User Experience Issues in Web Search

Daniel E. Rose
September 19, 2005
The goal of a search engine is to satisfy the user’s information need

(In other words, to help the user with an information-seeking task.)
Corollary

The more info the user gives the system about his or her information need, the better job the system can do to satisfy it.
What do we know about information-seeking?

• It’s an iterative process

  What’s the cure for AIDS?

  AIDS is caused by HIV

  What’s the cure for HIV?

  No cure, but treatments...

  Best treatments for HIV?
What do we know about information-seeking?

• Bates’ berrypicking model:

  “Each new piece of information [users] encounter gives them new ideas and directions to follow, and, consequently, a new conception of the query.”

What do we know about information-seeking?

• It is extremely subjective
What do we know about information-seeking?

- It depends on social and cultural context
  - “pants” in UK vs. US
  - “madonna and child” for
    - art historian
    - pop music fan
What do we know about information-seeking?

• *It depends on your goal*

User Input: Queries
Some Terminology

- address bar
- built-in query box (some browsers)
- search engine query box
More Terminology: Boolean Queries

penn OR teller

penn AND teller
People Don’t Understand Boolean Logic!

- Lots of studies, esp. Wason’s selection task.
- “If I search for cats AND dogs, I’ll get all the pages about cats and all the pages about dogs, right?”
  
  *(Wrong!)*

assum! /5 risk
/p ic* snow*** snowfall
/s slip! fell fall***
Alternative: Best Match

• Also known as
  – partial match
  – Natural Language
  – Scored OR

• The more the pages match my query, the higher they’ll rank – even if they don’t include ALL the words.

• Featured in most general-purpose search tools in the early 90s (Verity Search 97, Applesearch) and most original web search engines (Infoseek, AltaVista, Excite, etc.) until about 1998.
Query: *teak outdoor patio furniture*

Results:

1. “Patio World contains outdoor furniture in teak and redwood”
2. “Teak Warehouse – beautiful indoor and outdoor furniture.”
3. “Outdoor Furniture Showroom – patio chairs and benches in a variety of materials”
4. “Save the rainforest – don’t buy teak!”
How most web search engines interpret queries today

- Query words have implicit Boolean AND
- Quoted words must appear adjacent
- “-” before a word means “exclude pages containing that word”
User Behavior
What Are Users Doing?

• Not typing many words
  – Average query was 2.6 words long (in 2001), up from 2.4 words in 1997

• Moving toward e-commerce
  – less sex (down from 17% to 9%), more business (up from 13% to 25%)

Why are queries so short?

• Several possible reasons:
  – Users minimize effort
  – Users don’t realize more info is better
  – Users learn that too many words means too few results (since implicit Boolean AND)
  – Query boxes are small
Does (query box) size matter?

• Yes! Belkin research:
  – A query box that held 5 lines of text yielded longer queries
  – Different instructions (describe information problem) yielded longer queries

• But, in a non-web, best match system

Why are users searching?

- **Navigational** 15%
- **Resource** 22%
- **Informational** 63%
  
  - Directed 3%
  - Undirected 31%
  - Obtain 8%
  - Interact 6%
  - Entertain 4%
  - Download 4%
  - List 3%
  - Locate 24%
  - Advice 2%

---

Do users know how to search?

- Hargittai has many examples of confusion:
  - Confusion about address bar:
    - my wallet has been stolen.com
    - www.new york times.com
  - Queries without spaces
    - presidentialcampaign2000
    - Princetonhistoricalsocietyvolunteer

  “In email and Web addresses there are no spaces, so I tend not to use them in searches either.”

Do users know how to search?

- Quotation marks:
  - Only 16% of participants used quotation marks, many incorrectly
  - Some quoted single terms
  - Some quoted all terms for all queries

- Other operators?
  - lactose intolerance -recipes

Novices vs. Experts

• Bhavnani
  – Looked at differences between expert and novice searchers doing medical search task:
    • Experts visited 3 sites, took 7 minutes, got all the info.
    • Novices used only Google, visited 13 sites, took 20 minutes, and missed lots of info.

• Hargittai
  – “The ability to draw on a range of strategies and the agility to switch between them easily seems to be a key ingredient to successful and efficient Web navigation.”

Search is hard!
Vocabulary Problem

- People use different words for the same thing.
  - <20% chance of choosing same word
  - Even “best” word has 65-85% failure

“…The data show that no single access word, however well chosen, can be expected to cover more than a small proportion of users’ attempts…”

The Vocabulary Problem in Web Search

• Two people are unlikely to use the same word to describe the same thing…

• So, a web page author is unlikely to have used the same words as the user searching for the page’s content

• But anchortext helps
Recognition vs. Recall

• Terminology:
  – Recall: Remembering something without seeing it.
  – Recognition: Identifying something that exists in memory.

• People are better at recognition than recall
  – Menus vs. Commands

Recognition vs. Recall in Web Search

• Users may not remember the correct term, but could select it if they saw it.
One Possible Solution: Offering Document Concepts

• AltaVista Prisma

Web search is even harder!
Why harder?
All those different goals…
Why Harder?
Web Corpus is Different

- Heterogeneous (in many ways)
  - Format, length, genre, authority, quality…

- Stability
  - Ongoing growth of content volume
  - Ongoing “linkrot” (2 year half-life)
  - Content change in same URL
  - *Don’t know whether info is there*
Why Harder?
Spammers

• Misleading content
  Repeating terms, adding competitor’s products, adding unrelated terms, cloaking…

• Misleading links
  Multiple domains, link farms, guestbook bots…
How to measure success of search user experience?
Still More Terminology

RECALL = $\frac{|\text{Relevant} \cap \text{Retrieved}|}{\text{Relevant}}$

PRECISION = $\frac{|\text{Relevant} \cap \text{Retrieved}|}{\text{Retrieved}}$
Measure relevance?

• Recall
  
  Breaks down in shift from thousands to billions
  
  Result set now bigger than entire TREC corpus

• Precision
  
  What user goals are we trying to satisfy?
  
  Harder to assess heterogeneous results
    » Pages that link to good pages?
    » Pages that had good content when indexed, but no longer?
Monitor user actions?

- What’s the right number of clicks for
  - Navigational query?
  - Research?
  - Question?
  - Entertainment?
- Problem of unexpected content
Observe users?

- Time on Task
  - Exploration/quality vs. speed
    What if task is open-ended?
    What if it’s fun?
Ask users (through surveys)?

- Feedback from existing users
  Popup & inline rating forms.
  **Incentive?**

- Paying users
  **Wrong incentive!**
  (Discarded 55% in one AltaVista test.)
Evolution of the Search UI?
Hierarchical Fuzzy Logic Control of a Double Inverted Pendulum

Project Management with Fuzzy Logic

Project Management with Fuzzy Logic

Project Management with Fuzzy Logic

CENTRAL NOTICE - Fuzzy Logic

Archimedes Fuzzy Logic Development Tool

Mach-Intensive Control, Neural Networks, and Fuzzy Logic
Innovation?
Conclusion

The web search user experience does not reflect what we know about user behavior.