for an intensive two days. Venue is the Hotel Plaza in central Nice. The meeting provides an international ...

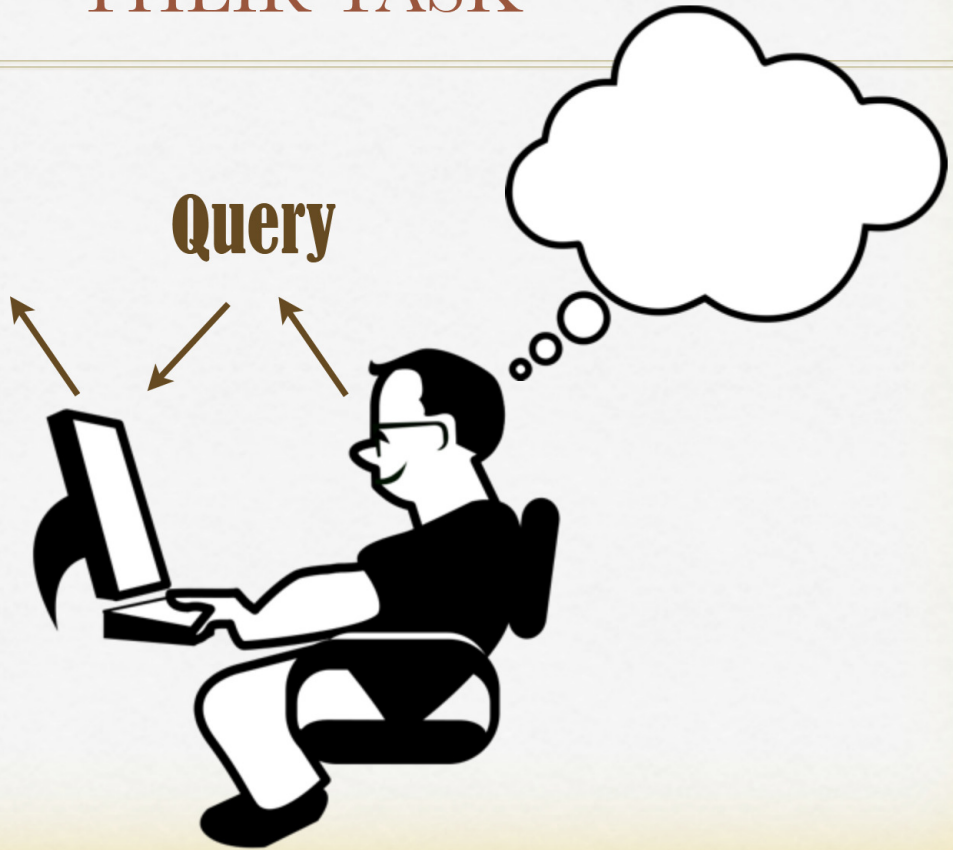
I-SOV 2013 - The Meeting - Dr. Haseef
www.haseef.com/ICIC/
 The International Information Conference on Search, Data Mining and Visualization, The 2013 I-SOV conference will take place in Nice, 15 -18 April 2013 ...

I-SOV 2012 : International Information Conference on Search, Data...
www.wikisp.com/2012/sov/sov.html?eventID=15277
 The I-SOV meeting takes place in Nice in April 2012 for an intensive two days. Venue is the Hotel Plaza in central Nice. The meeting provides an international ...

Programme I-SOV International Conference - Search, Data Mining...
www.i-sov.org/ - I-PLUG Space - PLUG Space - Forum - PLUG-OP
 Apr 22, 2012 - The programme for the April 2012 International Information Conference on Search, Data Mining and Visualization (I-SOV) is now complete and ...

I-SOV
www.conferencesstarts.com/show-event?id=1041040
 International Information Conference on Search, Data Mining and Visualization for information professionals. Of particular interest is patent information ...

Query



IT'S ALL ABOUT THE USER AND THEIR TASK

I-SOV Conference
www.i-sov.com/
 enter the site. I-SOV 2013, Nice, Côte d'Azur, France.

Search, Data Mining, and Visualization - I-SOV Conference
www.i-sov.com/submit_homepage.html
 The I-SOV meeting takes place in Nice in April 2013 for an intensive two days, at the Hotel Plaza in central Nice. The meeting provides an international ...

I-SOV 2013 - The Meeting - Dr. Haseef
www.haseef.com/ICIC/
 The International Information Conference on Search, Data Mining and Visualization 2013 I-SOV conference will take place in Nice, 15 -18 April 2013 ...

I-SOV 2012 : International Information Conference on Search, Data...
www.wikisp.com/2012/04/15/2012-i-sov-conference/
 The I-SOV meeting takes place in Nice in April 2012 for an intensive two days, at the Hotel Plaza in central Nice. The meeting provides an international ...

Programme I-SOV International Conference - Search, Data Mining...
www.i-sov.com/2012/04/15/2012-i-sov-conference/
 Apr 22, 2012 - The programme for the April 2012 International Information Conference on Search, Data Mining and Visualization (I-SOV) is now complete and ...

I-SOV
www.conferencesstarts.com/show-event?i=1041040
 International Information Conference on Search, Data Mining and Visualization for information professionals. Of particular interest to patent information ...

Critically important!

Query



- * Successful search technologies are statistically based.
- * Words in query must match words in document
 - * 'husband'? ... 'wants'? ... 'know' ... 'about' ?
- * Enterprise v. Web
 - * 'Deposit' or 'file' rather than 'publish' or 'promote'
 - * Granularity not designed for answers to queries
 - * SEO?
 - * Popularity / Authority?
- * Important that we run the right query

WHY POOR QUERIES?

- * 1. Misspellings
 - * E.g. 'serch enjin', 'IIS-DV', 'Manchester Untied'
- * 2. Acronyms and abbreviations
 - * E.g. 'DNA' v. 'deoxyribonucleic acid'
- * 3. Regional spellings
 - * E.g. 'color' v. 'colour', 'immunisation v. 'immunization'
- * 4. Compounds
 - * E.g. 'gesundheitswissenschaft', 'airforce'
- * 5. Language mismatch
 - * E.g. 'gun licence' v. 'permit to acquire a longarm'

SOME INFORMATION NEEDS MAKE IT HARD TO CHOOSE A QUERY

- ✱ I need to find descriptions of some projects my consulting company has undertaken where we showed creativity and innovation.
- ✱ My company (well it's not actually incorporated yet) has created and is selling a new form of financial derivative. What advice can I give our Australian customers on their taxation liability?
- ✱ How much does an Australian university receive from the government per student enrolled in a B.Comm?

SOME INFORMATION NEEDS MAKE IT HARD TO CHOOSE A QUERY

- ✱ I need to find descriptions of some projects my consulting company has undertaken where we showed creativity and innovation.
- ✱ My company (well it's not actually incorporated yet) has created and is selling a new form of financial derivative. What advice can I give our Australian customers on their taxation liability?
- ✱ How much does an Australian university receive from the government per student enrolled in a B.Comm?

... and the answer is:



Australian Government
Department of Education, Employment
and Workplace Relations

**Administrative
information for
providers:
Commonwealth Grant
Scheme**

[http://www.innovation.gov.au/HigherEducation/
ResourcesAndPublications/Resources/Documents/
AIPCGS24Dec2010.pdf](http://www.innovation.gov.au/HigherEducation/ResourcesAndPublications/Resources/Documents/AIPCGS24Dec2010.pdf)

THE UNKNOWN INFORMATION SPACE

- ✱ Effective search strategy: Start with a broad query.
- ✱ When I submit such a query,
 - ✱ What is the scale of the information being searched?
 - ✱ What is the scope of the information being searched?
 - ✱ What types of information units are in there? (huge manuals?)
 - ✱ What audiences does the information target?
 - ✱ What sub-topics relate to my initial probe?

CAN SCIENCE HELP?

- * Query completion
- * Query correction
- * Query substitution
- * Query expansion
- * Query shortening
- * Query segmentation
- * Query translation
- * Query blending
- * Query suggestion
- * Facets
- * Result diversification
- * Presentation

CAN SCIENCE HELP?

- * Query completion
- * Query correction
- * Query recommendation
- * Query expansion
- * Query shortening
- * Query segmentation
- * Query translation
- * Query blending
- * Query suggestion
- * Facets
- * Result classification
- * Presentation

ENTERPRISE DEMO

- * Query completion -- helps (a) explore the information space, (b) guide searchers to better queries, and (c) avoid misspellings.
- * Query blending -- running multiple variants of an original query and blending the results from the queries. (Need to prioritise what the user typed.)
- * Query suggestion -- faceting, contextual navigation, “Did you mean?”

OUR CURRENT RESEARCH FOCUS:

BLENDING

WHY BLEND RATHER THAN REPLACE?

- ✱ Always good to humour the user by doing what they asked!
- ✱ Replacement may be too complex.
 - ✱ User may not understand what happened.
 - ✱ It may be difficult to activate implicit phrase or annotation matching
- ✱ Blending can make use of **multiple** transformations independently.

WHAT TO BLEND

- * SPEL – ‘parc de triomphe’ ‘arc de triomphe’
- * DECO – ‘workflow’ ‘work flow’
- * COMP – ‘hoch schule’ ‘hochschule’
- * ACRO – ‘ministry of defence’ ‘MOD’
- * SYNS – ‘gun licence’ ‘permit to acquire a longarm’
- * USUK – ‘color’ ‘colour’
- * RELQ – ‘overseas banking’ ‘IBAN’
- * ...

HOW TO BLEND

- * What variants to run?
 - * Consider candidates from all the suggestion mechanisms.
 - * Assign and accumulate weights.
 - * Rank the candidates.
 - * Decide how many to run and with what weights
- * Guarantee primacy of user-typed query. How?
- * Always give the user the 'plain vanilla' option.

WHEN TO BLEND

- ✱ Whenever there is a viable candidate?
- ✱ Decision based on weights of candidates and assessment of the results from the primary query?

WHEN NOT TO BLEND

- * Try not to make things worse.
- * RELQ
 - * Greater potential to make non-overlapping suggestions
 - * Greater risk
 - * Can't make suggestions for new queries.
- * When in doubt, suggest rather than blend.

CONCLUSIONS

- ✱ Techniques for guiding and improving queries can help a lot, e.g.
 - ✱ Completion
 - ✱ Blending
- ✱ Particularly in enterprise rather than Web search.
- ✱ Challenges remain:
 - ✱ Sparsity of user interaction data
 - ✱ Complexity of data
 - ✱ Vagueness of human needs

But we're getting there!

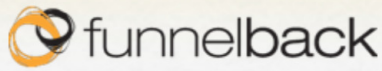




ANY QUERIES?

David.Hawking@funnelback.com

II-SDV Conference, Nice, 15 April 2013



ANY QUERIES?

“Perhaps I can improve
on them!”

David.Hawking@funnelback.com

II-SDV Conference, Nice, 15 April 2013