Modeling User Behavior and Interactions

Lecture 5: **Search Interfaces + New Directions**

Eugene Agichtein

Emory University

Lecture 5 Plan

1. **Generating result summaries (abstracts)**
   – Beyond result list

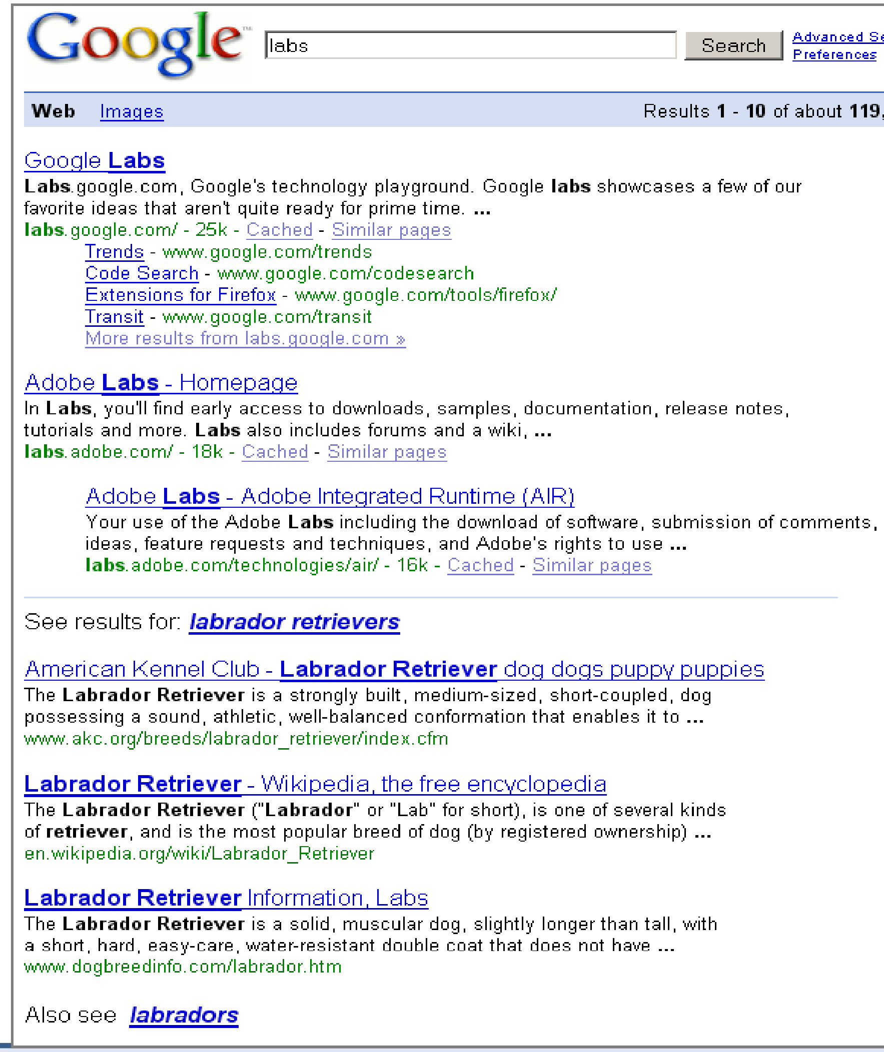
2. **Spelling correction and query suggestion**

3. **New directions in search user interfaces**
   – Collaborative Search
   – Collaborative Question Answering

- **PhD studies in the U.S. (and in Emory U)**

1. Generating Result Summaries

- How to present search results list to a user?
- Most commonly, a list of the document titles plus a short summary, aka “10 blue links”



Good Summary Guidelines

- All query terms should appear in the summary, showing their relationship to the retrieved page
- When query terms are present in the title, they need not be repeated
  - allows snippets that do not contain query terms
- Highlight query terms in URLs
- Snippets should be readable text, not lists of keywords

How to Generate Good Summaries?

- The title is typically automatically extracted from document metadata. What about the summaries?
  - This description is crucial.
  - User can identify good/relevant hits based on description.
- Two main kinds of summaries:
  - **Static summary:** always the same, regardless of the query that hit the doc
  - **Dynamic summary**: *query-dependent* attempt to explain why the document was retrieved for the query at hand

Dynamic Summary Generation

Tropical Fish

One of the U.K.s Leading suppliers of **Tropical**, Coldwater, Marine **Fish** and Invertebrates plus.. . next day **fish** delivery service ...

www.**tropicalfish**.org.uk/**tropical_fish**.htm Cached page

- Query-dependent document summary
- Simple summarization approach
  - rank each sentence in a document using a *significance factor*
  - select the top sentences for the summary
  - first proposed by Luhn in 50’s

Sentence Selection

- Significance factor for a sentence is calculated based on the occurrence of *significant words*
  - If $f_{d,w}$ is the frequency of word *w* in document *d*, then *w* is a significant word if it is not a stopword and

$$f_{d,w} \geq \begin{cases} 7 - 0.1 \times (25 - s_d), & \text{if } s_d < 25 \\ 7, & \text{if } 25 \leq s_d \leq 40 \\ 7 + 0.1 \times (s_d - 40), & \text{otherwise} \end{cases}$$

where s_d is the number of sentences in document *d*

  - text is **bracketed** by significant words (limit on number of non-significant words in bracket)

Sentence Selection

- Significance factor for bracketed text spans is computed by dividing the **square** of the number of **significant words** in the span by the **total number of words**
- e.g.,

```
w   w   w   w   w   w   w   w   w   w   w.
```
(Initial sentence)

```
w   w   s   w   s   s   w   w   s   w   w.
```
(Identify significant words)

```
w   w  [s   w   s   s   w   w   s]  w   w.
```
(Text span bracketed by significant words)

- Significance factor = $4^2/7 = 2.3$

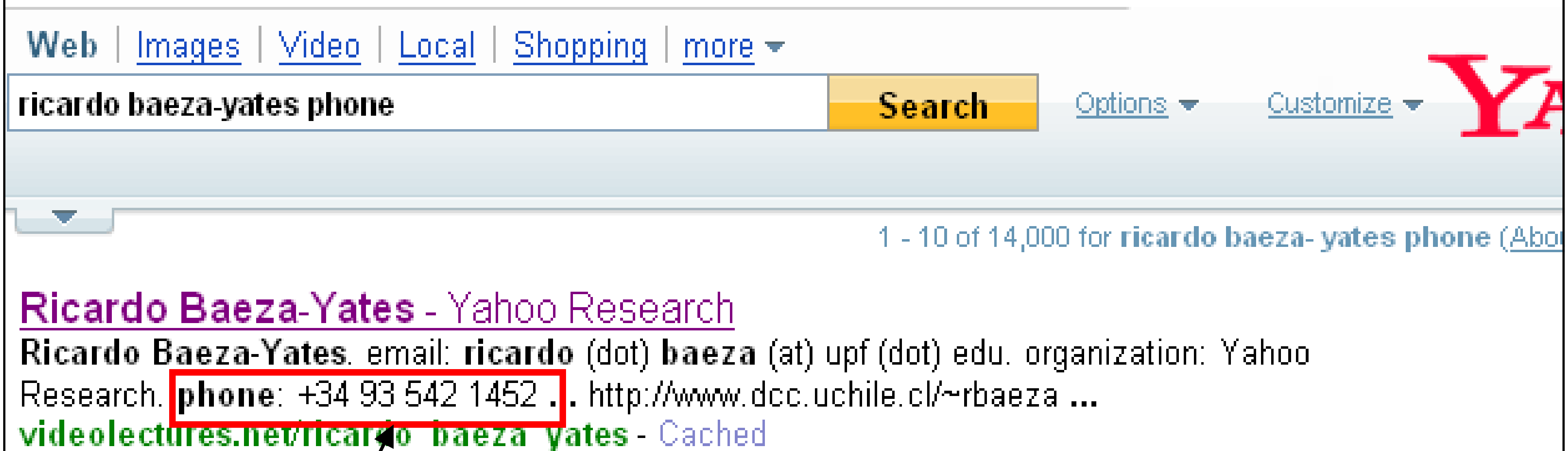
Dynamic Snippet Generation (Cont'd)

- Involves more features than just significance factor
- e.g. for a news story, could use
  - whether the sentence is a heading
  - whether it is the first or second line of the document
  - the total number of query terms occurring in the sentence
  - the number of unique query terms in the sentence
  - the longest contiguous run of query words in the sentence
  - a density measure of query words (significance factor)
- Weighted combination of features used to rank sentences

Static Summary Generation

- Web pages are less structured than news stories
  - can be difficult to find good summary sentences
- Snippet sentences are often selected from other sources
  - metadata associated with the web page
    - e.g., <meta name="description" content= ...>
  - external sources such as web directories
    - e.g., Open Directory Project, http://www.dmoz.org
  - Wikipedia: summary paragraph, infoboxes, …

Problem? Very Good Summaries May Not Get **Clicks**!

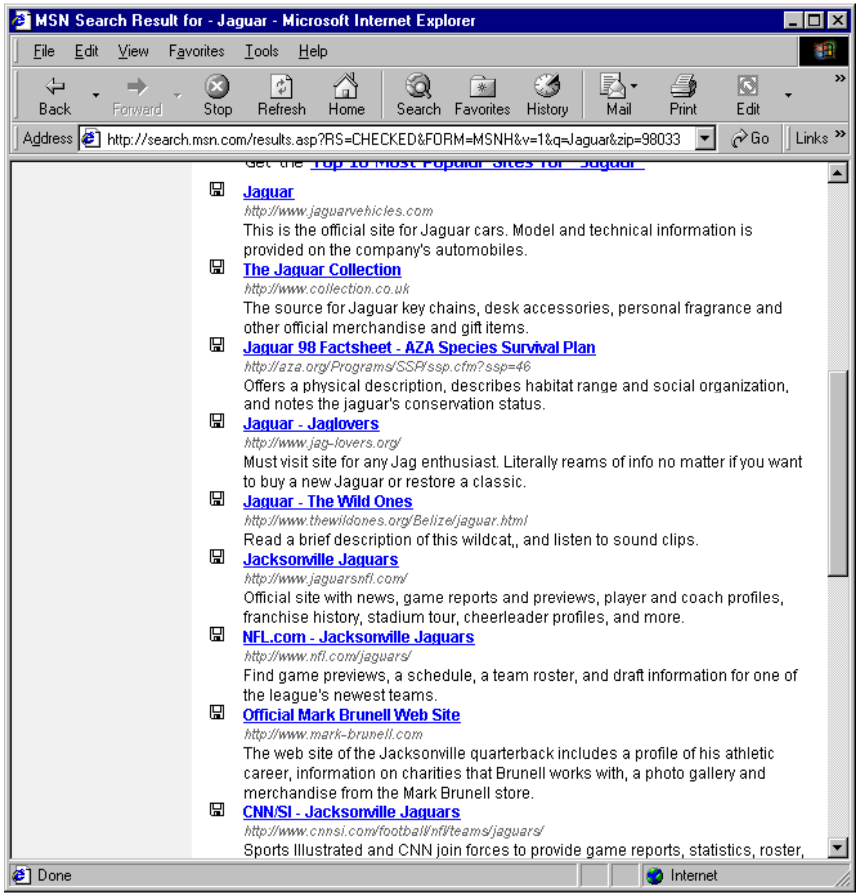


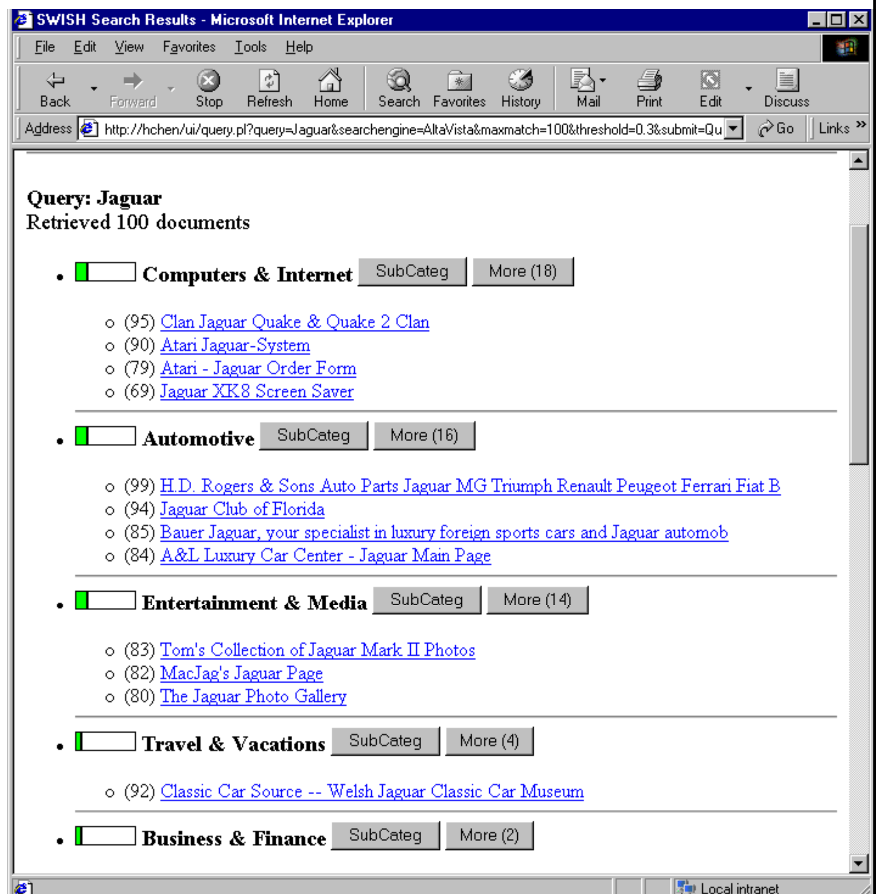
Everything you needed is in the summary

Organizing Search Results

Dumais, S, E. Cutrell, and H. Chen. *Optimizing search by showing results in context*, CHI 2001

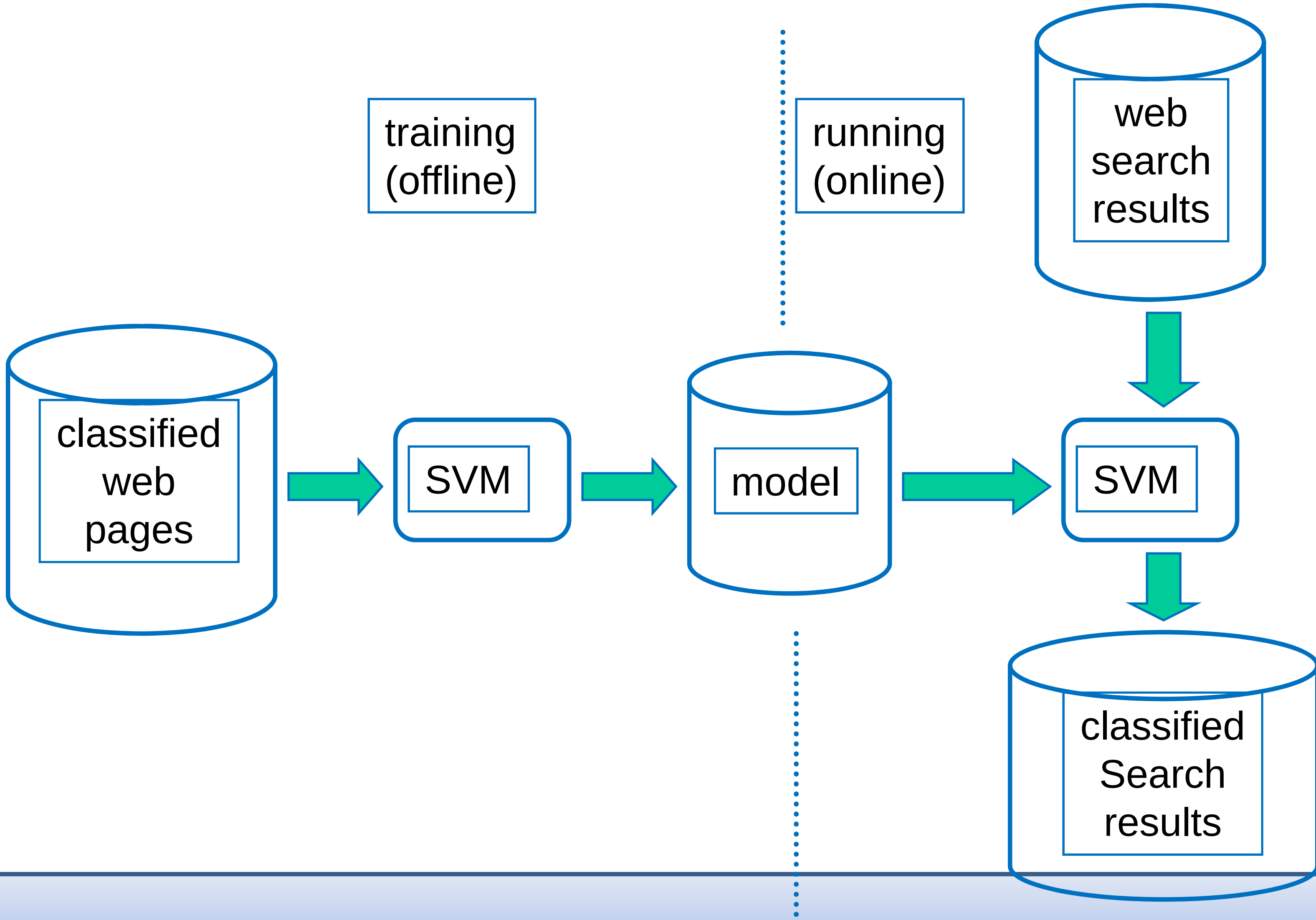
List Organization   Query: `jaguar`   Category Org (SWISH)





System Components

Dumais, S, E. Cutrell, and H. Chen. *Optimizing search by showing results in context*, CHI 2001



Text Classification

Dumais, S, E. Cutrell, and H. Chen. *Optimizing search by showing results in context*, CHI 2001

- Text Classification
  - Assign documents to one or more of a predefined set of categories
  - E.g., News feeds, Email - spam/no-spam, Web data
  - Manually vs. automatically
- Inductive Learning for Classification
  - Training set: Manually classified a set of documents
  - Learning: Learn classification models
  - Classification: Use the model to automatically classify new documents

Learning & Classification

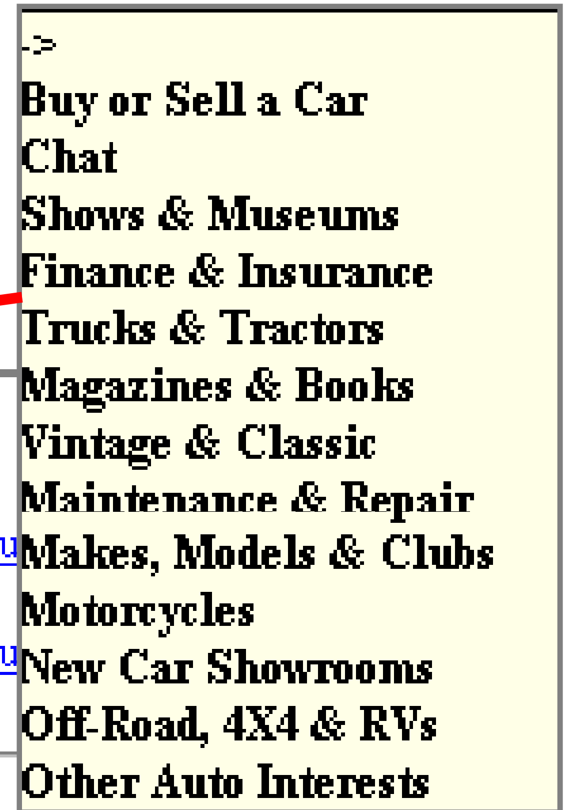
Dumais, S, E. Cutrell, and H. Chen. *Optimizing search by showing results in context*, CHI 2001

- Support Vector Machine (SVM)
  - Accurate and efficient for text classification (Dumais et al., Joachims)
  - Model = weighted vector of words
    - "Automobile" = motorcycle, vehicle, parts, automobile, harley, car, auto, honda, porsche …
    - "Computers & Internet" = rfc, software, provider, windows, user, users, pc, hosting, os, downloads ...
- Hierarchical Models
  - 1 model for N top level categories
  - N models for second level categories
  - Very useful in conjunction w/ user interaction

Information Overlay

Dumais, S, E. Cutrell, and H. Chen. *Optimizing search by showing results in context*, CHI 2001

– Use tooltips to show
  - Summaries of web pages
  - Category hierarchy



Expansion of Category Structure

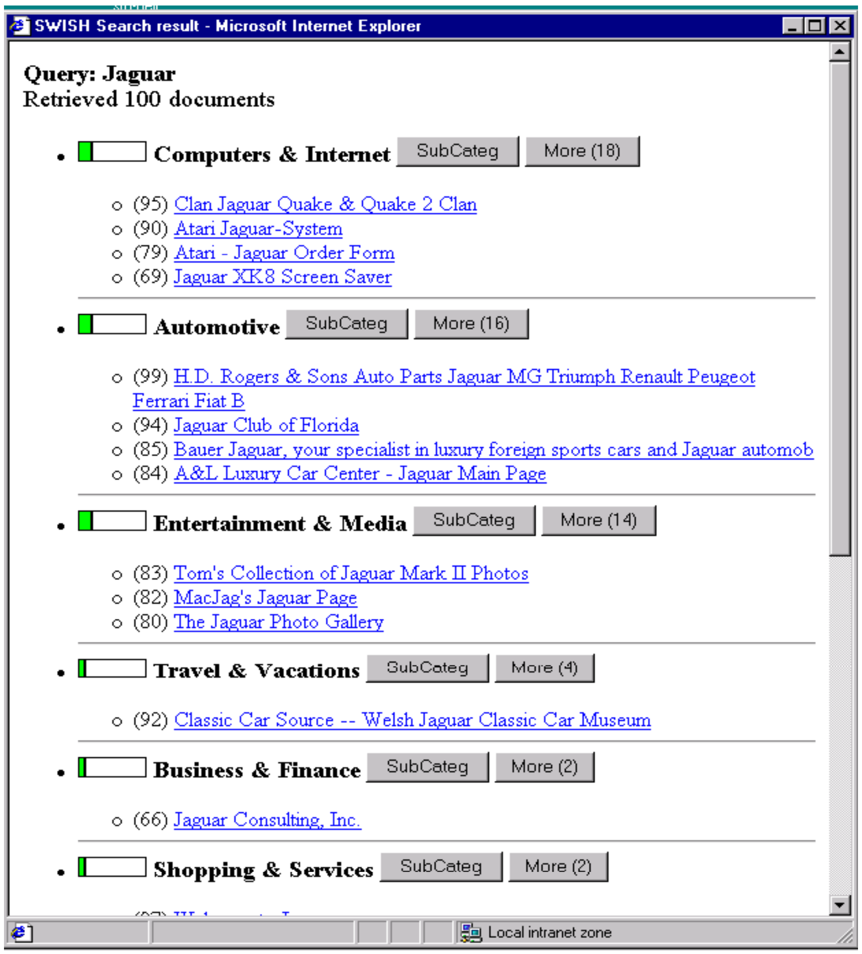
Dumais, S, E. Cutrell, and H. Chen. *Optimizing search by showing results in context*, CHI 2001

- **Automotive** MainCateg
  - **Maintenance & Repair** More (7)
    - (99) H.D. Rogers & Sons Auto Parts Jaguar MG Triumph Renault Peugeot Ferrari Fiat B
    - (85) Bauer Jaguar, your specialist in luxury foreign sports cars and Jaguar automob
  - **Buy or Sell a Car** More (6)
    - (85) Bauer Jaguar, your specialist in luxury foreign sports cars and Jaguar automob
    - (84) A&L Luxury Car Center - Jaguar Main Page
  - **Vintage & Classic** More (2)
    - (99) H.D. Rogers & Sons Auto Parts Jaguar MG Triumph Renault Peugeot Ferrari Fiat B
  - **Makes, Models & Clubs** More (1)
    - (94) Jaguar Club of Florida

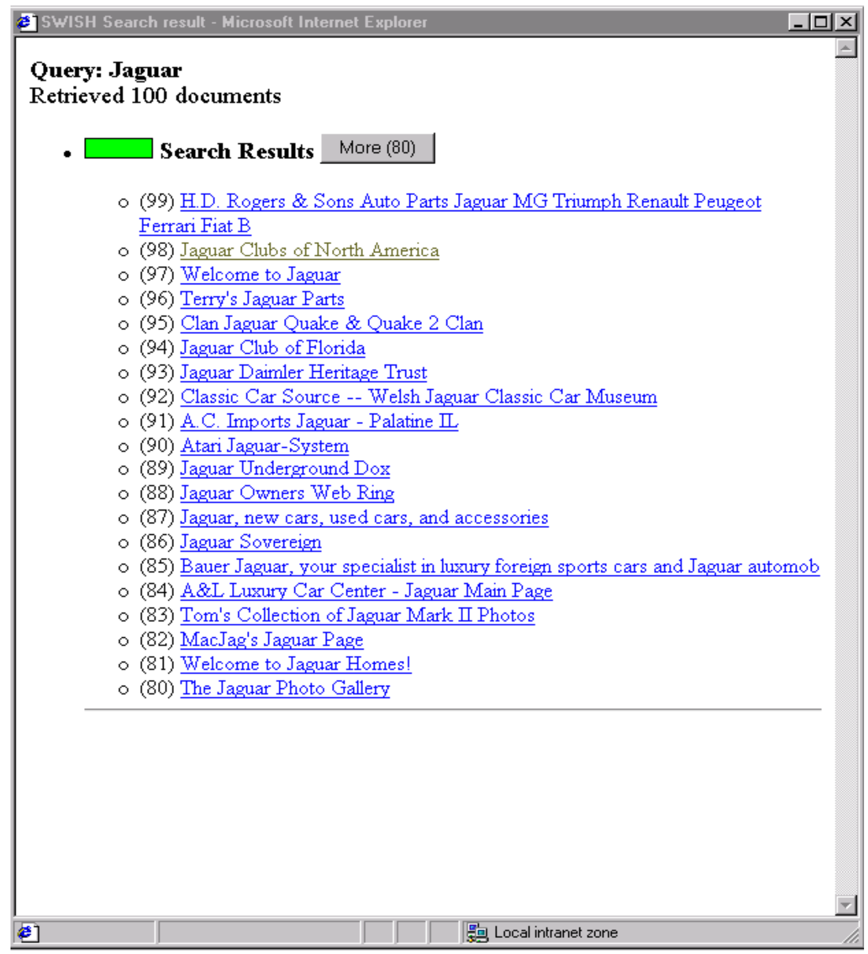
User Study - Conditions

Dumais, S, E. Cutrell, and H. Chen. *Optimizing search by showing results in context*, CHI 2001

Category Interface



List Interface



User Study

Dumais, S, E. Cutrell, and H. Chen. *Optimizing search by showing results in context*, CHI 2001

Subjective Results

Dumais, S, E. Cutrell, and H. Chen. *Optimizing search by showing results in context*, CHI 2001

7-point rating scale (1=disagree; 7=agree)

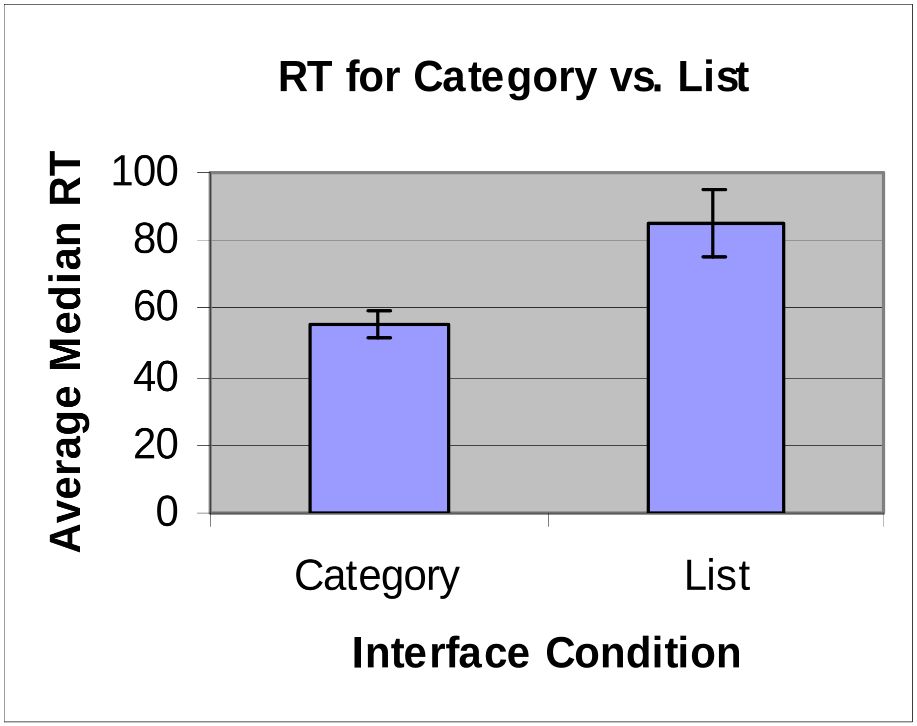
Question	Category	List	significance
It was easy to use this software.	6.4	3.9	p<.001
I liked using this software	6.7	4.3	p<.001
I prefer this to my usual Web Search engine	6.4	4.3	p<.001
It was easy to get a good sense of the range of alternatives	6.4	4.2	p<.001
I was confident that I could find information if it was there.	6.3	4.4	p<.001
The "More" button was useful	6.5	6.1	n.s.
The display of summaries was useful	6.5	6.4	n.s.

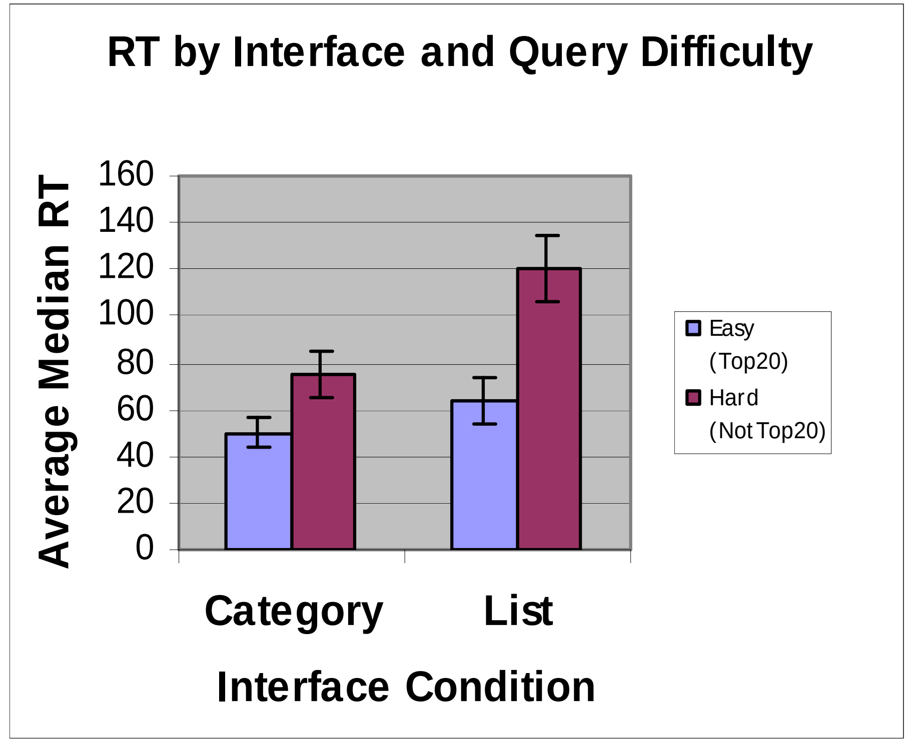
Average Number of Uses of Feature per Task

Interface Features	Category	List	significance
Expansing / Collapsing Structure	0.78	0.48	p<.003
Viewing Summaries in Tooltips	2.99	4.60	p<.001
Viewing Web Pages	1.23	1.41	p<.053

Results: Search Time

Dumais, S, E. Cutrell, and H. Chen. *Optimizing search by showing results in context*, CHI 2001





Category: 56 secs
List: 85 secs $p < .002$

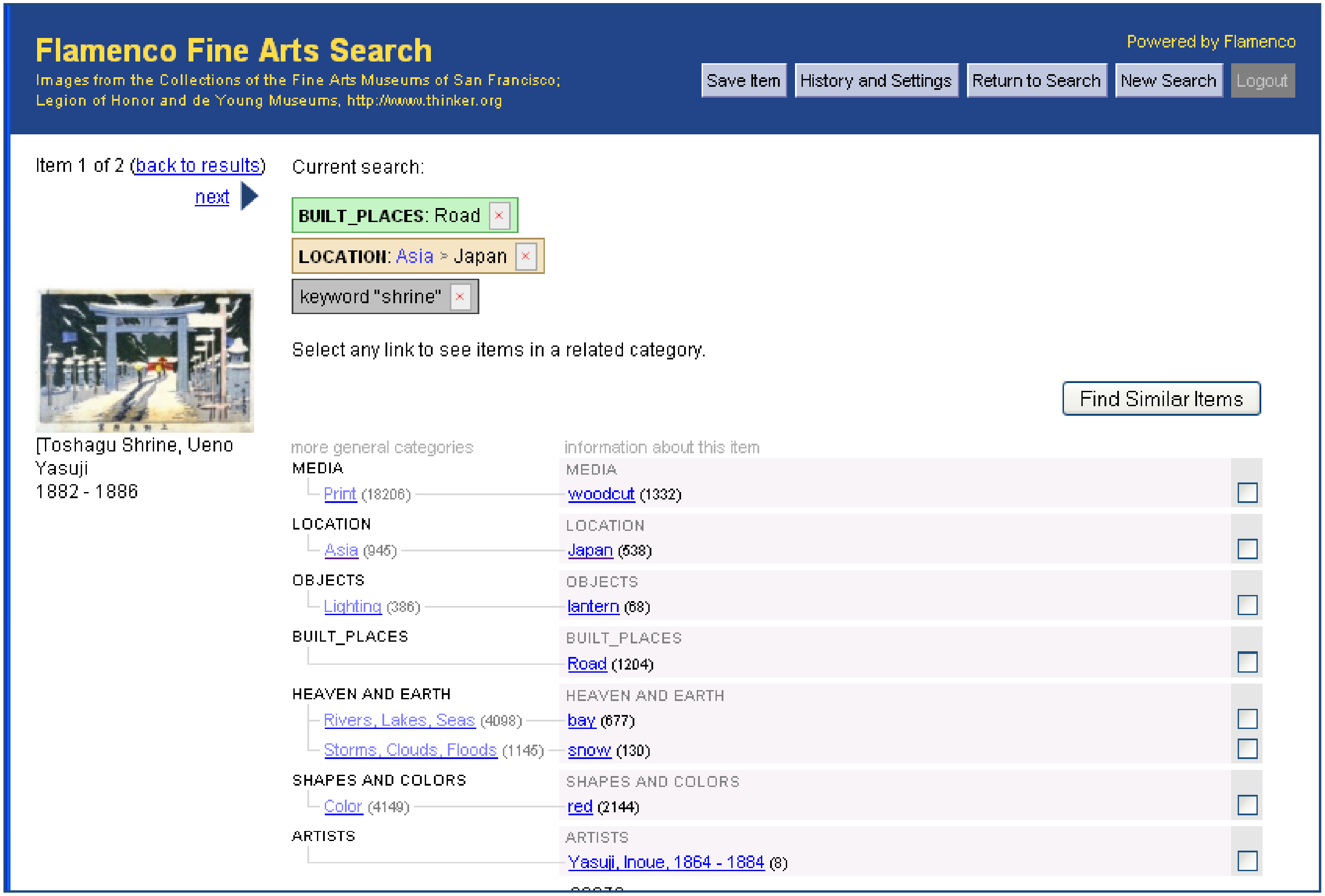
50% faster with Category interface

Top20: 57 secs
NotTop20: 98 secs

- No reliable interaction between query difficulty and interface condition
- Category interface is helpful for both easy and difficult queries

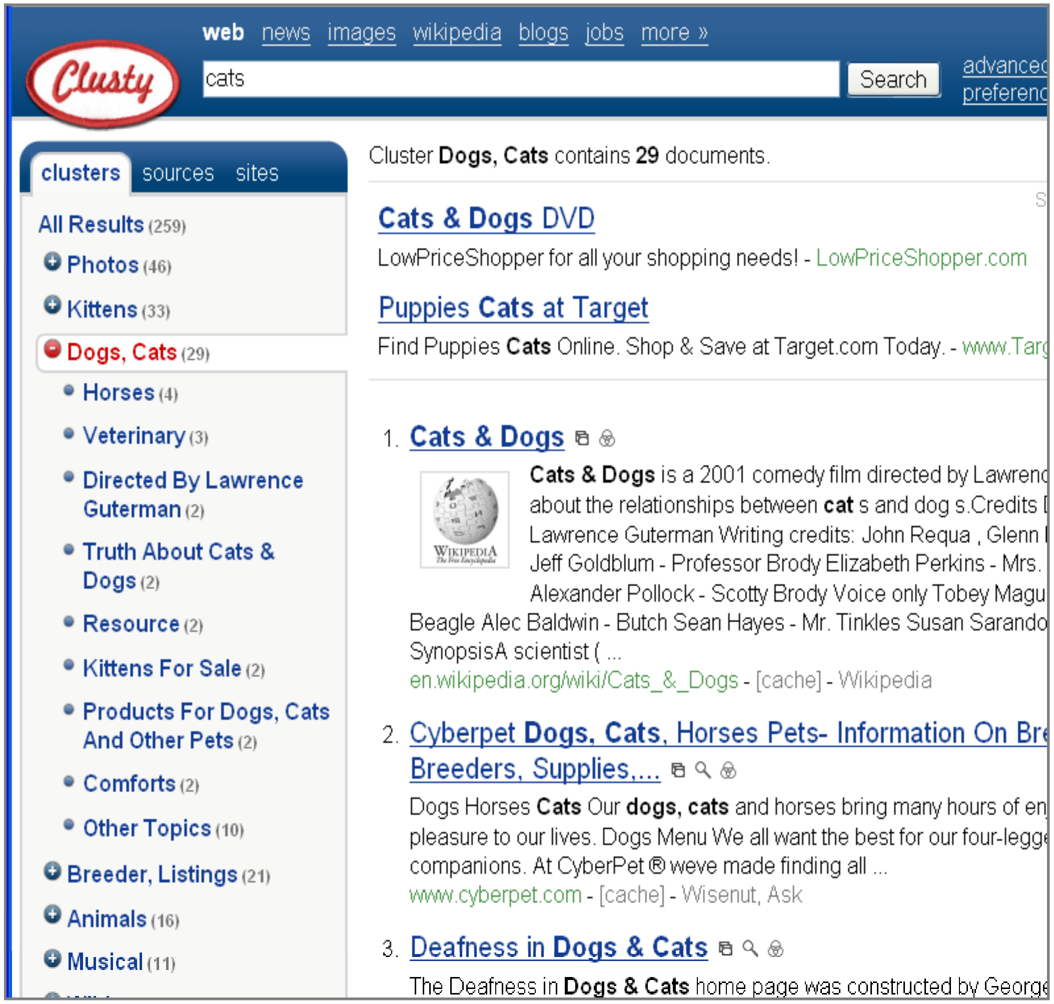
Faceted Navigation (Flamenco)

Marti Hearst, SUI 2009



Clustering Search Results

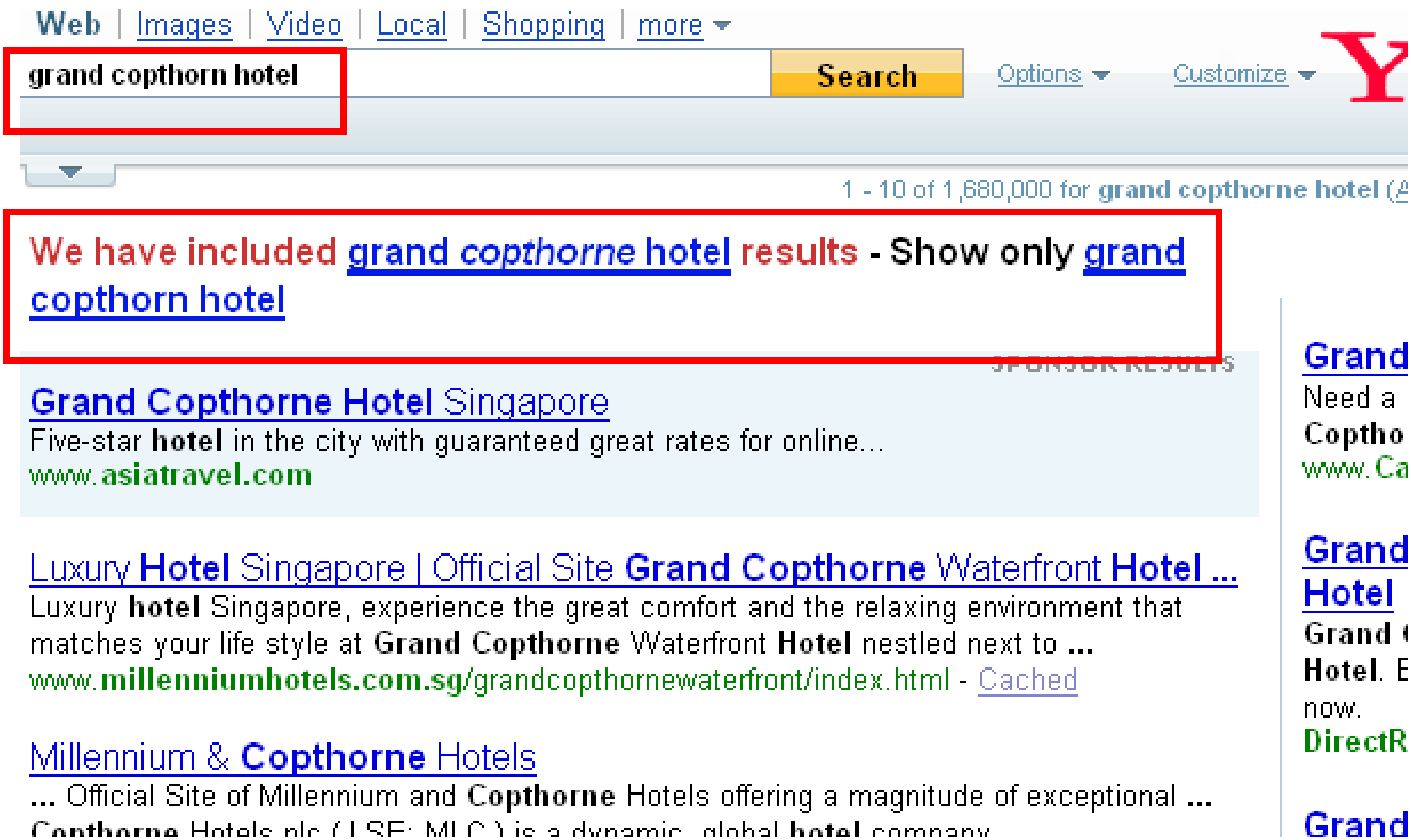
Marti Hearst, SUI 2009



Lecture 5 Plan

- ✓ **Generating result summaries (abstracts)**
  - ✓ Beyond result list
- ➢ **Spelling correction and query suggestion**
- **New directions in search user interfaces**
  - Collaborative Search
  - Collaborative Question Answering
- **PhD studies in the U.S.**

Query Spelling Correction



Reformulations from Bad to Good Spellings

Type	Example	%
non-rewrite	mic amps -> create taxi	53.2%
insertions	game codes -> video game codes	9.1%
substitutions	john wayne bust -> john wayne statue	8.7%
deletions	skateboarding pics → skateboarding	5.0%
spell correction	real eastate -> real estate	7.0%
mixture	huston's restaurant -> houston's	6.2%
specialization	jobs -> marine employment	4.6%
generalization	gm reabtes -> show me all the current auto rebates	3.2%
other	thansgiving -> dia de acconde gracias	2.4%

[Jones & Fain, 2003]

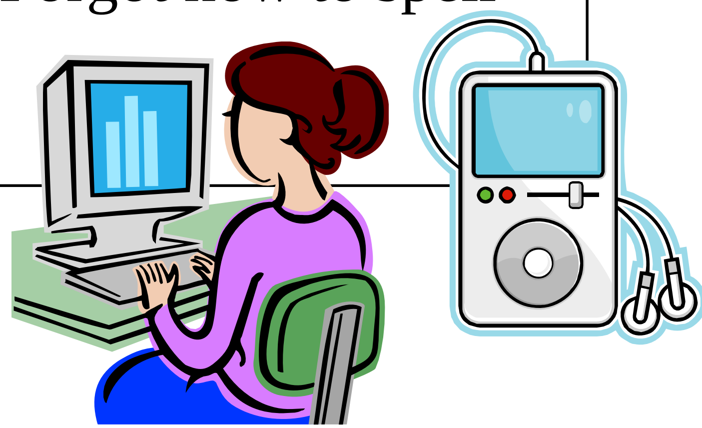
Spelling Correction: Noisy Channel Model

Platonic concept
of query

Correct Spelling

Typing quickly
Distracted

Forgot how to spell



Typos/spelling errors

Reconstruct original query by "reversing this process"

Modeling Errors

$$P(q_{correct} \mid q_{error}) = p(q_{error} \mid q_{correct})\, p(q_{correct})$$

Error model

Character level: p(m|n) p(s|z) etc

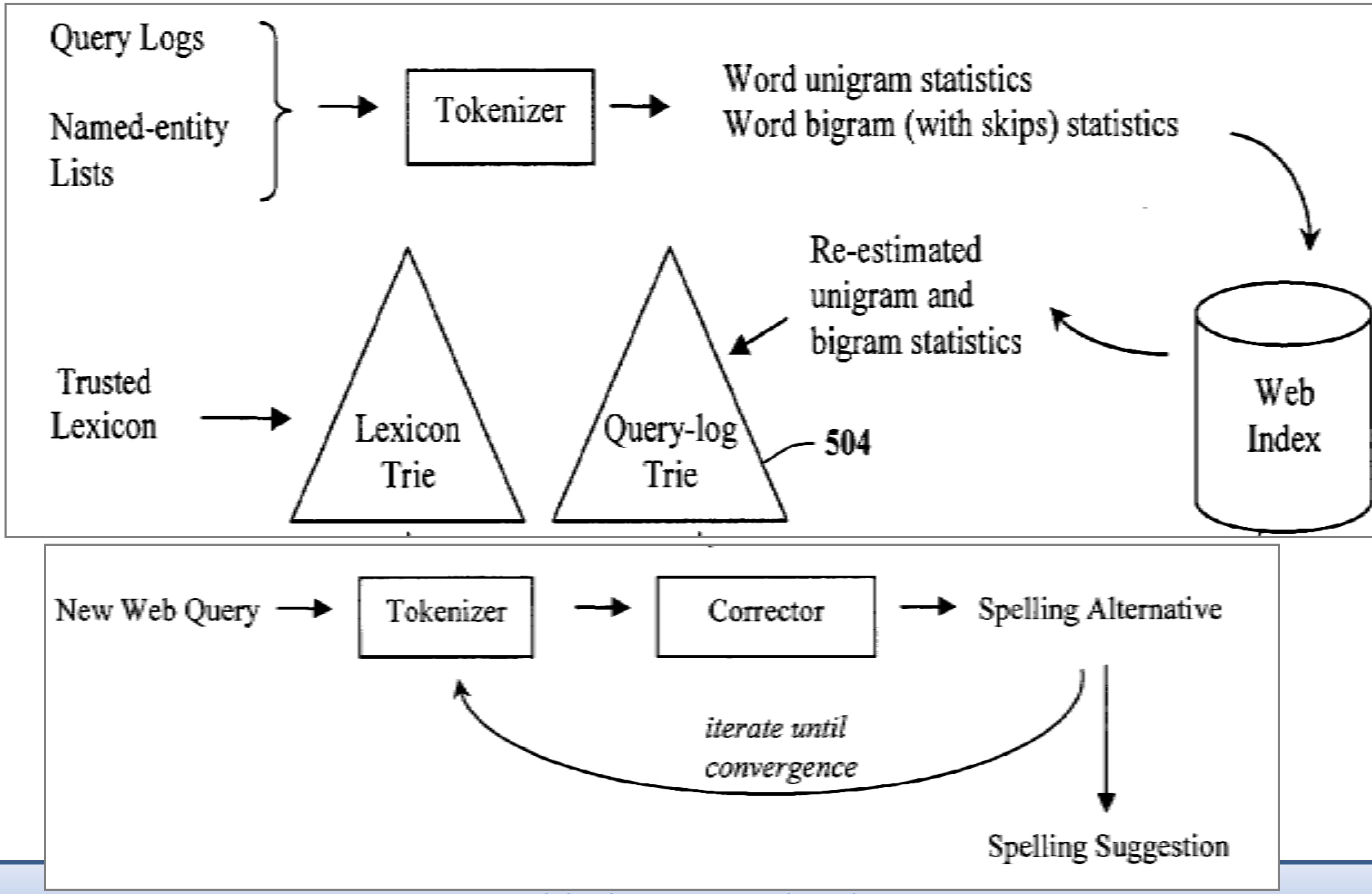
Language Model

Query level: p("sigir 2008"), p("sigir iraq")…

Mine web data sources for these probabilities

Learning Spell Checker from Query Logs

[Cucerzan and Brill, 2004]



Spelling Correction: Iterative Approach

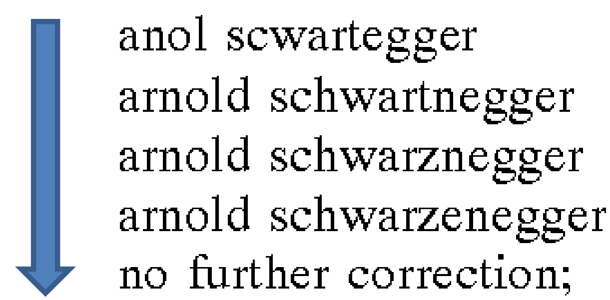
[Cucerzan and Brill, 2004]

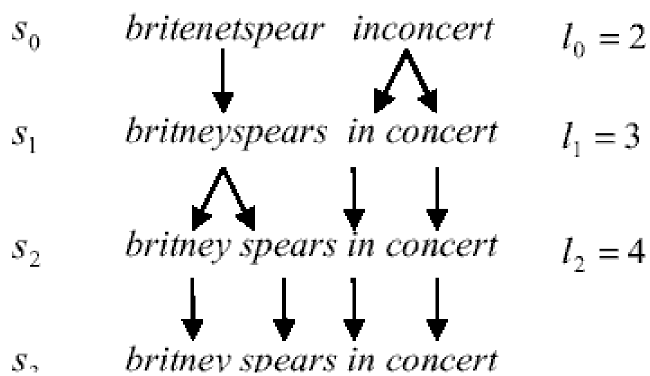
- Main idea:
  - Iteratively transform the query into other strings that correspond to more likely queries.
  - Use statistics from query logs to determine likelihood.
    - Despite the fact that many of these are misspelled
    - Assume that the less wrong a misspelling is, the more frequent it is, and correct > incorrect
- Example:
  - ditroitigers ->
    - detroittigers ->
      - detroit tigers

albert einstein	4834
albert einstien	525
albert einstine	149
albert einsten	27
albert einsteins	25
albert einstain	11
albert einstin	10
albert eintein	9
albeart einstein	6
aolbert einstein	6
alber einstein	4
albert einseint	3
albert einsteirn	3
albert einsterin	3
albert eintien	3
alberto einstein	3
albrecht einstein	3
alvert einstein	3

Spelling Correction Algorithm
[Cucerzan and Brill, 2004]

- Compute the set of all possible alternatives for each word in the query
  - Stats on word unigrams, bigrams from logs
  - Handles word concatenation and splitting
- Find the best possible alternative string to the input
  - Use modified Viterbi algorithm
- Constraints:
  - No 2 adjacent in-vocabulary words can change simultaneously
  - Short queries have further (unstated) restrictions
  - In-vocabulary words can't be changed in the first round of iteration





Spelling Correction Algorithm (cont'd)

[Cucerzan and Brill, 2004]

- Comparing string similarity
  - Damerau-Levenshtein edit distance:
    - The minimum number of point changes required to transform a string into another
- Trading off distance function leniency:
  - A rule that allows only one letter change can't fix:
    - dondal duck -> donald duck
  - A too permissive rule makes too many errors:
    - log wood -> dog food
- Actual measure:
  - "a modified context-dependent weighted Damerau-Levenshtein edit function"
    - Point changes: insertion, deletion, substitution, immediate transpositions, long-distance movement of letters
    - "Weights interactively refined using statistics from query logs"

Spelling Correction Evaluation

[Cucerzan and Brill, 2004]

- Emphasizing recall
- First evaluation:
  - 1044 randomly chosen queries
  - Annotated by two people (91.3% agreement)
  - 180 misspelled; annotators provided corrections
  - 81.1% system agreement with annotators
    - 131 false positives
      - 2002 kawasaki ninja **zx6e** → 2002 kawasaki ninja **zx6r**
    - 156 suggestions for the misspelled queries
  - 2 iterations were sufficient for most corrections
  - **Problem: annotators were guessing user intent**

Spelling Correction Evaluation

[Cucerzan and Brill, 2004]

- Second evaluation:
  - Try to find a misspelling followed by its correction
    - Sample *successive pairs* of queries from the log
      - Must be sent by same user
      - Differ from one another by a small edit distance
    - Present the pair to human annotators for verification and placement into the gold standard
      - Paper doesn’t say how many total

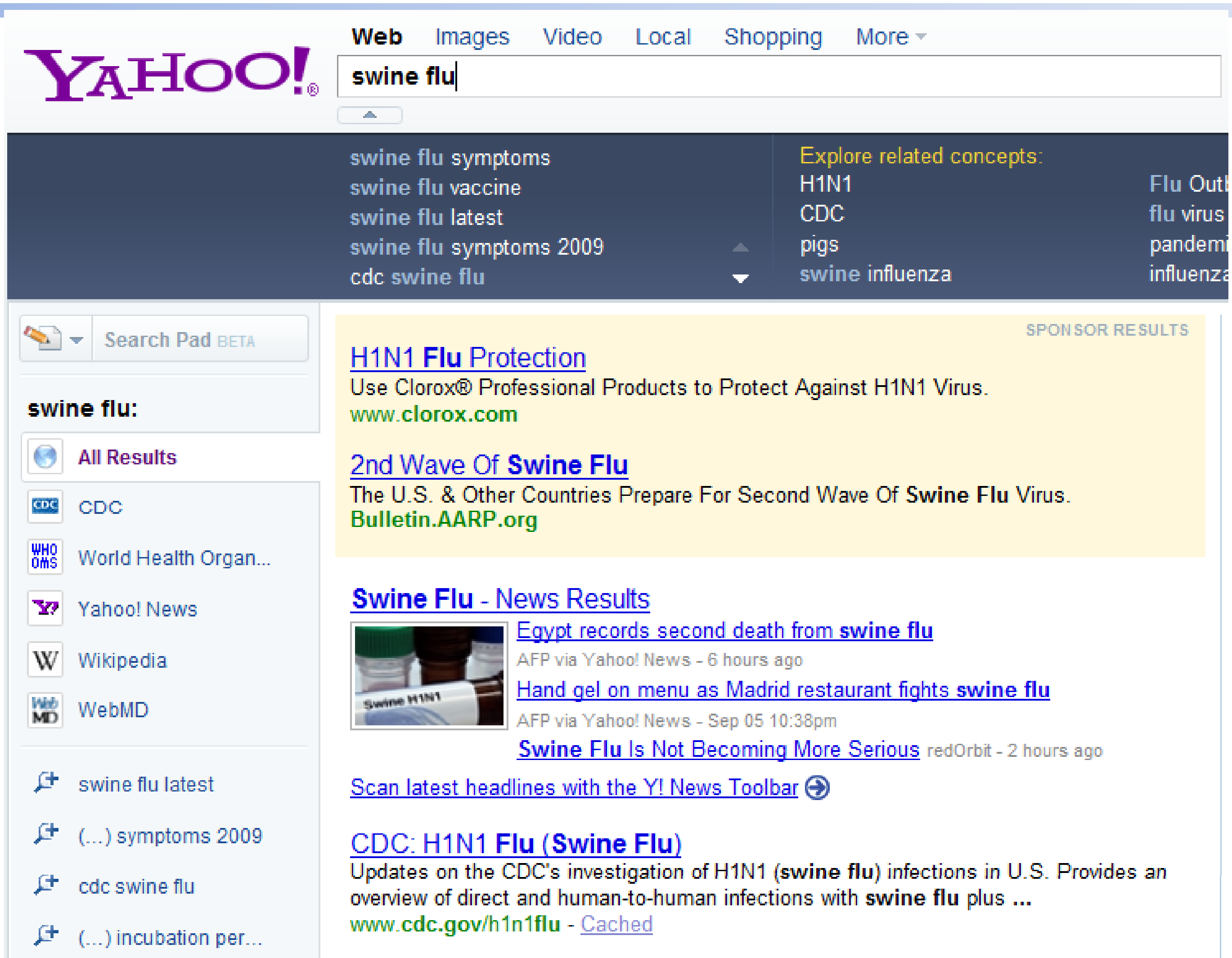
Spelling Correction Results

[Cucerzan and Brill, 2004]

- Results on 2nd evaluation:
  - 73.1% accuracy
  - Disagreed with gold standard 99 times; 80 suggestions
    - 40 of these were bad
    - 15 functionally equivalent (audio file vs. audio files)
    - 17 different valid suggestions (phone listings vs. telephone listings)
    - 8 found errors in the gold standard (**brandy sniffers)**
  - 85.5% correct: speller correct or reasonable
  - Sent an unspecified subset of the errors to Google's spellchecker
    - Its agreement with the gold standard was slightly lower

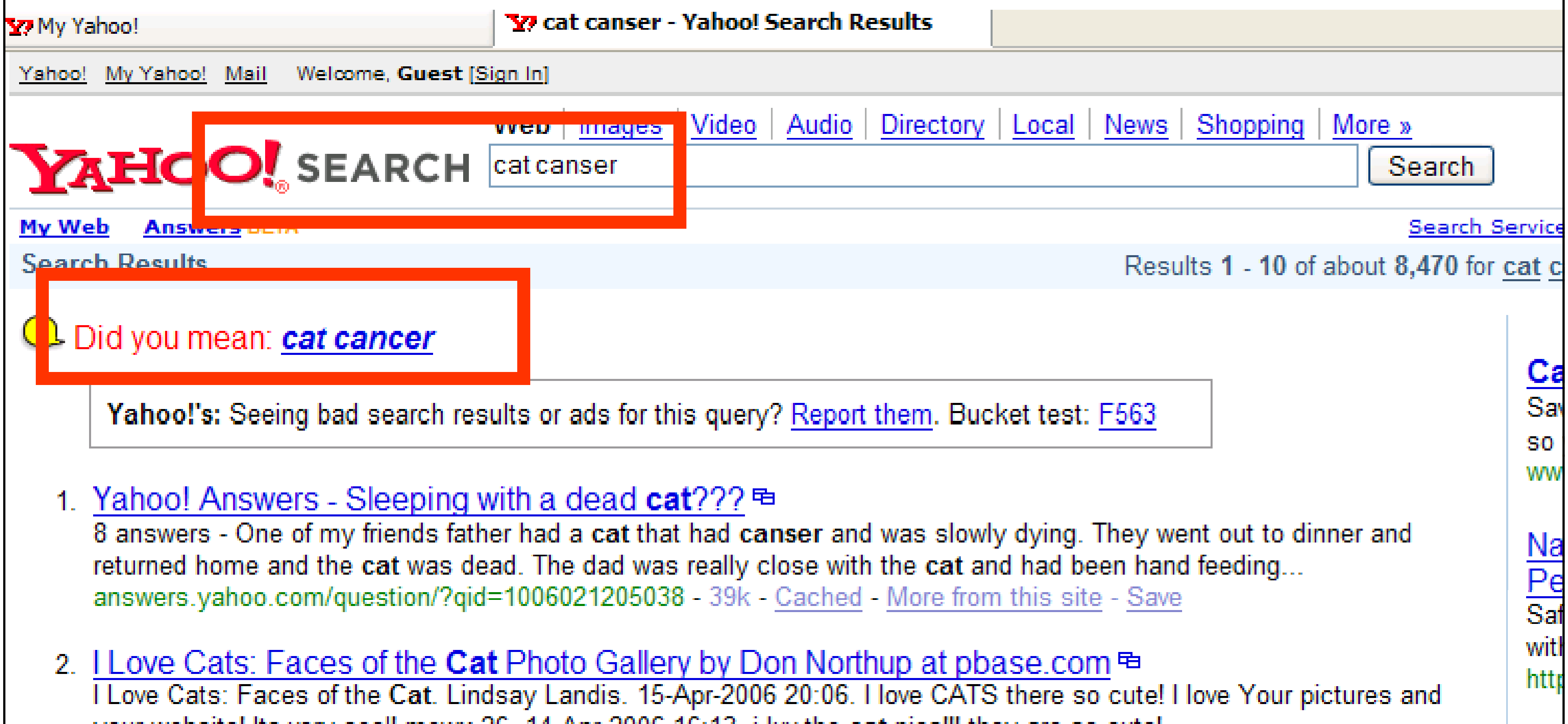
General Query Suggestion

[Slides adapted from Jones et al., 2006]



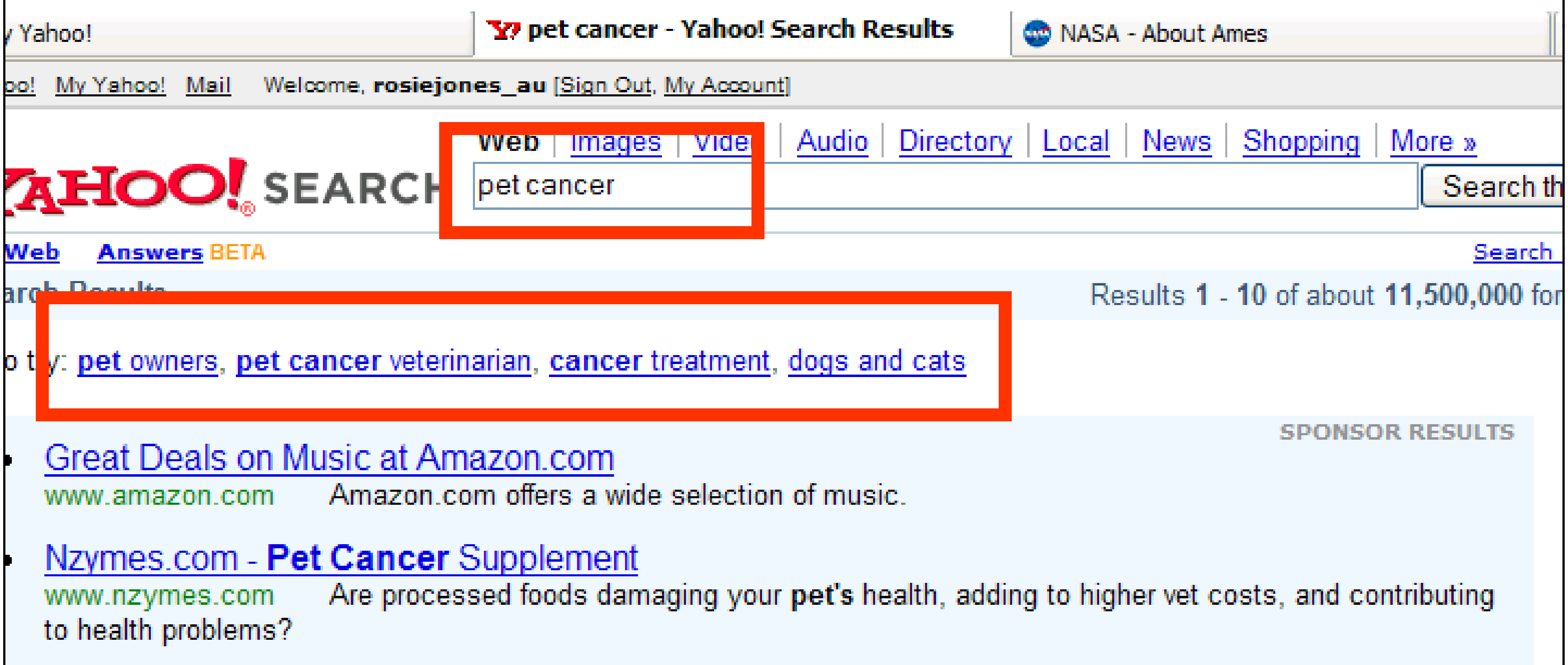
Query Substitutions

[Slides adapted from Jones et al., 2006]



Query Substitutions

[Slides adapted from Jones et al., 2006]



Functions of Rewriting

[Slides adapted from Jones et al., 2006]

- Enhance meaning
  - Spell correction
  - Corpus-appropriate terminology
    - Cat cancer $\rightarrow$ feline cancer
- Change meaning
  - Narrow
    - [lexical entailment: fruit $\rightarrow$ apple]
  - Broaden
    - [alternatives, common interests]
    - Conference proceedings $\rightarrow$ textbooks

Example: Trying to Find Nathan Welsh, who lives and works in Edinburgh

[Slides adapted from Jones et al., 2006]

- nathan welsh edinburg scotland
- nathan welsh edinburgh scotland
- financial consultants edinburg scotland
- financial consultants edinburgh scotland
- financial consultants
- nathan welsh 16-18 pennwell place edinburgh
- nathan welsh 16-18 pennywell place edinburgh
- international phone directory
- white pages
- edinburgh scotland phone directory
- edinburgh scotland uk
- nathan welsh investment consultant edinburg
- nathan welsh investment consultant edinburgh
- investment consultants edinburgh scotland
- nathan welsh
- kansas virginia
- herndon virginia

Spell correction

Name →profession

Spell correction

Delete terms, generalize

Try second approach, using his address

Spell correction

Try looking up addresses

rephrase

specialization

Generalize to location

Switch to new topic

Half of Query Pairs are Related

[Slides adapted from Jones et al., 2006]

Type	Example	%
non-rewrite	mic amps -> create taxi	53.2%
insertions	game codes -> video game codes	9.1%
substitutions	john wayne bust -> john wayne statue	8.7%
deletions	skateboarding pics → skateboarding	5.0%
spell correction	real eastate -> real estate	7.0%
mixture	huston's restaurant -> houston's	6.2%
specialization	jobs -> marine employment	4.6%
generalization	gm reabtes -> show me all the current auto rebates	3.2%
other	thansgiving -> dia de acconde gracias	2.4%

[Jones & Fain SIGIR 2003]

Substitutions are repeated

[Slides adapted from Jones et al., 2006]

- car insurance → auto insurance
  - 5086 times in a sample
- car insurance → car insurance quotes
  - 4826 times
- car insurance → geico [brand of car insurance]
  - 2613 times
- car insurance → progressive auto insurance
  - 1677 times
- car insurance → carinsurance
  - 428 times

Different Users, Different Days

Statistical Test to Find Significant Rewrites
[Slides adapted from Jones et al., 2006]

Test whether

$$p(q2 \mid q1) >> p(q2)$$

P(britney spears|brittney spears) >> P(britney spears)

8% >> 0.01%

Log likelihood ratio test (GLRT) gives χ^2 distributed score

About 90% of query pairs are related after filtering with LLR > 100

Many Types of Substitutable Rewrites

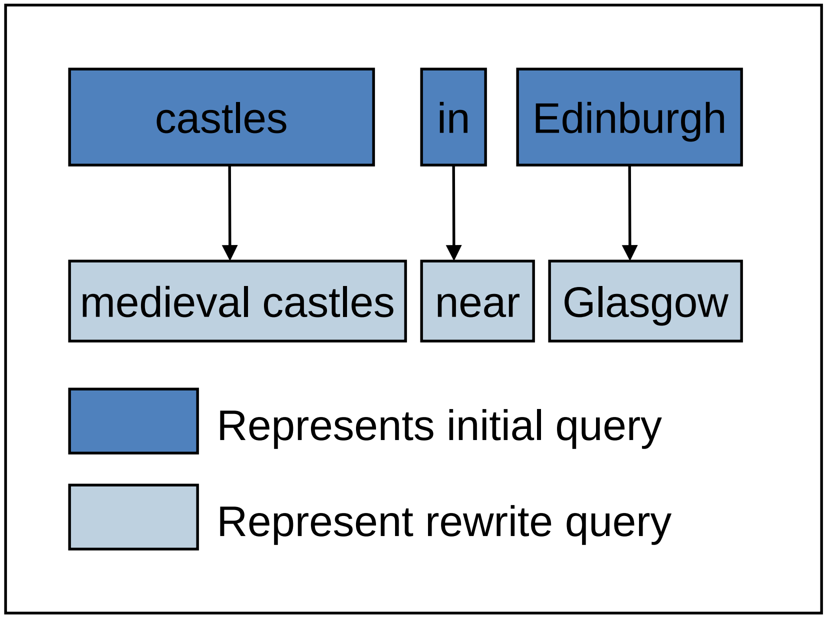
[Slides adapted from Jones et al., 2006]

dog -> dogs	9185	pluralization
dog -> cat	5942	both instances of 'pet‘
dog -> dog breeds	5567	generalization
dog -> dog pictures	5292	more specific
dog -> 80	2420	random junk in query processing
dog -> pets	1719	generalization -- hypernym
dog -> puppy	1553	specification -- hyponym
dog -> dog picture	1416	more specific
dog -> animals	1363	generalization -- hypernym
dog -> pet	920	generalization -- hypernym

Increase Tail Coverage with Query Segmentation

[Slides adapted from Jones et al., 2006]

- Query segmented using high mutual information terms
- Most frequent queries: replace whole query
- Infrequent queries: replace constituent phrases



Defining Query Relatedness for Sponsored Search

[Slides adapted from Jones et al., 2006]

1- Precise Match	A near-certain match. *E.g.: automotive insurance - automobile insurance;*
2- Approximate Match	A probable, but inexact match with user intent. E.g.*: apple music player - ipod shuffle*
3- Marginal Match	A distant, but plausible match to a related topic. E.g.: *glasses - contact lenses*
4- Mismatch	A clear mismatch.

Call {1,2} Precise and {1,2,3} Broad

Generating Query Substitutions

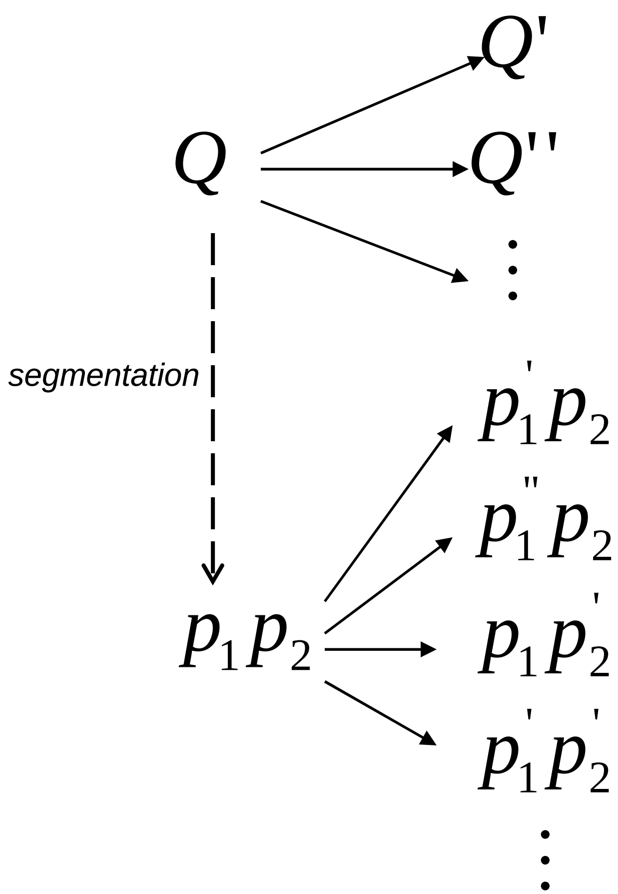
[Slides adapted from Jones et al., 2006]

- Q1 →{q2,q3,q4,q5,q6}
- "catholic baby names" →

  {christian baby names, christian baby boy names, catholic names, …}
- Learn model to rank and score

Increase Tail Coverage with Query Segmentation

[Slides adapted from Jones et al., 2006]



- Query segmented using high mutual information terms
- Most frequent queries: replace whole query
- Infrequent queries: replace constituent phrases

Generating Query Substitutions

[Slides adapted from Jones et al., 2006]

- Q1 -> {q2,q3,q4,q5,q6}
- "catholic baby names" -> {christian baby names, christian baby boy names, catholic names, …}
- All are statistically relevant (log likelihood ratio on successive queries)

Find a model to

- rank substitutions, to be able to pick the best ones

$$score(Q->u_1^{''}u_2) < score(Q->Q'') < \ldots$$

- associate a probability of correctness

$$P(Q->Q' \; is \; correct \mid score(Q->Q'))$$

Train/Test Data

[Slides adapted from Jones et al., 2006]

- Sample 1000 queries (q1)
- Select a single substitution for each (q2)
- Manually label the <q1,q2> pairs
- Learn to score <q1,q2> pairs
- Order by score
- Assess Precision/Recall
  - Precise task {1,2} vs {3,4}
  - Broad task {1,2,3} vs {4}

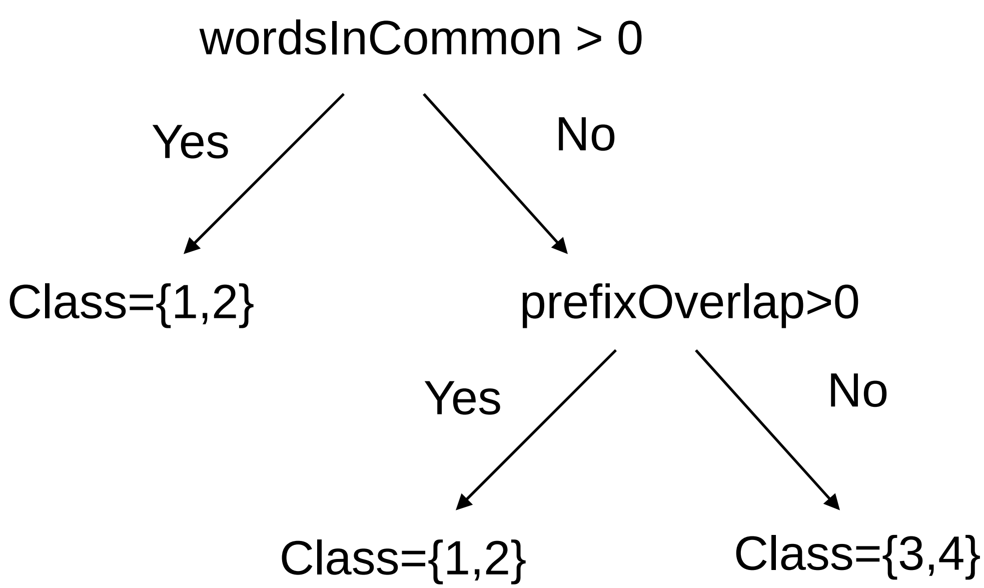
Predicting High Quality Query Suggestions

[Slides adapted from Jones et al., 2006]

- Used labels to fit model
- Tried 37 features for model:
  - Lexical features including
    - Levenshtein character edit distance
    - Prefix overlap
    - Porter-stem
    - Jaccard score on words
  - Statistical features including
    - Probability of rewrite
    - Frequency of rewrite
  - Other
    - Number of substitutions (numSubst)
      - Whole query = 0
      - Replace one phrase = 1
      - Replace two phrases = 2
    - Query length, existence of sponsored results…

Simple Decision Tree

[Slides adapted from Jones et al., 2006]



Interpretation of the decision tree:
- substitution must have at least 1 word in common with initial query
- the beginning of the query should stay unchanged

Linear Regression Model

[Slides adapted from Jones et al., 2006]

Regression: continuous output in [1,4]

$$LMScore = intercept + \sum_{f = features} w_f . f$$

Classification:

If(*LMScore* < *T*) then *Good*, else *Bad*

For each T, we have a precision and a recall

Evaluation:

Average precision / recall on 100 times 10-fold cross validation

Learned Function

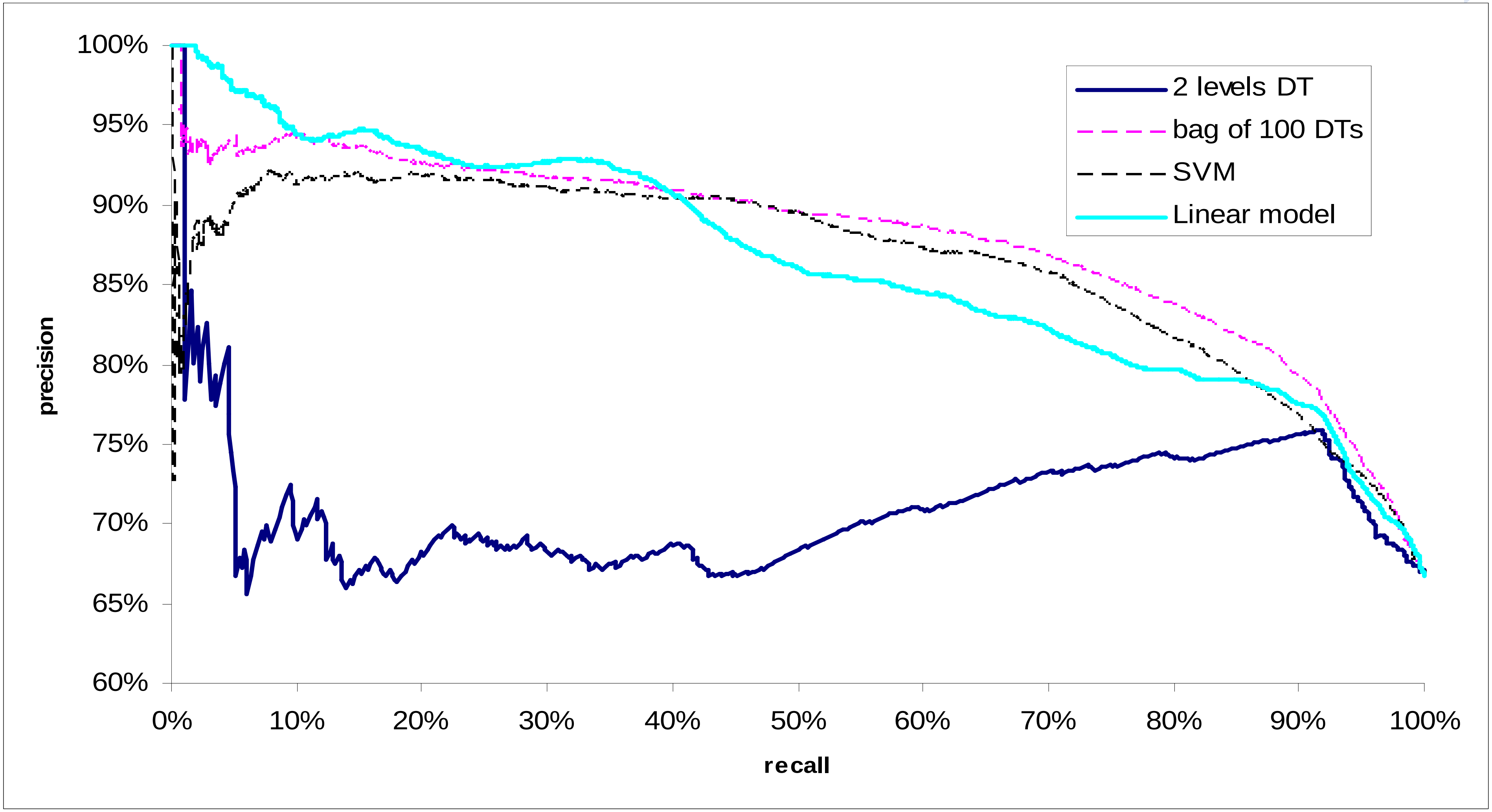
[Slides adapted from Jones et al., 2006]

$$f(q_1,q_2)=0.74+1.88\times editDist(q_1,q_2)$$
$$+0.71\times wordDist(q_1,q_2)$$
$$+0.36\times numSubst(q_1,q_2)$$

- Outputs continuous score [1..4]
- Like decision tree
  - Prefer few edits
  - Prefer few word changes
  - Prefer whole-query or few phrase changes
- Normalize output to a probability of correctness using sigmoid fit

SVM, Bags of Trees, Linear Model Trade-offs

[Slides adapted from Jones et al., 2006



Example Query Substitutions

[Slides adapted from Jones et al., 2006]

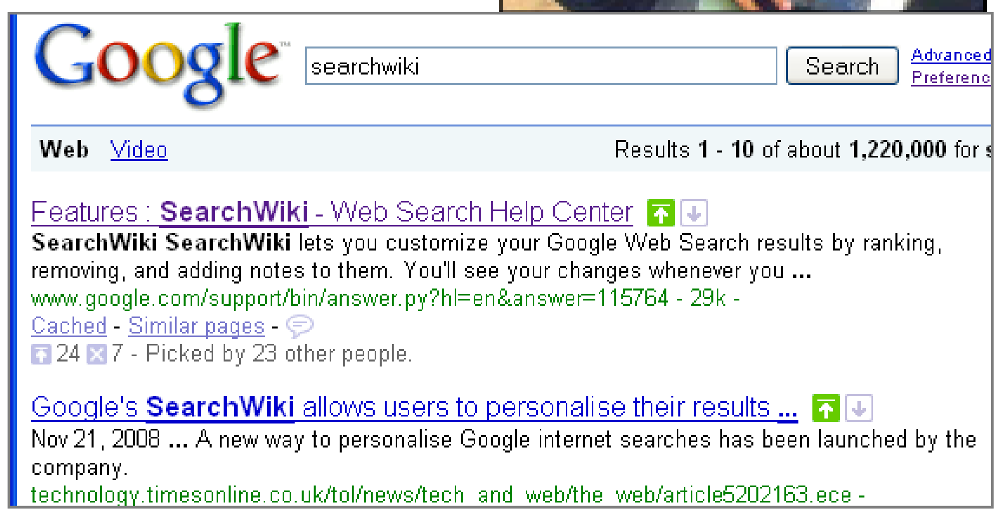
Initial Query	Substitution	Hand-label	Alg. Prob
anne klien watches	anne klein watches	1	92%
sea world san diego	sea world san diego tickets	2	90%
restaurants in washington dc	restaurants in washington	2	89%
nash county	wilson county	3	66%
frank sinatra birth certificate	elvis presley birth	4	17%

Lecture 5 Plan

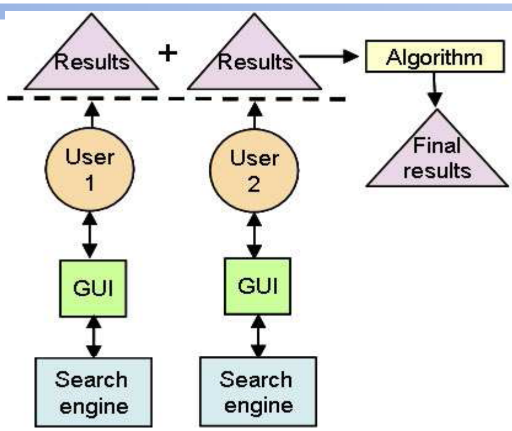
- ✓ **Generating result summaries (abstracts)**
  - ✓ Beyond result list
- ✓ **Spelling correction and query suggestion**
- ➢ **New directions in search user interfaces**
  - Collaborative Search
  - Collaborative Question Answering
- **PhD studies in the U.S.**

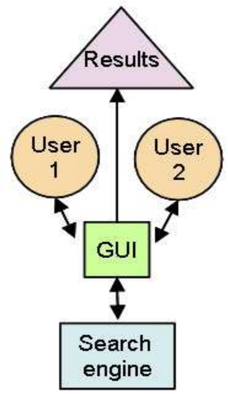
Collaborative Web Search

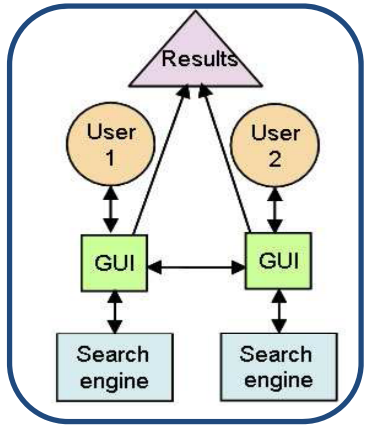
- People collaborate during Web search (Morris, 2008)
- Tools have been developed to support collaborative Web search (Morris, 2007; Pickens et al., 2008)
- Information seeking can be more effective as a collaboration than as a solitary activity.
  - Different perspectives, experiences, expertise, and vocabulary to the search process.

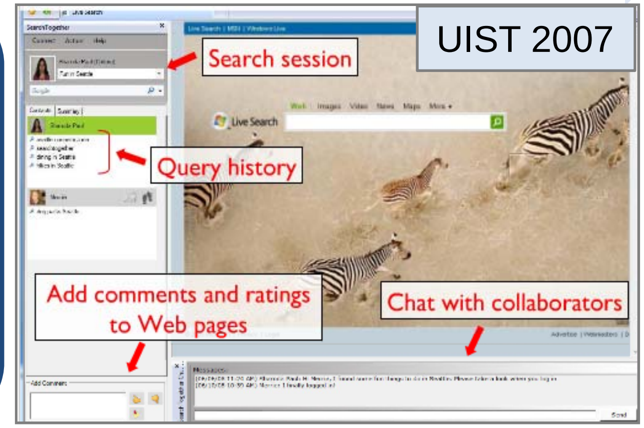


Algorithmically Mediated Social Search





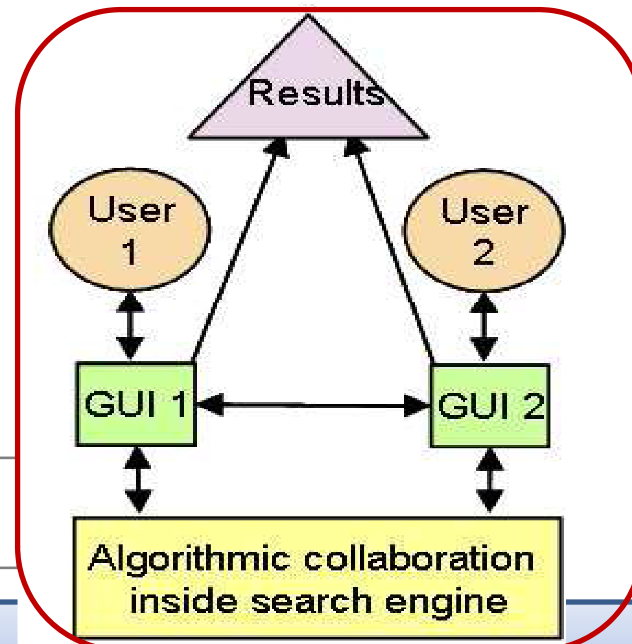




- Previous approaches (above): merge searching results from different individuals or let multiple people share a single user interface and cooperatively formulate queries

- **Pickens et al.**: algorithmically-mediated retrieval in **search engine level** to focus and enhance the team's search and communication activities

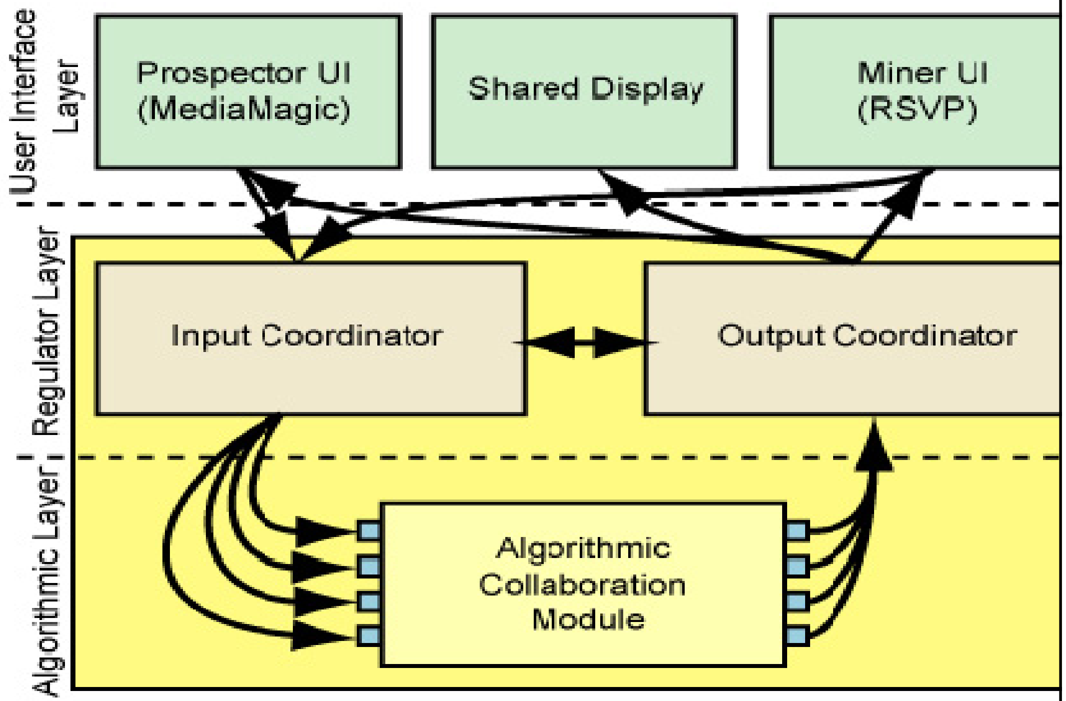
J. Pickens, G. Golovchinsky, C. Shah, P. Qvarfordt, and M. Back. *Algorithmic mediation for collaborative exploratory search*, SIGIR 2008



Algorithmically Mediated Social Search II

J. Pickens, G. Golovchinsky, C. Shah, P. Qvarfordt, and M. Back.
Algorithmic mediation for collaborative exploratory search, SIGIR 2008

- Two search roles:
  - Prospector: opens new fields for exploration into a data collection.
  - Miner: view and assess the documents returned by Prospector.



- System architecture
  - User Interface Layer
    - A query interface for Prospector to issue queries.
    - A visualization result browsing interface for Miner to assess relevance.
  - Regulator Layer
    - Input regulator is responsible for capturing and storing searcher's searching results.
    - Output regulator accepts information from the algorithmic layer and routes it to appropriate roles.

System Design

J. Pickens, G. Golovchinsky, C. Shah, P. Qvarfordt, and M. Back. *Algorithmic mediation for collaborative exploratory search*, SIGIR 2008

- Algorithmic Layer
  - Weight Definition
    - L_k: a ranked list of documents retrieved by query k.
    - Relevance: $w_r(L_k) = |rel \in L_k| / |nonrel \in L_k|$
    - Freshness: $w_f(L_k) = |unseen \in L_k| / |seen \in L_k|$
  - Miner Algorithm
    - As Prospector generates new search results, new list (L_k) is added to the whole results collection (L).
    - The documents retrieved by Prospector will be queued for Miner to assess their relevance. The queue is ordered by the following formula in w tance of document

$$score(d) = \sum_{L_k \in \{L\}} w_r(L_k) w_f(L_k) borda(d, L_k)$$

    - Both Prospector and Miner will view and judge documents, so the weights (w_f and w_r) will change over time.
    - As a result, the documents with higher scores will have more chances to be evaluated by the Miner.

System Design (cont'd)

J. Pickens, G. Golovchinsky, C. Shah, P. Qvarfordt, and M. Back. *Algorithmic mediation for collaborative exploratory search*, SIGIR 2008

- Prospector Algorithm
  - Prospector focuses on coming up with new avenues for exploration into the collection. This is accomplished by real-time query term suggestion.
  - Each term in the whole document corpus has a score which is defined by the following formula. rlf() function means the number of documents in L_k in which term t is found.

$$score(t) = \sum_{L_k \in \{L\}} w_r(L_k) w_f(L_k) rlf(t, L_k)$$

  - As Miner's algorithm affect w_f and w_r, the system will reorder term suggestions.
    - The more the Miner digs into fresher and more relevant documents, the more terms associated with those documents will appear in term suggestion.
    - Once one document proves to be not fresh and relevant, the associated terms will be gradually replaced by others.
- Collaboration is accomplished by the dynamic change of freshness value and relevance value.

Experimental Setup

J. Pickens, G. Golovchinsky, C. Shah, P. Qvarfordt, and M. Back. *Algorithmic mediation for collaborative exploratory search*, SIGIR 2008

- Goal: test the hypothesis that mediated collaboration search offers more effective searching capability than simple merging of independently produced results
- 4 teams, each team has 2 persons. Every time, one team searches in for one topic in two ways:
  - simple merging and mediated collaboration search. Each experiment lasts 15 min.
- 24 topics from TREC collection into two groups based on the total number of relevant documents available for that topic.
  - Topics that fell below the median (130) were deemed “sparse” (average of 60 relevant documents per topic).
  - Topics above the median were “plentiful” (average of 332 relevant documents per topic).
  - Searching “sparse” topics is an exploratory search process, more difficult

Results

J. Pickens, G. Golovchinsky, C. Shah, P. Qvarfordt, and M. Back. *Algorithmic mediation for collaborative exploratory search*, SIGIR 2008

	3.75 min	7.5 min	11.25 min	15 min
	Avg%Chg	Avg%Chg	Avg%Chg	Avg%Chg
P_s				
Overall	+9.8	+21.5	+22.4	+30.2
Plentiful	-2.6	+6.1	+4.2	+0.4
Sparse	+22.4	+36.8	+40.7	+60.1
R_s				
Overall	+15.2	+35.7	+19.2	+29.7
Plentiful	+13.9	+13.5	+3.8	-4.4
Sparse	+16.4	+57.9	+34.7	+63.8
P_v				
Overall	+13.6	+65.4	+41.1	+51.1
Plentiful	+16.6	+9.1	+2.3	-9.7
Sparse	+10.6	+121.6	+79.9	+111.9

Lecture 5 Plan

- ✓ **Generating result summaries (abstracts)**
  - ✓ Beyond result list
- ✓ **Spelling correction and query suggestion**
- ➢ **New directions in search user interfaces**
  - – Collaborative Search
  - ➢ **Collaborative Question Answering**
- **PhD studies in the U.S.**

jumpcut™
Make Amazing Movies Online

askville BETA
by amazon

GARAGEBAND

blinkx

flickr™ BETA

YouTube

WIKIPEDIA

twitter

Mahalo
Beta

YAHOO!® ANSWERS

Do I have a shot at Emory University?

Anthony

I have an unweighted 3.73 GPA on a 4.0 scale, and a weighted 3.82. I've only taken a couple honors classes throughout high school (Chemistry and Math 9) and no APs, but I'm taking two APs this year (senior year) (Economics and Psychology). I've taken the ACTs twice and scored a 29 Composite with a 9 out of 12 on the writing my first time, and a 30 Compoisite with a 9 on the essay on my second time. I'm a pretty well-rounded student as I have been on missions trips to 3

Ranto

TOP CONTRIBUTOR

Best Answer - Chosen by Voters

Your GPA is average for Emory. However, the average Emory student has more AP classes than you do. You are on the right track taking more -- but you aren't there yet.

Your ACT score corresponds to an SAT score of about 1920-1980. Over 75% of those who are accepted at Emory have higher SAT scores.

Bottom line -- you are close to where you should be and have a shot at at Emory -- but I would put your odds at less than 50%. While I think you have a decent shot at getting into Emory, I think it is pretty unlikely that you will get in Early Decision when your stats are below the average for students who are admitted.

If you take the SATs and score above 2100, then you have a better chance.

You will also need a killer admissions essay.

1 year ago

Source(s):
College Professor

Finding Information Online (Revisited)

Next generation of search:

Algorithmically-mediated information exchange

CQA (collaborative question answering):

- **Realistic information exchange**
- Searching archives
- Train NLP, IR, QA systems
- **Study of social behavior, norms**

Content **quality**, asker **satisfaction**

Current and future work

Finding High Quality Content in SM

E. Agichtein, C. Castillo, D. Donato, A. Gionis, and G. Mishne, *Finding High Quality Content in Social Media, in* WSDM 2008

- **Well-written**
- **Interesting**
- **Relevant (answer)**
- **Factually correct**
- Popular?
- Provocative?
- Useful?

As judged by professional editors

Do I have a shot at Emory University?

I have an unweighted 3.73 GPA on a 4.0 scale, and a weighted 3.82. I've only taken a couple honors classes throughout high school (Chemistry and Math 9) and no APs, but I'm taking two APs this year (senior year) (Economics and Psychology). I've taken the ACTs twice and scored a 29 Composite with a 9 out of 12 on the

Text analysis | **Clicks** | **Community**

Ranto

TOP CONTRIBUTOR

Best Answer - Chosen by Voters

Your GPA is average for Emory. However, the average Emory student has more AP classes than you do. You are on the right track taking more -- but you aren't there yet.

Your ACT score corresponds to an SAT score of about 1920-1980. Over 75% of those who are accepted at Emory have higher SAT scores.

Text analysis

Readability statistics

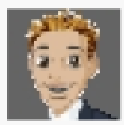
Language modeling

Punctuation density

Help! math! histogram! asap?

In Mathematics - Asked by Markyme123 - 0 answers - 3 minutes ago

Capitalization errors



WHAT is heidi montag thinking WITH THIS MUSIC VIDEO?

In Celebrities - Asked by chris_bann88 - 0 answers - 3 minutes ago

Number of words

Help!!!!!!!!!!!!?

In General - Asked by *So Confused* - 1 answer - 6 minutes ago

+ spacing density, sylablles per word,...

E. Agichtein, C. Castillo, D. Donato, A. Gionis, G. Mishne: Finding High-Quality Content in Social Media. WSDM'08.

Text analysis

Readability statistics

Language modeling

Language model disagreement

Distributions of word n-grams
and part-of-speech sequences

when|how|why -- “to” -- verb

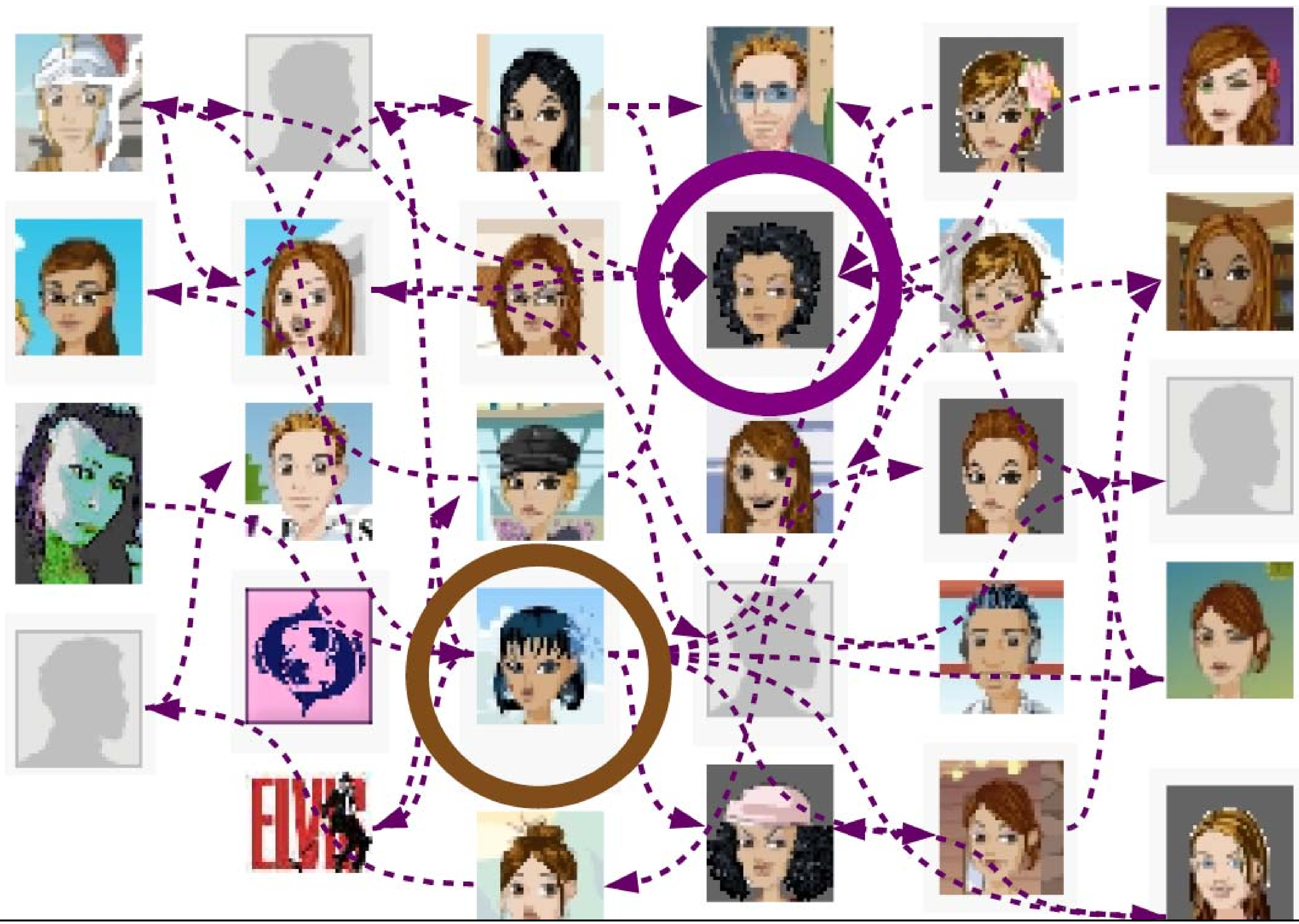
“how to identify …”

when|how|why – verb – verb – pronoun – verb

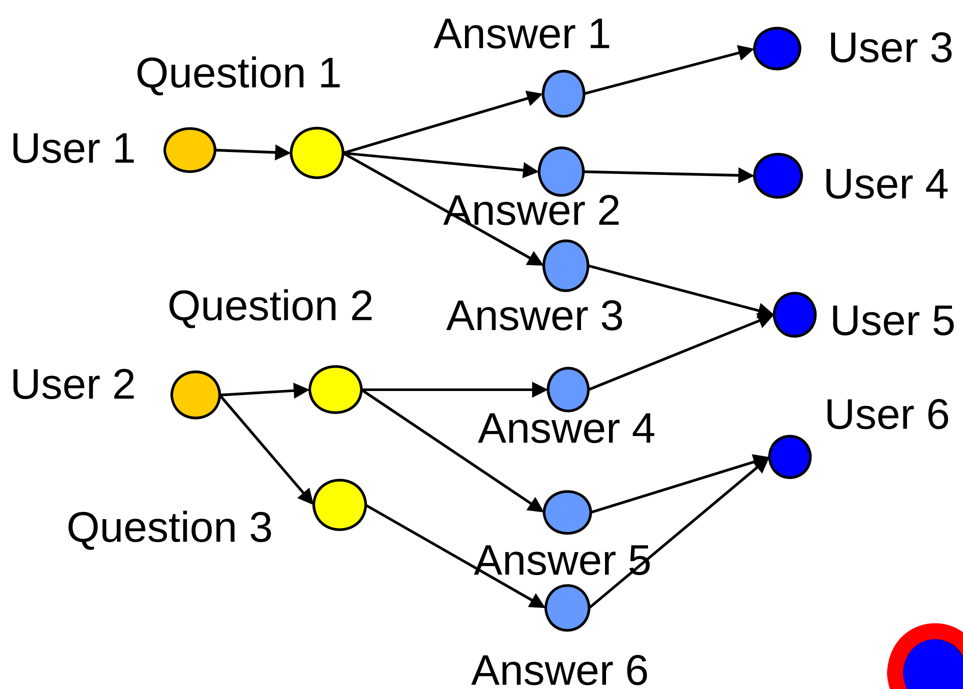
“how do I remove …”

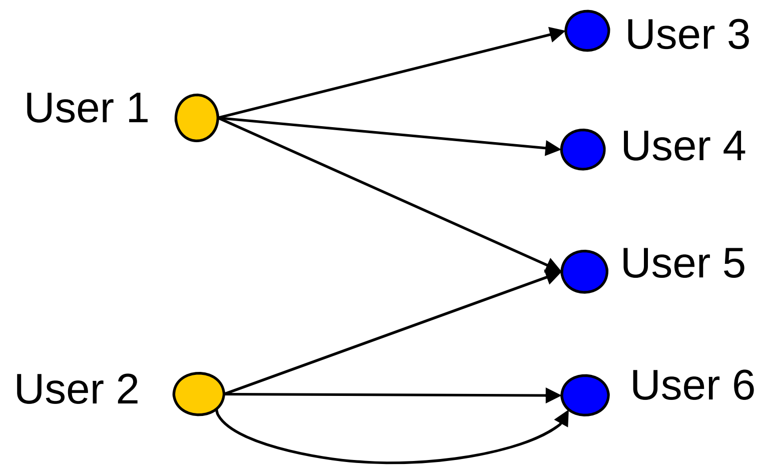
Agichtein, C. Castillo, D. Donato, A. Gionis, G. Mishne: Finding High-Quality Content in Social Media. WSDM'08.

Community



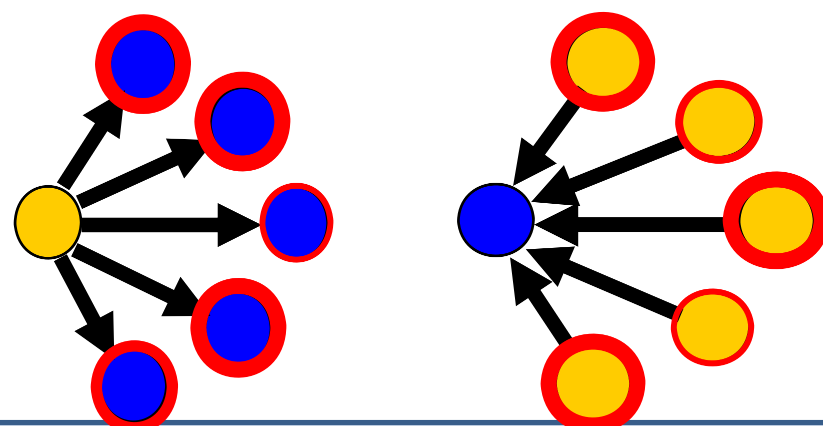
Link Analysis for Authority Estimation



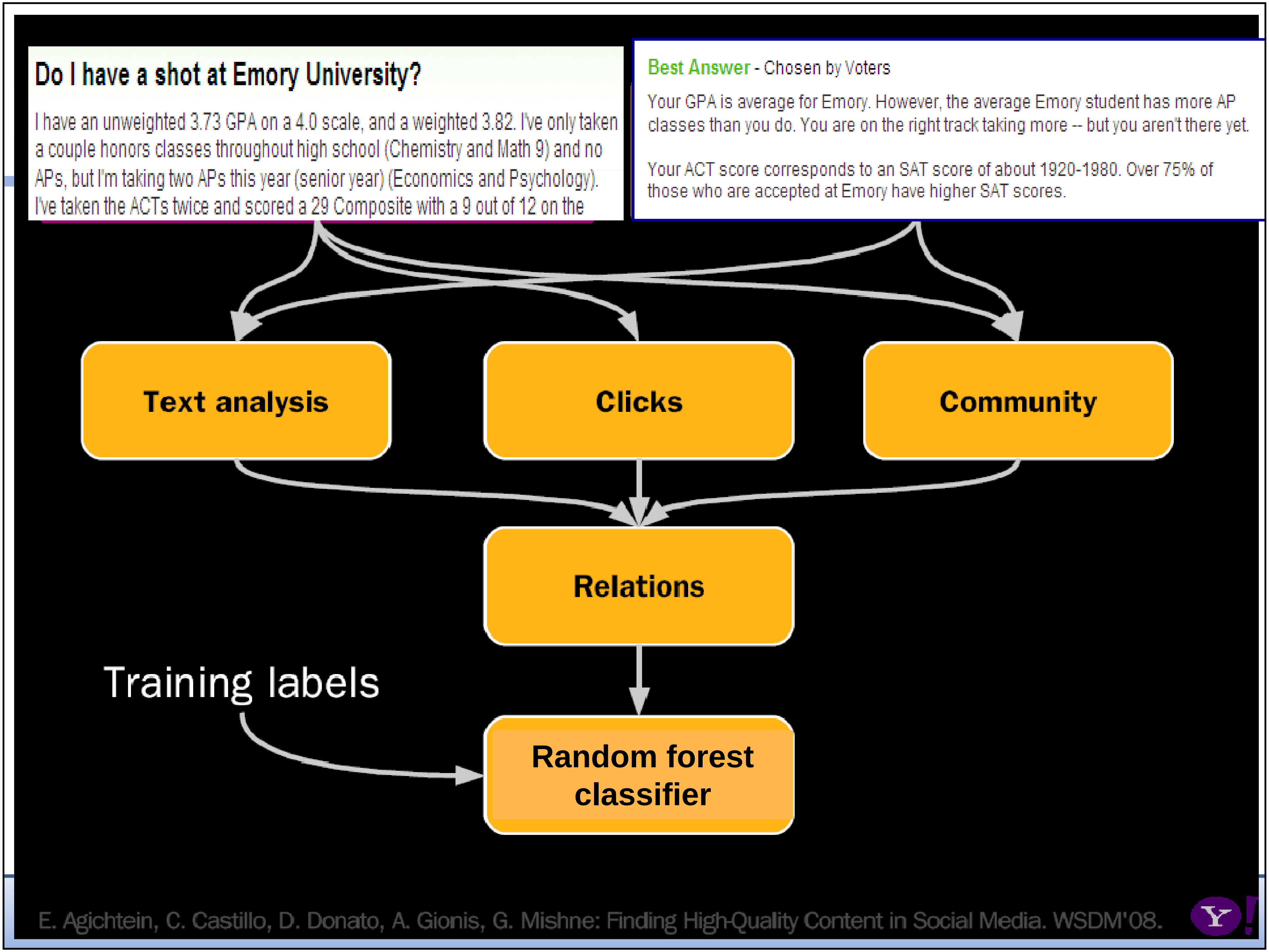


$$A(j) = \sum_{i=0..M} H(i)$$

$$H(i) = \sum_{j=0..K} A(j)$$



Hub (asker) Authority (answerer)



Do I have a shot at Emory University?
I have an unweighted 3.73 GPA on a 4.0 scale, and a weighted 3.82. I've only taken a couple honors classes throughout high school (Chemistry and Math 9) and no APs, but I'm taking two APs this year (senior year) (Economics and Psychology). I've taken the ACTs twice and scored a 29 Composite with a 9 out of 12 on the
Best Answer - Chosen by Voters
Your GPA is average for Emory. However, the average Emory student has more AP classes than you do. You are on the right track taking more -- but you aren't there yet.
Your ACT score corresponds to an SAT score of about 1920-1980. Over 75% of those who are accepted at Emory have higher SAT scores.
Text analysis
Clicks
Community
Relations
Training labels
Random forest classifier
E. Agichtein, C. Castillo, D. Donato, A. Gionis, G. Mishne: Finding High-Quality Content in Social Media. WSDM'08.

Yahoo! Answers: The Good News

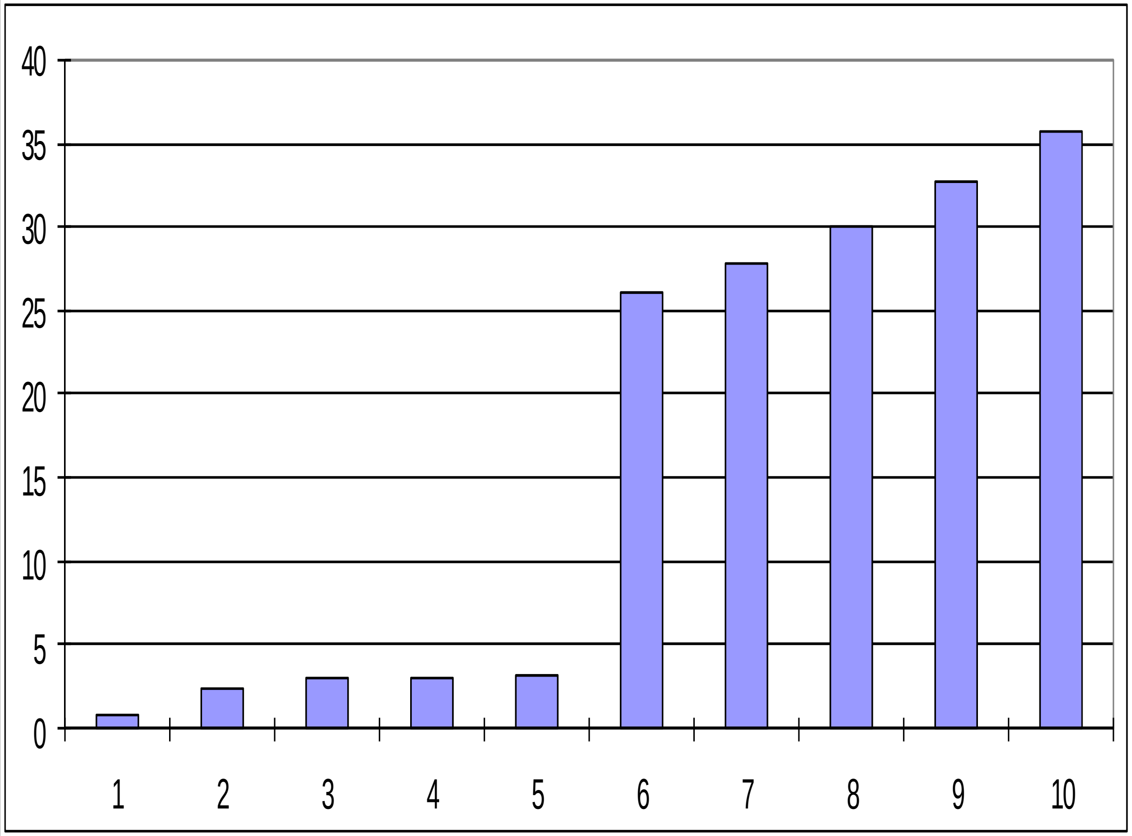
- Active community of millions of users in many countries and languages

- Effective for **subjective** information needs
  - Great forum for socialization/chat

- Can be invaluable for hard-to-find information not available on the web

Yahoo! Answers: The Bad News

May have to wait a **long** time to get a satisfactory answer



Time to close a question (hours)

1. FIFA World Cup
2. Optical
3. Poetry
4. Football (American)
5. Soccer
6. Medicine
7. Winter Sports
8. Special Education
9. General Health Care
10. Outdoor Recreation

May **never** obtain a satisfying answer

Predicting Asker Satisfaction

Y. Liu, J. Bian, and E. Agichtein, in SIGIR 2008

Yandong Liu

Jiang Bian

Given a question submitted by an asker in CQA, predict whether the user will be **satisfied** with the answers contributed by the community.

- ***"Satisfied"*** :
  - The **asker** has closed the question **AND**
  - Selected the best answer **AND**
  - *Rated best answer >= 3 "stars"* (# not important)
- Else, ***"Unsatisfied***

Satisfaction by Topic

Topic	*Questions*	*Answers*	*A per Q*	*Satisfied*	*Asker rating*	*Time to close by asker*
2006 FIFA World Cup	**1194**	**35,659**	**329.86**	**55.4%**	**2.63**	**47 minutes**
Mental Health	**151**	**1159**	**7.68**	**70.9%**	**4.30**	**1.5 days**
Mathematics	**651**	**2329**	**3.58**	**44.5%**	**4.48**	**33 minutes**
Diet & Fitness	**450**	**2436**	**5.41**	**68.4%**	**4.30**	**1.5 days**

Satisfaction Prediction: Human Judges

- Truth: asker's rating
- A random sample of 130 questions
- Researchers
  - **Agreement: 0.82 F1: 0.45 → 2P*R/(P+R)**
- Amazon Mechanical Turk
  - Five workers per question.
  - **Agreement: 0.9 F1: 0.61**
  - Best when at least 4 out of 5 raters agree

Performance: ASP vs. Humans (F1, *Satisfied*)

Best Human Perf	**0.61**

Human F1 is lower than the random baseline!

ASP is significantly more effective than humans

Top Features by Information Gain

- **0.14 Q: Askers' previous rating**
- **0.14 Q: Average past rating by asker**
- **0.10 UH: Member since (interval)**
- 0.05 UH: Average # answers for by past Q
- 0.05 UH: Previous Q resolved for the asker
- 0.04 CA: Average asker rating for category
- 0.04 UH: Total number of answers received

…

Current Work (in Progress)

- Partially supervised reinforcement models of expertise (Bian et al., WWW 2009)
- Real-time CQA
- Sentiment, temporal sensitivity analysis
- Mining forum post for health informatics (disease co-morbidity, drug side-effects, …)

Lecture 5 Plan

- ✓ **Generating result summaries (abstracts)**
  - ✓ Beyond result list
- ✓ **Spelling correction and query suggestion**
- ✓ **New directions in search user interfaces**
  - ✓ Collaborative Search
  - ✓ Collaborative Question Answering
- ➢ **PhD studies in the U.S.**

PhD Studies in the U.S.

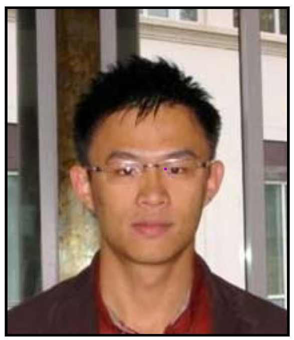
- Variants:
  - BS/BA (4-years) → MS (2 years) → PhD (4-6 years, 5 year MLE)
  - BS/BA (4-years) → MS + PhD (4-7 years, 5 year MLE)
- Application process
  - Deadline: Late Dec → Mid January
  - Standard Exam Scores:
    - GRE general
    - TOEFL
  - Application:
    - Personal statement/research interests
    - Reference letters
    - Transcript (grades).
- Other resources:
  - Pavel Dmitriev page:

    http://www.pavel-dmitriev.org/faq/question001_ru.xml

Emory Intelligent Information Access Lab (IRLab) (we are hiring…)

- Text and data mining
- Modeling information seeking behavior
- Web search and social media search
- Tools for medical informatics and public health

Ablimit Aji
(2nd year PhD)



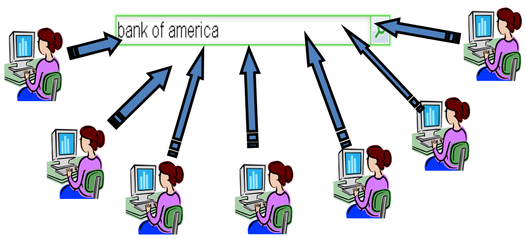
Qi Guo
(3rd year Phd)

1st year graduate students: Julia Kiseleva, Dmitry Lagun, Qiaoling Liu, Wang Yu

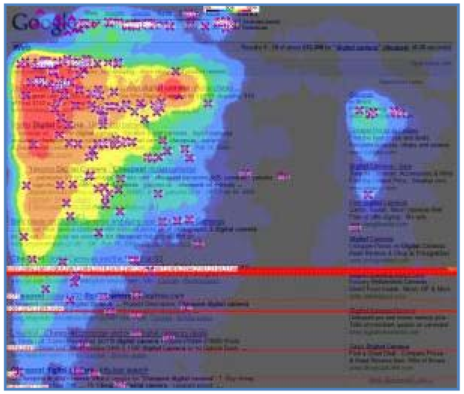
In collaboration with:
- Beth Buffalo (Neurology)
- **Charlie Clarke** (Waterloo)
- Ernie Garcia (Radiology)
- Phil Wolff (Psychology)
- Hongyuan Zha (GaTech)

Online Behavior and Interactions

Information sharing:
blogs, forums, discussions

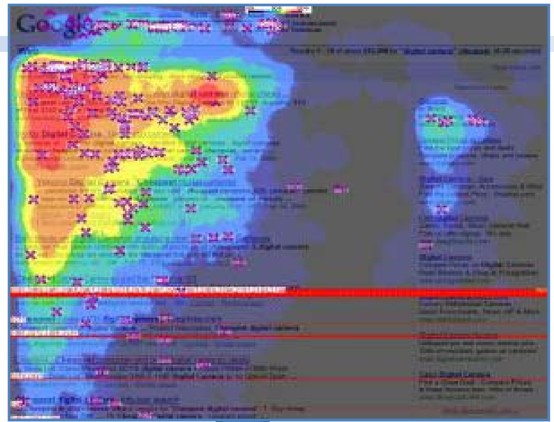


Search logs:
queries, clicks



Client-side behavior:
Gaze tracking, mouse movement, scrolling

Research Overview



Discover Models of Behavior
(machine learning/data mining)

- **Intelligent search**
- **Information sharing**
- **Health Informatics**
- **Cognitive Diagnostics**

Main Application Areas

- **Search**: ranking, evaluation, advertising, search interfaces, medical search (clinicians, patients)
- **Collaborative information sharing**: searcher intent, success, expertise, content quality
- **Health informatics**: self reporting of drug side effects, co-morbidity, outreach/education
- **Automatic cognitive diagnostics**: stress, frustration, other impairments …

References and Further Reading

- **Hearst**, Marti, ***Search User Interfaces***, 2009, Chapters 5, 6, 8, : "Presentation of Search Results", "Query Reformulation" http://searchuserinterfaces.com/
- **Croft**, Bruce, Metzler D, and Strohman, T, ***Search Engines:*** *Information Retrieval in Practice,* 2009, Chapters 6 and 10: "Queries and Interfaces", "Social Search"**,** http://www.search-engines-book.com/

Dumais, S, E. Cutrell, and H. Chen. *Optimizing search by showing results in context*, CHI 2001

Cucerzan, S and Brill, E, *Spelling Correction as an Iterative Process that Exploits the Collective Knowledge of Web Users*, EMNLP 2004

Jones, R., Rey, B., Madani, O., and Greiner, W. *Generating query substitutions*, WWW 2006

Pickens, J, G. Golovchinsky, C. Shah, P. Qvarfordt, and M. Back., *Algorithmic mediation for collaborative exploratory search*, SIGIR 2008

Agichtein, E, Gabrilovich, E, and Zha, H, E. Agichtein, E. Gabrilovich, and H. Zha, *The Social Future of Web Search: Modeling, Exploiting, and Searching Collaboratively Generated Content,* in IEEE Data Engineering Bulletin, 2009