What Can Search Tell Us?
A Human-Centered Perspective

Alejandro (Alex) Jaimes
Yahoo! Research (Barcelona, Cat, Spain)

Social Media Engagement Research Group (SOMER)

Search Solutions, Nov. 28th, 2012
Human-Centered Computing

Data

User

Hypotheses

Design

Context
Human-Centered Computing

- User Experience
- Data analysis
- Machine Learning, User Modeling, Personalization

Human aspects
I. What is Search?
Search

- Trying to find something or someone
- By seeking carefully, scrutinizing thoroughly

from late Latin circare ‘go around,’ from Latin circus ‘circle.’
NOT A Punctual Activity!
More than a query and a click

It is a series of activities around a specific goal
BUT
We often get distracted, our “goal” is often not specific, and we frequently change goals during the search process.
look for, seek, forage for, fish around/about for, look high and low for, ferret around/about for, root around/about for, rummage around/about for, cast around/about for, look through, hunt through, explore, scour, rifle through, go through, sift through, comb, go through with a fine-tooth comb; turn upside down, turn inside out, leave no stone, examine, inspect, check, frisk, hunt, look, quest, pursuit, manhunt, searching for, hunting for, seeking, looking for, on the lookout for, in pursuit of.
Interaction experience
WE’RE GREAT AT ADAPTING TO BAD TECHNOLOGY
Information Retrieval
Information Retrieval
Information Interaction
New Paradigm

• Search (pull):
  • $Q(\phi) \rightarrow$

    Explicit HUMAN

• Discovery (push):
  • $D[\Delta] \rightarrow$

    Implicit NON-HUMAN
I. In order to *really* support search, we need to consider its complex, practical aspects
Main Points

- Not single query
- Not single session
- Context (implicit, explicit, internal, external)
II. How Can Search Analysis Be Useful?
Two “main” research areas

• User experience

• Analytics
Five Dimensions

- Personal (Personal Information Management)
- Algorithmic (improve results)
- Consumer insights
- Design (user experience)
- Macro (marketing and social aspects)
Five Dimensions

- Personal (Personal Information Management)
- Algorithmic (improve results)
- Consumer insights
- Design (user experience)
- Macro (marketing and social aspects)
Does in-network behaviour depend on how users “enter” the network?
Session Length

Long sessions
(mail)

Short sessions
(aggregators)
Clusters
Main Observations I

- Mail
  - Longer sessions
  - Strong focus on managing contacts
- Search
  - Search task “moves” to Flickr
- Aggregators
  - Short sessions
  - Image browsing
Main Observations II

- The entry point of a user influences her behaviour
  - Session length
  - Type of page visited
  - Actions taken
  - Engagement
Comparing different ranking methods

[SIGIR 2012]
Image Search [with Roelof von Zwol, R. Sierra]

• 1.4 Billion anonymous search queries
  • 75 M unique queries

• 100K most frequent queries
Observations

• Most people that “search” for images are actually browsing!

• Impact on experience design, features, algorithms, and metrics
Challenges

- What do users do? How do they search?
- Which categories?
- Scalability
- Noise
- Diversity
- Freshness
- Accuracy

Image Search

- Cognitive psychology
- Library sciences
- HCI
- IR
- Computer Vision
- Multimedia
- Engineering
Parallel Search
Macro

- Social aspects
- Marketing
Search Behavior [with Ingmar Weber]

- 2.3 million “active” registered web search users
- Anonymous, aggregate analysis
- Who? What? How?
Search Behavior

Results

• “Baby boomers”: older users, business & finance, more navigational
• “Liberal females”: women, shopping, more likely to have voted Obama, longer sessions
• “White conservatives”: older men, auto, business, cons. Electronics, less navigational
• “Challenged youth”: younger, lower income, music, more navigational
Search Behavior: Information Flows

Results II  (w/Ingmar Weber, CIKM ’10)

• “sonia sotomayor”
  • hispanic -> everybody (burst) -> hispanic

  ![Graph showing information flow]

• “avatar movie”
  • young/male -> everybody (burst) -> everybody

  ![Graph showing information flow]
Yahoo! Clues (http://clues.yahoo.com)

Most Popular Search Queries

1. India
2. Philippines
3. Iran
4. China
5. Japan

1. India
- India’s Nifty Stock Futures Advance as Valuations ..., Bloomberg
- India’s Infosys posts profit rise but shares fall, AFP
- India Official: Last Fiscal Year Power Transmission..., Nasdaq

2. Philippines
- Philippines: No word from 2 Americans' abductors, Boston Herald
- Philippines says China sea claim shaky, AFP
- Earthquake Off Philippines Measuring 6.2: Official..., ThirdAge

3. Iran
Yahoo! Clues (http://clues.yahoo.com)

Find Trends

angelina jolie
Compare with ... (optional)

Popular Searches

Today
- alison krauss
- courtney cox
- katherine jenkins
- applebees
- dolly parton

Past 7 Days

*Anonymous, based on census data*
Yahoo! Clues

Previous Queries
angelina jolie movies
angel
craigslist los angeles
jolie
angelina

Next Queries
angelina jolie twins
brad pitt angelina jolie
angelina jolie movies
angelina jolie hot
angelina jolie news

Related Searches
brad pitt, james haven, faye dunaway, ethan hawke, jennifer aniston, megan fox, jennifer lopez, lindsay lohan, britney spears, anne hathaway

*Anonymous, based on census data
yahoo! the signal

romney on 47 percent
mitt romney 47%
romney 47 video
romney 47% comment
mitt romney 47% statement
47% comment
47 percent
romney 47% quote
romney 47% comment video
47%
Yahoo! The Signal

obama vs romney polls
obama romney poll numbers
romney obama polls 2012 florida
obama vs romney ohio poll
latest obama vs romney polls
current obama vs. romney poll
obama / mitt romney opinion poll
latest 2012 presidential polls
<table>
<thead>
<tr>
<th>obama vs romney 2012</th>
<th>obama and romney ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>obama vs romney views</td>
<td>obama vs romney</td>
</tr>
<tr>
<td>romney vs obama comparison</td>
<td>map of obama and romney</td>
</tr>
<tr>
<td>mitt romney vs obama facts</td>
<td>obama vs romney debate 2012</td>
</tr>
<tr>
<td>obama vs romney who is ahead?</td>
<td>obama versus romney 2012</td>
</tr>
</tbody>
</table>
What happens if Obama wins
Obama big bird
Obama's birth certificate
Obama will win

Romney's dog
Romney flip flops
Why Romney will win
Can Romney win in 2012
Challenges

• How to group? What is the right level?
• Which categories?
• Automation
• Noise
• Freshness
• Interestingness (context)
• Interaction & interpretation

The Signal

• Design
• NLP
• Entity recognition
Insights
III. What are the “hot,” “new” areas?
Key areas

- Mobile+muliscreen
- Multimedia
- Social (?)
Mobile
(it’s all mobile)
Multimedia
(conversations around media)
Social
(to be seen)
IV. Big Data
The New Data Paradigm

- Mobile production
- Fragmentation
- Aggregation
- Multi-authorship
Search Has Changed
Have we?
IT’S all ABOUT THE CONNECTIONS, CONNECTIONS, CONNECTIONS, CONNECTIONS!
On Implicit Links

“unapparent connections are stronger than ones which are obvious”

Heraclites
Fifth century BC
Main Points

1. Not single query, session; context
2. How can search analysis be useful
3. “Hot,” “new” areas
4. Big Data
Human-Centered Computing

Data → User → Hypotheses → Design → Context

- Data
- User
- Hypotheses
- Design
- Context
Human-Centered Computing

Data
- [HT 2012]
- [ICME 2012]
- [I-Pat, KDD 2012]

User
- [TIST 2011]
- [CIKM 2011]
- [WSDM 2011]
- [WISE 2012]
- [MM 2012]
- [SIGIR 2012]

Design
- [DIS 2012]
- [MM 2012]

Hypotheses

Context
- [SIGIR 2012]
Human-Centered Computing

User Experience

Data analysis

Human aspects

Machine Learning, User Modeling, Personalization
THANK YOU!
Alex Jaimes

Social Media Engagement Group
Neil O’Hare, Luca Aiello, Amin Mantrach, Daniele Quercia, Pancho Tolchinsky, Jordi Vallmitjana, Ricard Sierra, Jordi Atserias, Luca Chiardini, Eduardo Graells, Michele Trevisol, Ruth Garcia.

Collaborators
Ingmar Weber, Barbara Poblete, Marcelo Mendoza

Papers listed at: http://www.jaimes-larrarte.com