Towards Supporting Search over Trending Events with Social Media

Sanjay Kairam, Stanford University
Meredith Ringel Morris, Jaime Teevan, Dan Liebling, & Susan Dumais, MSR
Trending Events
Trending Events

Events that serve as novel or evolving sources of widespread online activity
simpson trial verdict
“Follow our liveblog of the Simpson Trial...”

“did u hear oj simpson is on trial?”

“the simpson trial is rigged!”
United States Trends · Change

Walmart
#stlcards
New York
Sounds
Texas
Kevin Hart
San Francisco
Netflix
Florida
#Egypt
Follow our liveblog of the Simpson Trial...

The Simpson Trial is rigged!

Did you hear OJ Simpson is on trial?
Related Work
Related Work

• Temporal Patterns in Search [ADAR 2007, KULKARNI 2011]
Related Work

• Temporal Patterns in Search [Adar 2007, Kulkarni 2011]

• Trends in Social Media [Kwak 2010, Zubiaga 2011]
Related Work

• Temporal Patterns in Search [ADAR 2007, KULKARNI 2011]

• Trends in Social Media [KWAK 2010, ZUBIAGA 2011]

• Queries over Social Resources:
Related Work

• Temporal Patterns in Search [Adar 2007, Kulkarni 2011]

• Trends in Social Media [Kwak 2010, Zubiaga 2011]

• Queries over Social Resources:
  • People, named entities, & temporally-relevant [Teevan 2011]
Related Work

- Temporal Patterns in Search [ADAR 2007, KULKARNI 2011]
- Trends in Social Media [KWAK 2010, ZUBIAGA 2011]
- Queries over Social Resources:
  - People, named entities, & temporally-relevant [TEEVAN 2011]
  - Churn rates up to 4x higher than search engines [LIN & MISHNE 2012]
Research Questions
Research Questions
Research Questions

1. How does search behavior vary with prior awareness of a trending event?
Research Questions

1. How does search behavior vary with prior awareness of a trending event?
2. What trends where, and when?
Research Questions

1. How does search behavior vary with prior awareness of a trending event?
2. What trends where, and when?
3. How does interest in an event change over time?
Data
Data

218 *Trending Queries* 

+ 

113 *Twitter Trends* 

= 

331 *Trending Events*
Data
Data

The Simpson Trial
Data

The Simpson Trial
Data

simpson trial
simpson trial

Data

- o j simpson trial
- black glove simpson trial
- simpson murder trial
- simpson trial verdict
Data

-o j simpson trial
-black glove simpson trial
-simpson murder trial
-simpson trial verdict

“Follow our liveblog of the Simpson Trial…”

“the simpson trial is rigged!”

“did u hear oj simpson is on trial?”
Data

simpson trial

- o j simpson trial
- black glove simpson trial
- simpson murder trial
- simpson trial verdict

Two Weeks

"Follow our liveblog of the Simpson Trial…"
"the simpson trial is rigged!"
"did u hear oj simpson is on trial?"
Data

- 5 Days
- 288 Turkers
- 49 Events

Query Volume: 'recall toyota' (Log−scaled)
Results

RQ1: How does search behavior vary with prior awareness of an event?
Results

**RQ1:** How does search behavior vary with prior awareness of an event?
RQ1: How does search behavior vary with prior awareness of an event?
Results

RQ1: How does search behavior vary with prior awareness of an event?
Results

**RQ1:** How does search behavior vary with prior awareness of an event?
RQ1: How does search behavior vary with prior awareness of an event?
Results

“Manual” Queries

- simpson murder trial
- o j simpson trial

Already aware of event

“Link” Queries

- Prompted by search engine

O.J. Simpson trial

RQ1: How does search behavior vary with prior awareness of an event?
RQ1: How does search behavior vary with prior awareness of an event?

<table>
<thead>
<tr>
<th>Entry Point</th>
<th>% Click Instant Answer</th>
<th>% Click Standard Result</th>
<th>Click Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link</td>
<td>17.98%</td>
<td>4.64%</td>
<td>2.93</td>
</tr>
<tr>
<td>Manual</td>
<td>31.73%</td>
<td>29.28%</td>
<td>4.13</td>
</tr>
</tbody>
</table>
RQ1: How does search behavior vary with prior awareness of an event?
RQ1: How does search behavior vary with prior awareness of an event?

Results

- User behavior varies in relation to prior awareness
- Real-time and expert commentary are key
- Especially for users new to an event.
Results

*RQ2: What trends where, and when?*
Results

<table>
<thead>
<tr>
<th>Trending Queries</th>
<th>Twitter Trends</th>
</tr>
</thead>
<tbody>
<tr>
<td>News</td>
<td>News</td>
</tr>
<tr>
<td>Entertainment</td>
<td>Entertainment</td>
</tr>
<tr>
<td>Politics</td>
<td>Politics</td>
</tr>
<tr>
<td>Celebrity</td>
<td>Celebrity</td>
</tr>
<tr>
<td>Sports</td>
<td>Sports</td>
</tr>
<tr>
<td>Holiday</td>
<td>Holiday</td>
</tr>
<tr>
<td>Deaths</td>
<td>Deaths</td>
</tr>
<tr>
<td>Unknown</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

*RQ2: What trends where, and when?*
Results

Trending Queries
News
Entertainment
Politics
Celebrity
Sports
Holiday
Deaths
Unknown

Twitter Trends

RQ2: What trends where, and when?
RQ2: What trends where, and when?

Results

Trending Queries
- News
- Politics

Twitter Trends
- Entertainment
  - Celebrity
  - Sports
  - Holiday
  - Deaths
  - Unknown
Results

Trending Queries
- News
- Politics

Twitter Trends
- Celebrity
- Sports
- Holiday
- Deaths
- Unknown

RQ2: What trends where, and when?
Results

RQ2: What trends where, and when?

Search Activity

Social Media

Query Volume: 'recall toyota' (Log−scaled)

Tweet Volume: 'dame notre uniform' (Log−scaled)
Results

RQ2: What trends where, and when?

Search Activity

Social Media

Query Volume: 'recall toyota' (Log-scaled)

Tweet Volume: 'dame notre uniform' (Log-scaled)
Results

RQ2: What trends where, and when?

Distribution of Delays of Search Activity from Social Media
Results

• Different topics ‘trend’ in different streams
• Sharing activity leads searching activity by a few hours

RQ2: What trends where, and when?
Results

RQ3: How does interest in an event change over time?
RQ3: How does interest in an event change over time?
RQ3: How does interest in an event change over time?
Results

RQ3: How does interest in an event change over time?
**RQ3:** How does interest in an event change over time?
Results

- Topical interest diverges around period of peak activity
- Overlap between content shared and searched increases, as well.

*RQ3: How does interest in an event change over time?*
Limitations

• Limited view of trends (filtering)
• Limited collection of query and tweet content (matching)
• Limited analysis of events
• Assumptions about user behavior from logs
Summary

• Users want real-time content, especially when they are new.
• Social media activity tends to precede search activity.
• While topical interest diverges during peak, overlap increases, as well.
related stories on msnNOW

See ballplayers' priceless reaction to female gymnast's amazing pitch

2 days ago

Mrs_Officer1 @Mrs_Officer1
Me gustó un video de @YouTube youtu.be/RVeO6QMcbW4?a South Korean rhythmic gymnast Shin Soo-ji's first pitch

Marcio612 Marcio Bispo
Gostei de um video @YouTube youtu.be/RVeO6QMcbW4?aS... Korean rhythmic gymnast Shin Soo-ji's first pitch

RumenGanev Rumen P Ganev
I guess we should've watched this video after all...@s2kev @JacobWarren youtube.com/watch?v=RVeo6Q...

MesutChobitsu Hideki
Me ha gustado un video de @YouTube (youtu.be/RVeO6QMcbW4?a South Korean rhythmic gymnast Shin Soo-ji's first pitch).

bbolan1 Brian Bolan
@robertAbooeys slightly better than your attempts at a first pitch. youtube.com/watch?v=RVeo6Q...
Thanks!

- Sanjay Kairam
- Email: sanjay.kairam@gmail.com
- Twitter: @skairam