A Comparison of Microblog Search and Web Search
Welcome to #NewTwitter! Read up on what's new. You can still access old Twitter for a limited time.

Results for #TwitterSearch

Tweets Tweets with links Tweets near you People

jteevan Jaime Teevan
Made it to Hong Kong for #wsdm2011. Taking advantage of my jet lag to finish my #TwitterSearch slides. Personal navigation slides next... 3 minutes ago

merrierm Merrie Ringel Morris
Our #TwitterSearch paper is being presented at #WSDM2011 this week: http://bit.ly/cd0eQA 6 minutes ago

RonaldGrey Ronald Grey
@CoffeyBrowne Would you please share my article? Re: creationStories and the economy: http://wp.me/pZiAD-1Z0 Thanks! #twitterSearch
TwitSCOOP - STAY ON TOP OF TWITTER!

Home  Hot Trends  Search

Search for:  Search  Clear All  Auto Refresh  Pause  Graph

15,105 people like this. Be the first of your friends.

Follow @twitscoop
Follow @twitscooptrends
Follow @hbookmark

Language:  English

#twittersearch

Tweet a link to this search

jteevan: Made it to Hong Kong for #wsdm2011. Taking advantage of my jet lag to finish my #TwitterSearch slides. Personal navigation slides next...
7 minutes ago via web

merrierm: Our #TwitterSearch paper is being presented at #WSDM2011 this week: http://bit.ly/cd0eQA
10 minutes ago via web

ronaldgrey: @CoffeyBrowne Would you please share my article? Re: #creationStories and the economy: http://wp.me/pZiAD-1Z0 Thanks!
#twitterSearch
6 hours ago via web
@M005 He was at some Retromobile show in Paris actually... #twittersearch

please share my article? Re: creationStories and the economy: http://wp.me/pZJD-1ZQ Thanks! #twitterSearch

Made it to Hong Kong for wsdm2011. Taking advantage of my jet lag to finish my #TwitterSearch slides. Personal navigation slides next...

"@yehrightdude: Tweets that don't show up in a trending session #aretheworst Why can't I see them? Can you see my #tweet? #twittersearch

#TwitterSearch: a comparison of microblog search and web search | ACM http://bit.ly/eeEtB6 thx @giustini

@dan_vh i have the #twittersearch for #robertkubica open - does that get me back in the good books #f1

'Totally Free Websites"
WSDM Microblog Papers

- TwitterRank: Finding topic-sensitive influential Twitterers [Weng, Lim, Jiang, He]
- Everyone's an influencer: Quantifying influence on Twitter [Bakshy, Hofman, Mason, Watts]
- Learning concept importance using a weighted dependence model [Bendersky, Metzler, Croft]
- Early online identification of attention gathering items in social media [Mathioudakis, Koudas, Marbach]
- Identifying topical authorities in microblogs [Pal, Counts]
- Transient crowd discovery on the real-time social Web [Yeshwanth, Caverlee]
- We feel fine and searching the emotional Web [Kamvar, Harris]
- Topical semantics of Twitter links [Welch, Schonfeld, He, Cho]
But How is Microblog Search Used?
Past Surprises About Web Search

- Early log analysis ([Jansen et al. 2000; Broder 1998])
- Queries are not 7 or 8 words long
- Advanced operators not used or “misused”
- Nobody used relevance feedback
- Lots of people search for sex
- Navigation behavior common

➢ Prior experience was with library search
Microblog Search Surprises?

Welcome to #NewTwitter! Read up on what’s new. You can still access old Twitter for a limited time.

Results for #twittersearch

jteevan  Jaime Teevan
Made it to Hong Kong for #wsdm2011. Taking advantage of my jet lag to finish my #TwitterSearch slides. Personal navigation slides next...
52 minutes ago

merrierm  Merrie Ringel Morris
Our #TwitterSearch paper is being presented at #WSDM2011 this week: http://bit.ly/cdoeQA
55 minutes ago
Microblog Search Surprises?

Ordered by time

Ordered by relevance
Methodology

• Survey of 54 Microsoft Twitter users
• Query log analysis
  – Twitter queries issued to http://twitter.com
    • Sample of 33k users over 2 weeks
    • 126k queries
  – Web queries issued to Bing, Google, and Yahoo!
    • For the users who issued Twitter queries
    • 2.5 million queries
• Comparison of search results
Methodology

• Survey of 54 Microsoft Twitter users

• Query log analysis
  – Twitter queries issued to [http://twitter.com](http://twitter.com)
    • Sample of 33k users
    • 126k queries over 2 weeks
  – Web queries issued to Bing, Google, and Yahoo!
    • For the users who issued Twitter queries
    • 2.5 million queries

• Comparison of search results
Methodology

• Survey of 54 Microsoft Twitter users

• Query log analysis
  – Twitter queries issued to http://twitter.com
    • Sample of 33k users
    • 126k queries over 2 weeks
  – Web queries issued to Bing, Google, and Yahoo!
    • For the users who issued Twitter queries
    • 2.5 million queries

• Comparison of search results

Queries issued
Temporal patterns
Cross-corpus behavior
## Top Web Queries Issued

<table>
<thead>
<tr>
<th>Web</th>
</tr>
</thead>
<tbody>
<tr>
<td>twitter</td>
</tr>
<tr>
<td>youtube</td>
</tr>
<tr>
<td>facebook</td>
</tr>
<tr>
<td>google</td>
</tr>
<tr>
<td>myspace</td>
</tr>
<tr>
<td>youtube.com</td>
</tr>
<tr>
<td>yahoo</td>
</tr>
<tr>
<td>ebay</td>
</tr>
<tr>
<td>craigslist</td>
</tr>
<tr>
<td>myspace.com</td>
</tr>
</tbody>
</table>

- Top Web queries navigational
  - Learn more today after lunch
- Biased towards social networking sites because of our user sample
- Twitter queries cannot be navigational
### Top Twitter Queries Issued

<table>
<thead>
<tr>
<th>Web</th>
<th>Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>twitter</td>
<td>new moon</td>
</tr>
<tr>
<td>youtube</td>
<td>#youknowyouruglyif</td>
</tr>
<tr>
<td>facebook</td>
<td>justin bieber</td>
</tr>
<tr>
<td>google</td>
<td>adam lambert</td>
</tr>
<tr>
<td>myspace</td>
<td>#theresway2many</td>
</tr>
<tr>
<td>youtube.com</td>
<td>taylor swift</td>
</tr>
<tr>
<td>yahoo</td>
<td>lady gaga</td>
</tr>
<tr>
<td>ebay</td>
<td>modern warfare 2</td>
</tr>
<tr>
<td>craigslist</td>
<td>thanksgiving</td>
</tr>
<tr>
<td>myspace.com</td>
<td>#wecoolandallbut</td>
</tr>
</tbody>
</table>

- People-focused
- Temporal aspects
- Specialized syntax
People and Time in Twitter Queries

• Lots of celebrity names
  – *lady gaga*

• Celebrities unlikely to just be part of a query
  – *lady gaga meat dress*

• Many references to individual user accounts

• Hashtags common
  – Memes, trending topics

<table>
<thead>
<tr>
<th></th>
<th>Web</th>
<th>Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is a celebrity name</td>
<td>3.1%</td>
<td>15.2%</td>
</tr>
</tbody>
</table>
Twitter Syntax: @ and #

- Specialized syntax very common for Twitter
- @ and # reduce ambiguity like advanced query operators
- Important differences:
  - Part of content creation
  - Hashtag queries often issued via a click

<table>
<thead>
<tr>
<th></th>
<th>Web</th>
<th>Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is a celebrity name</td>
<td>3.1%</td>
<td>15.2%</td>
</tr>
<tr>
<td>Mentions a celebrity</td>
<td>14.9%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Contains @</td>
<td>0.1%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Is a username without @</td>
<td>0.0%</td>
<td>2.4%</td>
</tr>
<tr>
<td>Contains #</td>
<td>0.1%</td>
<td>21.3%</td>
</tr>
<tr>
<td>Is a hashtag without #</td>
<td>3.0%</td>
<td>4.4%</td>
</tr>
</tbody>
</table>
Twitter Query Popularity

• Hashtag queries particularly popular
  – Most popular queries: Hashtag 51% of the time
  – Least popular queries: Hashtag 7% of the time

• Celebrity queries particularly popular
  – Most popular queries: Celebrity 25% of the time
  – Least popular queries: Celebrity 4% of the time

• Twitter queries less diverse than Web queries
  – Only 1 in 4 unique (v. 2 in 4 unique)
Temporal Patterns on Twitter

• Individuals repeat the same query on Twitter
  – 35% of Web queries are repeat
  – 56% of Twitter queries are repeat

• But sessions are shorter

<table>
<thead>
<tr>
<th></th>
<th>Web</th>
<th>Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of queries in session</td>
<td>2.9</td>
<td>2.2</td>
</tr>
<tr>
<td>Number of unique queries in session</td>
<td>2.67</td>
<td>1.5</td>
</tr>
<tr>
<td>Seconds between queries in session</td>
<td>13.6</td>
<td>9.4</td>
</tr>
</tbody>
</table>

> Twitter queries used for monitoring
Cross-Corpus Behavior

• Some users issued same query to Twitter & Web
• Overlapping queries highly informational
• Web used to explore
  – Overlapping query is 8 times more likely to appear in another query
• Twitter used to monitor
  hong kong
  weather in hong kong
  hong kong restaurants
  hong kong
  …
  hong kong
  …
  sheraton hong kong
How Microblog and Web Search Differ

- Time important
- People important
- Specialized syntax
- Queries common
- Repeated a lot
- Change very little

- Often navigational
- Time and people less important
- No syntax use
- Queries longer
- Queries develop