@martinberg
Martin Berg

I can haz qualitative research? #icwsm

57 minutes ago via Twitter for iPhone  Favorite  Retweet  Reply

You sure can!
What is Twitter?

Twitter is a service for friends, family, and co-workers to communicate and stay connected through the exchange of quick, frequent answers to one simple question: What are you doing?

Get Started—Join!

© 2009 Twitter  About Us  Contact  Blog  Status  Apps  API  Search  Help  Jobs  Terms  Privacy
See who's here

Top tweets View all

**TotallyJBiebs** Don't forget we're unfollowing Justin tonight for Apriks Fools Day. You can refollow after tomorrow, RT&Spread the word ^_^ 2 hours ago

**AusNeedsJB** Don't forget we're unfollowing Justin tonight for Apriks Fools Day. You can refollow after tomorrow, RT&Spread the word ^_^ about 1 hour ago

**joshmeatsix** Great idea guys, let's make #happybirthdayflinttown att!! One love x about 1 hour ago

**SUAREASY** Yooo! @infinityony #HappyBirthdayRylandRT it! 7 hours ago
How do I follow my Interests?
I Warned U Not To Cross Teh Streams.
Yet more Data
Meta Data, Profile Data, Linked Data
Yet more Data

Meta Data, Profile Data, Linked Data
Any of it Useful?

“I think the challenge not only for twitter, but for the technology industry at large. Is building more relevant filters, in real time. Like being able to surface valuable information immediately. No matter who it is, who’s listening or who’s broadcasting, is a really really hard problem, and it makes twitter alot more meaningful[... ]We’ve gotten really really good at being able to put content in, into media [...] getting it out in a relevant, valueable way, in real time is still very difficult.”

- Jack Dorsey (Creator of Twitter)
Any of it Useful?
Who cares how much data there is!

“I think the challenge not only for twitter, but for the technology industry at large. Is building more relevant filters, in real time. Like being able to surface valuable information immediately. No matter who it is, who’s listening or who’s broadcasting, is a really really hard problem, and it makes twitter alot more meaningful[... ]We’ve gotten really really good at being able to put content in, into media [...] getting it out in a relevant, valueable way, in real time is still very difficult.”

- Jack Dorsey (Creator of Twitter)
Why Twitter?
Where is the value?
Lets go back...
Lets go back...
Great Scott!
Asking Friends

Hey, what are you doing?
Social Search
What is everyone else doing?

you & me

Tuesday, 19 July 2011
Social Search

What is everyone else doing?
Existing Knowledge

No need to reinvent the wheel


Tuesday, 19 July 2011
Existing Knowledge
No need to reinvent the wheel

Let's go back to the network

Remember...

you & me
and if we take a step back...
Please mind the gap

friend

friend

friend

friend

you

me

Tuesday, 19 July 2011
and if we take a step back...

Please mind the gap
and if we take a step back...

Please mind the gap
We start to see interesting things...
We start to see interesting things...
Which have value!
Which have value!
Location, experiences, temporal data

Yardi, Sarita and Boyd, Danah. ICWSM 2010.
Tweeting from the Town Square: Measuring Geographic Local Networks

http://www.flickr.com/photos/gcaw/5445225362/
http://www.flickr.com/photos/24423474@N08/4999891492/
http://www.flickr.com/photos/24423474@N08/4999891492/
http://www.flickr.com/photos/gcaw/5445225362/
Location, experiences, temporal data

Political upheaval, emergency events .. so what are you tweeting now?

Yardi, Sarita and Boyd, Danah. ICWSM 2010. Tweeting from the Town Square: Measuring Geographic Local Networks

Tuesday, 19 July 2011
Twitter Search
How do you find useful information?
Displaying Results

Realtime
Displaying Results
Time, ReTweets, Location, Popularity?

http://www.flickr.com/photos/publicenergy/394124407/

Tuesday, 19 July 2011
Displaying Results

Time, ReTweets, Location, Popularity?
Displaying Results

Time, ReTweets, Location, Popularity?

http://www.flickr.com/photos/publicenergy/394124407/

Tuesday, 19 July 2011
Displaying Results
Making sense of the data.
Displaying Results

Making sense of the data.
Displaying Results
Making sense of the data.
Displaying Results
Making sense of the data.

Displaying Results
Making sense of the data.
Interestingness

Not necessarily useful!


http://www.flickr.com/photos/wwarby/2460655511/

Tuesday, 19 July 2011
How we are different?

What makes us unique?
Finding Usefulness!

What constitutes a useful Tweet?

http://www.flickr.com/photos/eddoddie/4346349664/

Tuesday, 19 July 2011
Finding Usefulness!

What constitutes a useful Tweet?
The Method

How did we go about this?
Information Seeking

3 Information Seeking Tasks

20 Participants

They were really nice people!
Search Interface
A simple, easy to understand interface
Results for: icwsm

gihuili: RT @alsothings: I forked @jakehofman's python script to grab the full #icwsm proceedings + metadata. Works with non-ascii encodings now http://t.co/Yz7xb1 just now from Twitter for Mac

有用

 gingdottwit: @icwsm is someone more, equal, or less likely to be infected by a spread if more than one of their friends is infected? A: more. #icwsm
1 minute ago from Twitter for iPad

有用

 thetarro: RT @ricardoaraujo: Glancing over laptops of #icwsm audience shows that not many people are browsing Facebook. Twitter and Google+ dominate.
3 minutes ago from web

有用

 gingdottwit: Lots of comparisons between observed network graphs and random graphs to see what's emergent. at #icwsm
4 minutes ago from Twitter for iPad

有用

 gingdottwit: @icwsm most cascades, or epidemics, are 1% size of network, using Digg data. Why? They investigated. #icwsm
5 minutes ago from Twitter for iPad

有用

 alsothings: excellent intro to comp-net theory of epidemics by Greg Ver Steeg as a ramp up to What Stops Social Epidemics http://t.co/s7qow1H #icwsm
6 minutes ago from Twitter for Mac

有用

 gingdottwit: @flwrsn @lintool because you only it if you are following @icwsm unless you

有用
Think aloud + Interviews

To help us provide more insight

It's useful because...

I didn't because...
Analysis

Lots and lots of it!
Grounded Theory

Inductive Coding = Lots of Post-its!


Tuesday, 19 July 2011
Kappa Analysis

Cohen... Fleiss....

<table>
<thead>
<tr>
<th></th>
<th>Useful Tweets Data Set</th>
<th>Not Useful Tweets Data Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Author 1</td>
<td>Author 2</td>
</tr>
<tr>
<td>Author 1</td>
<td>-</td>
<td>0.5065</td>
</tr>
<tr>
<td>Author 2</td>
<td>0.5065</td>
<td>-</td>
</tr>
<tr>
<td>Independent Coder</td>
<td>0.5097</td>
<td>0.4607</td>
</tr>
</tbody>
</table>

Extended Kappa Analysis

Multi Coded Kappa

0.73 (Substantial Agreement) Between Evaluators
&
0.62 (Substantial Agreement) with Independent Untrained Coder

What did we find?

Useful & Not-Useful
<table>
<thead>
<tr>
<th>In Tweet Content</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>Someone reporting a personal experience, but not necessarily suggestion / direction.</td>
</tr>
<tr>
<td>Direct</td>
<td>Someone making a direct recommendation, but not necessarily relaying a personal experience.</td>
</tr>
<tr>
<td>Social Knowledge</td>
<td>Containing information that is spreading socially, or becoming general knowledge.</td>
</tr>
<tr>
<td>Specific Information</td>
<td>Where facts are listed directly in tweets e.g. prices, times etc.</td>
</tr>
<tr>
<td>Reflection on Tweet</td>
<td></td>
</tr>
<tr>
<td>Entertaining</td>
<td>The reader finds them amusing.</td>
</tr>
<tr>
<td>Shared Sentiment</td>
<td>The reader agrees with the author of the tweet.</td>
</tr>
<tr>
<td>Relevant</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>The time is current</td>
</tr>
<tr>
<td>Location</td>
<td>The location is relevant to the query.</td>
</tr>
</tbody>
</table>
# Useful (cont.)

| **Trust** |  |
| --- |  |
| **Trusted Author** | The twitter account has a reputation / following |
| **Trusted Avatar** | The visual appearance cultivates trust. |
| **Trusted Link** | A link to a trustworthy recognisable domain. |

| **Links** |  |
| --- |  |
| **Actionable Link** | The user can perform a transaction by using the link (heavily dependent on trust) |
| **Media Link** | The link is to rich multimedia content. |
| **Useful Link** | The link provides valuable information content, e.g. authoritative information, educated reviews |

<p>| <strong>Meta Tweet</strong> |  |
| --- |  |
| <strong>Retweeted Lots</strong> | Its information that others have passed on lots |
| <strong>Conversation</strong> | Its part of a series of tweets, and they all need to be useful |</p>
<table>
<thead>
<tr>
<th>Tweet Content</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Information</td>
<td>Absence of anything, event, factual points</td>
</tr>
<tr>
<td>Introspective</td>
<td>Personal content and personal thoughts for no social benefit</td>
</tr>
<tr>
<td>Off Topic</td>
<td>Result not related to the query give / TF-IDF irrelevant</td>
</tr>
<tr>
<td>Too Technical</td>
<td>The content requires specific domain knowledge the reader doesn’t possess</td>
</tr>
<tr>
<td>Poorly Constructed</td>
<td>Tweets that may have grammatical / spelling errors, or malformed URLs.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bad Tweets</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPAM</td>
<td>Irrelevant or inappropriate messages</td>
</tr>
<tr>
<td>Wrong Language</td>
<td>Messages sent in a foreign language of that to the reader</td>
</tr>
<tr>
<td>Dead Link</td>
<td>A URL which does not work i.e. a 404</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Not Relevant</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Out of date content</td>
</tr>
<tr>
<td>Location</td>
<td>Wrong geographic location</td>
</tr>
</tbody>
</table>
### Not Useful (cont.)

<table>
<thead>
<tr>
<th>Trust</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Un-trusted Author</td>
<td>An author the reader feels at uneased by or suspicious of.</td>
</tr>
<tr>
<td>Un-trusted Link</td>
<td>A link the reader feels is suspicious</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subjective</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perspective Oriented</td>
<td>A tweet that is perspective centric, meaning the author is providing their view or projecting an attitude on a subject matter or to a subject / reader.</td>
</tr>
<tr>
<td>Disagree with Tweet</td>
<td>A conflict of agreement between the reader and the author</td>
</tr>
<tr>
<td>Not Funny</td>
<td>A tweet that is aimed to be humorous, which the reader does not feel is humorous.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Meta Tweet</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>QnA</td>
<td>Part of a conversation, reader desires the whole convo. not just the question or the answer.</td>
</tr>
<tr>
<td>Repeated</td>
<td>Content the reader has seen before.</td>
</tr>
</tbody>
</table>
Insights
Interesting finds

http://www.flickr.com/photos/foxmulderven/3063598624

Tuesday, 19 July 2011
The Possible Impact

Where could we see the impact of this work?
Search System
A work in progress

Tuesday, 19 July 2011
Conclusions

So just remember.
Thank you for Listening

Jonathan Hurlock
@jonhurlock

Max L. Wilson
@gingdottwit

Like the talk? Then please tweet it, by quickly visiting:
http://moourl.com/LikedTheTalk