Predicting Events based on Social Media Exposure: from the Arab Spring to Disease Outbreaks

Rishab Aiyer Ghosh
Co-Founder & Chief Scientist

TOPSYYLABS
Headlines, early 2011

Egyptians Defiant as Military Stands By Intelligence Chief Is Mubarak’s First Vice President

Some in Army Fear Repeal of Law on Protests

World pressure on Mubarak grows

Dutch cut Iran links over hanging

Irish Senate passes finance bill

Davos 2011

Everying

News

Shopping

Books

Paxos, CA

Change location

News for egypt

Egypt’s Tourism: The Army’s OK with the Protestors, for Now

Parts of Egypt’s capital city seemed almost normal on Saturday as the writers of Friday’s violence and shootings faded. After a night marked by gunfire and looting.

Published 9:12 AM ET

Updated 10:31 AM ET
Google, Twitter search :-(

{ not realtime, not fresh - or not ranked, not relevant }
Topsy search :-) 
{ realtime, fresh and relevant }

search result from 2011-10-11 at 15.40 EDT
{ validation : real people! }
underlying \textit{shift} in \textit{information dissemination}
Most intelligence historically was from open sources
For financial, business, brand, publishing, government use...

Previously open sources, public sources, were *published* sources: news in *publications*
Public conversations were private information

Human intelligence - inside information - insider trading
Public conversations are public information

Publicly accessible
we can all be

in tahrir square
we can all be

in #tahrir square

{ public conversations are publicly accessible }
Search results 1-20 out of 673 photos about #bahrain
{ public conversations are publicly accessible }

we can all have

reporters

on the ground
{ public conversations are publicly accessible }

we can all have reporters on the ground signal / noise
public conversations are publicly accessible

we can all have reporters on the ground

signal / noise

Topsy’s platform provides editing, curation...
{ historical context for back-testing... }
What evidence is there of failed planned protests?

{ ... helps predict future events with quantified confidence }

Share Of Voice Over Time
Mar - Protest Dates
Compared to Google Flu Trends, etc...

What people search for... (“Pinterest”)

Mirrors their public conversations

(source: [http://adaptmedia.clarify-it.com/d/ngpj7j](http://adaptmedia.clarify-it.com/d/ngpj7j))
Unlike search records, which are based on private IP addresses, private search query logs...

There are many more ways you can analyze public conversations.
Track activity:
- Mentions
- Exposure

Discover:
- Related new terms

Monitor conversations:
- Posts
- Articles shared
- Who is saying what?

See shared media ranked by influence.
Map activity for terms:
Raw mentions
Share of all social mentions

See conversations from particular locations
Extensible to all types of social content (Twitter, Facebook, blogs, etc.)

Fast Indexing - pipeline architecture capable of indexing 1M documents per second

Live Ranking - new citations impact search results/ranking in realtime

Native multi-language support
social media as human intelligence

Rishab Aiyer Ghosh Co-Founder & Chief Scientist: rishab@topsy.com