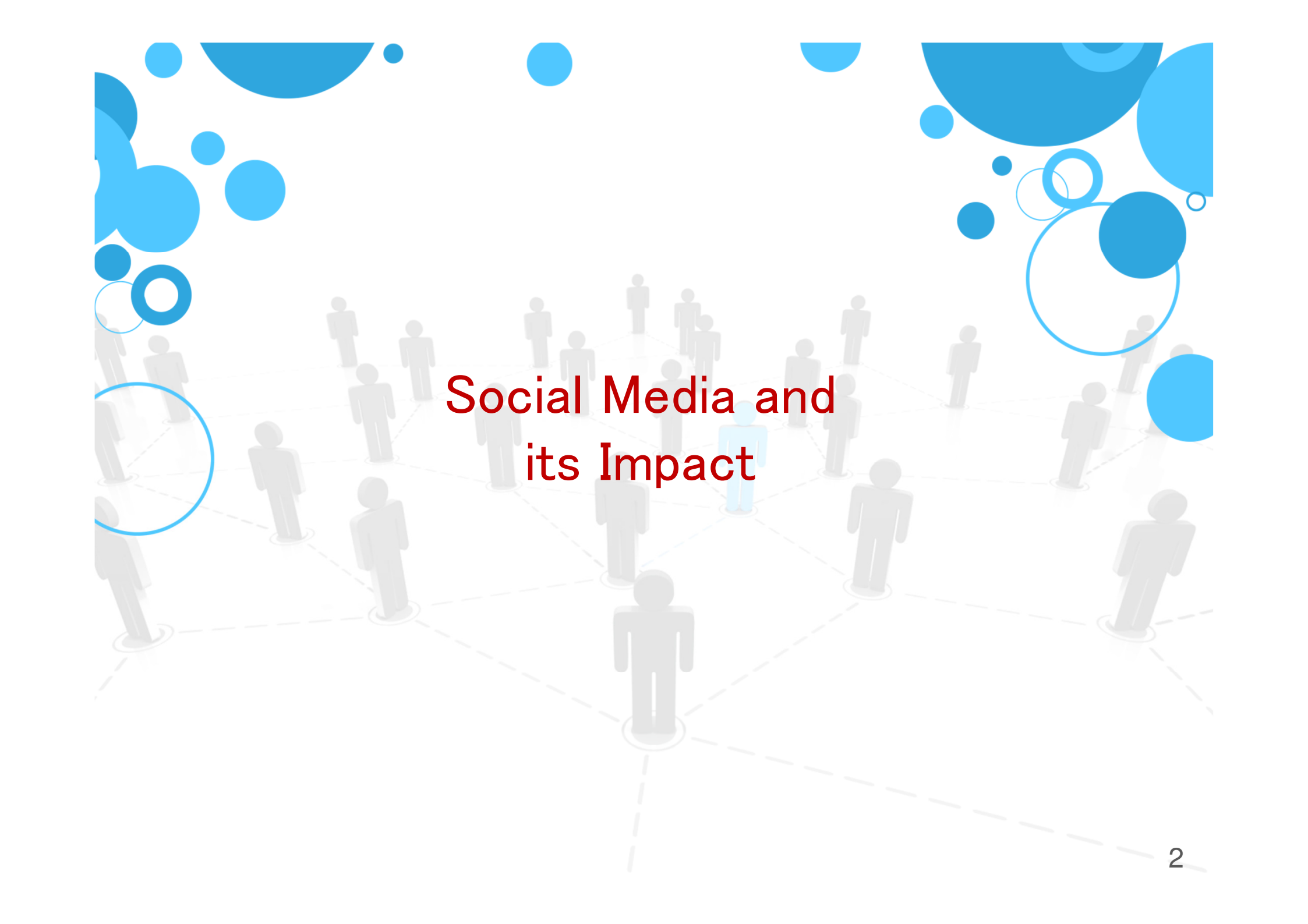




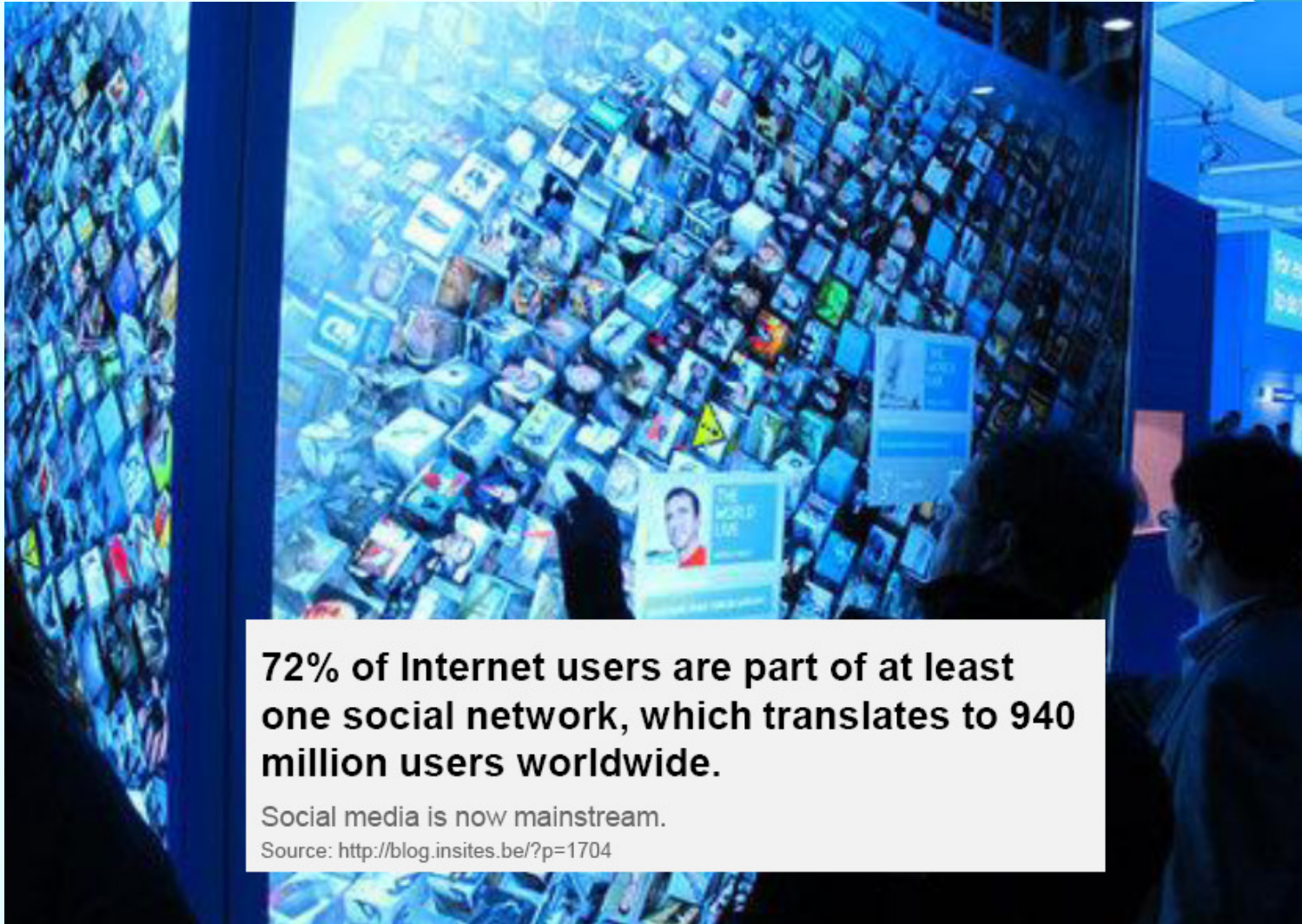
# Mining Opinion from Twitter

Lo Siaw Ling  
School of Information Technology  
Nanyang Polytechnic  
Singapore

The background features a network of stylized human figures in various shades of gray, connected by dashed lines. The figures are arranged in a circular pattern, with one figure in the center highlighted in a light blue color. The overall theme is social connectivity and digital communication. The text "Social Media and its Impact" is centered in a bold, red font.

# Social Media and its Impact

# How Big is Social Media?

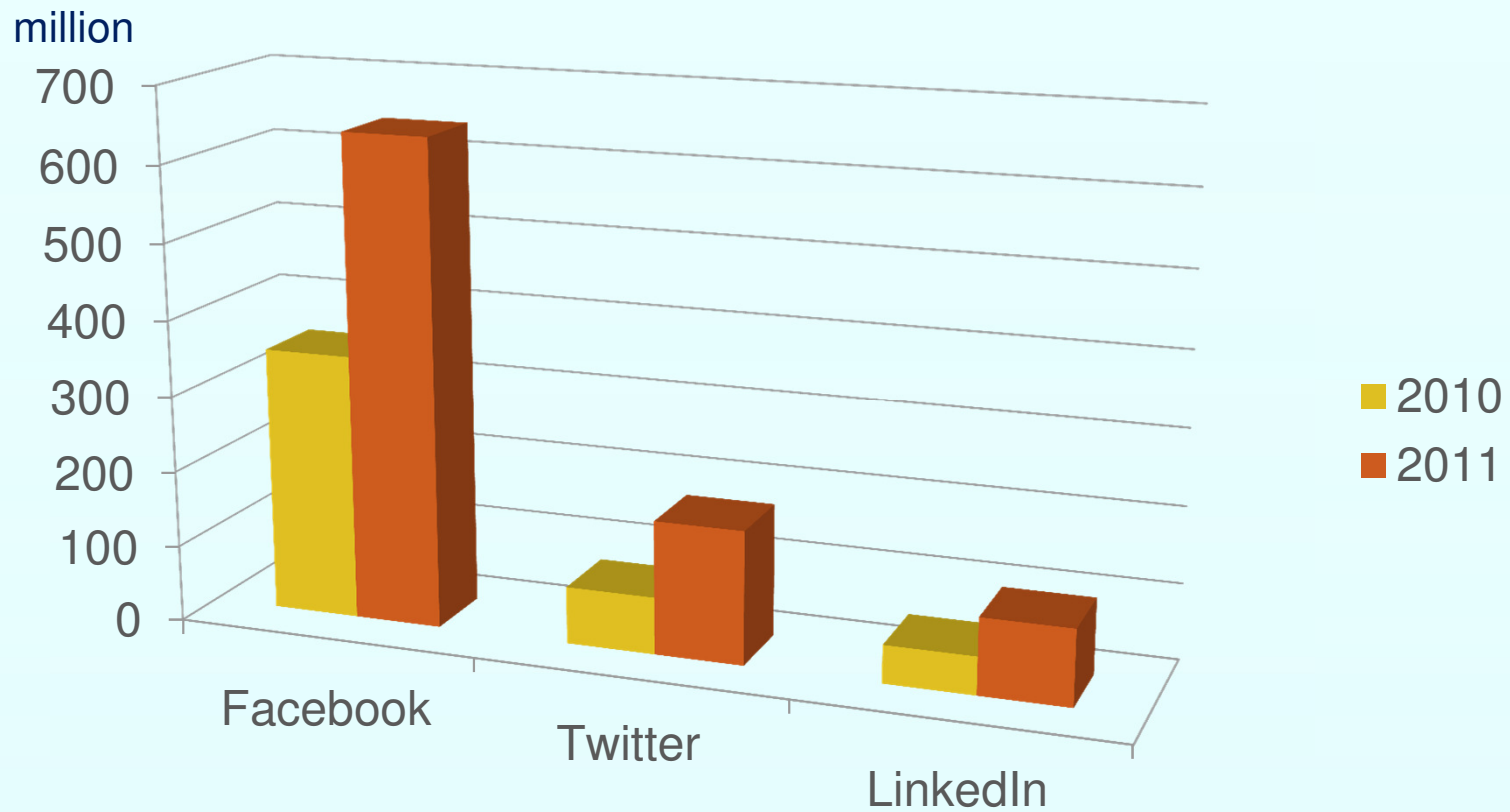


**72% of Internet users are part of at least one social network, which translates to 940 million users worldwide.**

Social media is now mainstream.

Source: <http://blog.insites.be/?p=1704>

# Social Media Growth



<http://www.browsermedia.co.uk/2011/03/30/2011-social-media-statistics-show-huge-growth/>

# Social Media

*Lots more sharing,  
talking and opinions!*



**More than 75 percentage of businesses using social media**

- 14 Jul 2011 by BusinessWire \*

**Majority of Fortune 100 companies using more than one social media platforms**

- 17 Feb 2011 by Burson Marsteller \*\*

# Landscape

“Social Media Marketing Spending to Hit \$3.1 Billion by 2014 (faster than any other form of online marketing)”

**Forrester, 15 September 2009**

“Only 14% of people trust advertisers yet 78% of consumers trust peer recommendations.”

**Erik Qualman's book Socialnomics, 2009**

# Impact of Social Media



# Impact of Social Media on Product

## GM Cancels 'Hideous' Buick SUV After Would-Be Customers Twitter!

**ONE week** after announcing a new Buick SUV Christopher Barger, GM's spokesman for social media said:

The decision was based on customers' input - Face-to-face, **blogs and tweets**. No matter how they expressed it "they just didn't like it."

<http://www.bloomberg.com/apps/news?pid=newsarchive&sid=aHsoNjdHUQLY>



# Impact of Social Media on Product

## GM Cancels 'Hideous' Buick SUV After Would-Be Customers Twitter!



# Impact of Social Media on Product



- Del Monte created a new “hot-selling” dog food snack in 6 weeks
- Used a **social community to source for creative ideas** and create a new product
- Demonstrates the potential power of social media marketing to influence product sales

[http://www.youtube.com/watch?v=yP\\_3bpCPZaQ](http://www.youtube.com/watch?v=yP_3bpCPZaQ)

# Impact of Social Media on Organisation

## Nestlé vs Greenpeace Palm Oil from Destroyed Rainforest

- Nestlé persuaded YouTube to remove Greenpeace's video
- Storm ensued on Nestlé **Facebook** page
- **TWO months** later, Nestlé announced a “zero deforestation” policy in partnership with The Forest Trust (TFT)

*“Social media: as you can see we're learning as we go.  
Thanks for the comments.”*

# Impact of Social Media on Organisation

## Nestlé vs Greenpeace Palm Oil from Destroyed Rainforest



# Impact of Social Media on Government

25<sup>th</sup> Jan 2011

**Egypt Blocked Twitter and Facebook!**



Egyptian protesters have openly thanked social media's role in the revolution against the country's ruling government.

# Impact of Social Media on Government

12<sup>th</sup> August 2011

**Herald Sun**  
Stories start here.

Meanwhile, British Prime Minister Cameron has given the option to disable social media sites.

Speaking in parliament, Mr Cameron said the government would consider violence after nights of riots and street mayhem erupting again.

Mr Cameron told MPs he was sending police to raid houses to round up suspects.

He said the government, police and the public should focus on the use of social media sites to spread disorder.

August 12, 2011

**THE HUFFINGTON POST**  
UNITED KINGDOM 



**Gordon MacMillan**

*Social media editor, Brand Republic Group*

GET UPDATES FROM GORDON MACMILLAN



## In the Riot Clean-up Social Media shows Itself as a Force of Social Good

Posted: 10/8/11 10:06 GMT



The background features a network of stylized human figures in grey, connected by dashed lines, with one figure highlighted in light blue. The top corners are decorated with various blue circles and semi-circles of different sizes. The text "Social Media Analytics" is centered in a bold, red, sans-serif font.

# Social Media Analytics

# Data and Metrics



Peter Drucker:

“That which is not measured is not managed.”

## Quantitative Metrics VS Qualitative Metrics



# Key Measurement Goals



# Reach & Buzz

aka getting attention....

## Reach

Size of your audience

How many saw your message

E.g. Twitter followers

Just implies popularity?

## Buzz

2 types:

Conversation Focus vs. Content Focus

# Identifying Influence

aka getting respect....

Reach & Buzz: Subscribers, Readers, Fans

**Influence:** REPEATERS

Your message is valuable when it is repeated  
aka “Honourable mentions”

High probability of others referencing & reproducing what  
you say

E.g. Twitter: reply/mention (@xxx) & retweet (RT)

# Recognising Sentiment

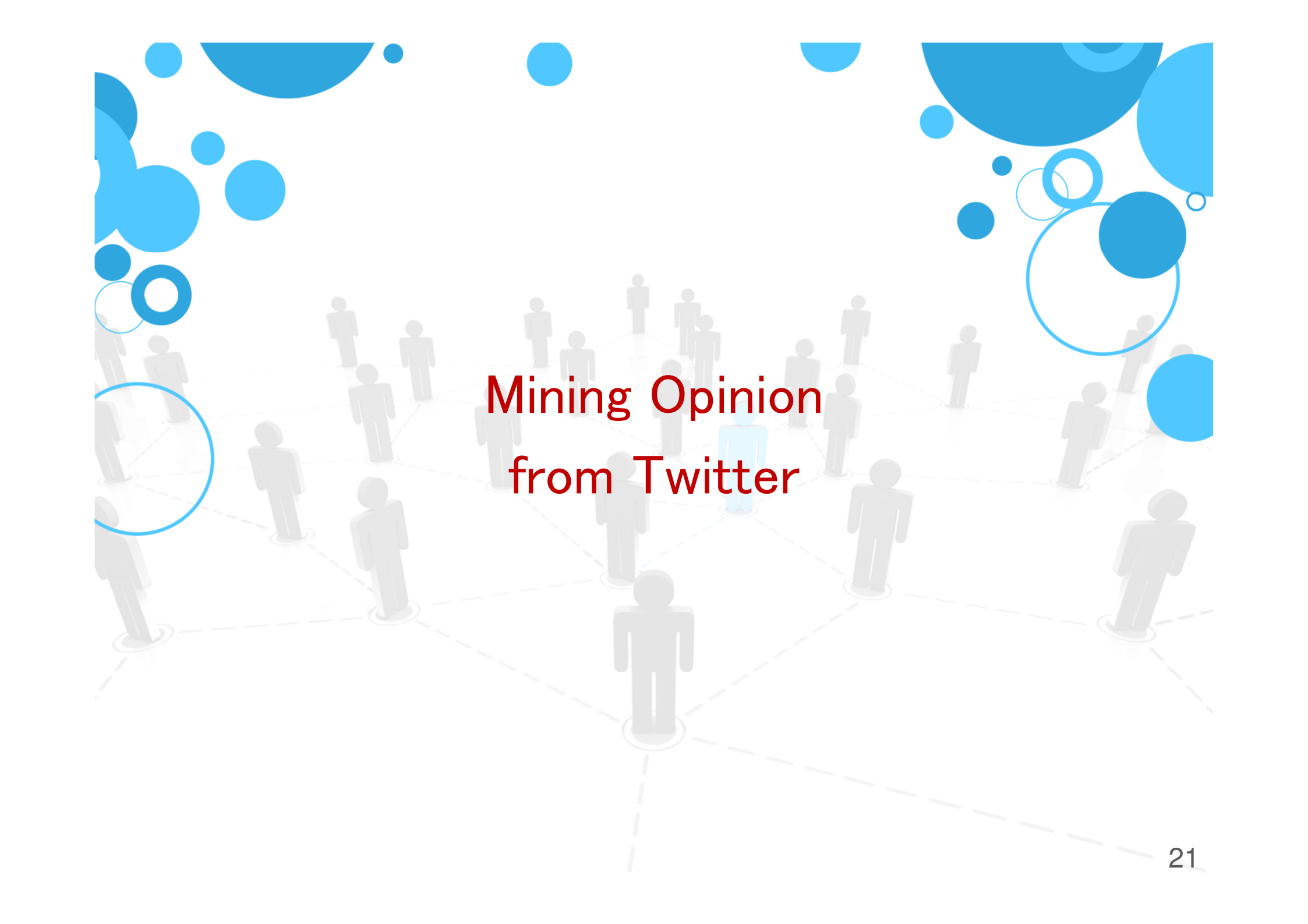
aka getting emotional....

## **Polarity**

Positive, Negative or Neutral

## **Intensity**

Degree of emotion -- !!!!



# Mining Opinion from Twitter

# Background

- Singapore Government National Funded Project
- Objective
  - Research and develop **better algorithms** in addressing the accuracy in sentiment analysis and discovering the underlying online relationship for influential study
  - Capability development for staff and students

# The platform and data

Twitter - one of the fastest growing social media platforms – 200 millions users and growing...

2007 – 0.5 million

2009 – 58 million

2011 – 200 million

It is **open** and influential – to company and even country.



Jan 26, 2011

## The latest on Twitter in Egypt

04:54 PM  Share  7  Yahoo! Buzz  Share  Print

By [Jon Swartz](#), USA TODAY

 1 Comments  
 2 Recommend

Twitter has acknowledged it is being blocked in Egypt because of massive protests in that economically battered country arranged through social-media sites like it and Facebook.

# Influential Analysis

Retweet and the Reply features of Twitter is used to enable real-time study

$$\text{Influential Index} = \frac{n(\text{Reply}) + n(\text{Retweet})}{n(\text{Tweet})}$$

For example, a tweet - Verizon will launch iPhone 4 on 10 Feb sent by user ABC

Retweet (think of it as forwarding)

**RT @ABC** Verizon will launch iPhone 4 on 10 Feb

Reply

**@ABC** thanks... I will be there to get one



# Sentiment Analysis

Twitter only allows 140 characters message  
Information shared can be

## Factual

Verizon will launch iPhone 4 on 10 Feb

## Advertisement

Win an iPad2 by visiting this website

## Opinions

Verizon iPhone 4 has the same short battery life as AT&T, no improvement  
I can't wait to get my hand on an iPhone 4... yippee

# Challenges

## Complexity of human language

Taken in whole context-

“'Sinful' is a good thing when applied to chocolate cake.”

“Great!!!! It rained! What should we do now?”

**Contextual sentiment** (*for example, double negative is positive in English*) and the role of content-word negator (*for example, eliminated has negative sentiment*)

**Human expression** – every generation and/or culture use different way to express (*for example, LOL*)

# Implementation

## Sentiment Analysis

A *combination approach* using Support Vector Machine [SVM](sigmoid kernel), SentiWordNet and key phrases is implemented

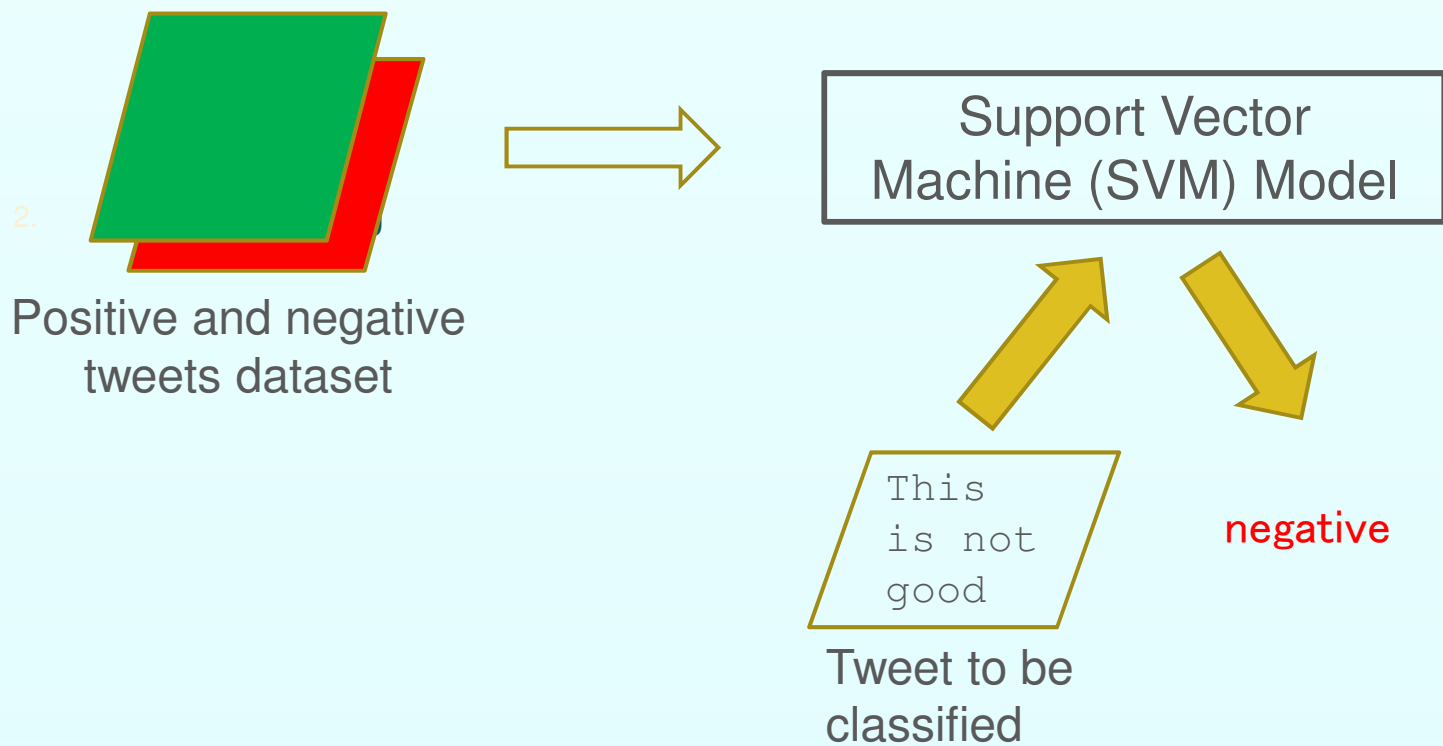
**SVM** – common approach in sentiment analysis

**SentiWordNet** – built on WordNet and annotate with positivity, negativity (e.g., *'better' has a more positive value than 'good'*)

**Key phrases** – list of positive and negative word(s) extracted from internet (e.g., *looking forward to, fallen behind*)

# Support Vector Machine

## 1. Build a model

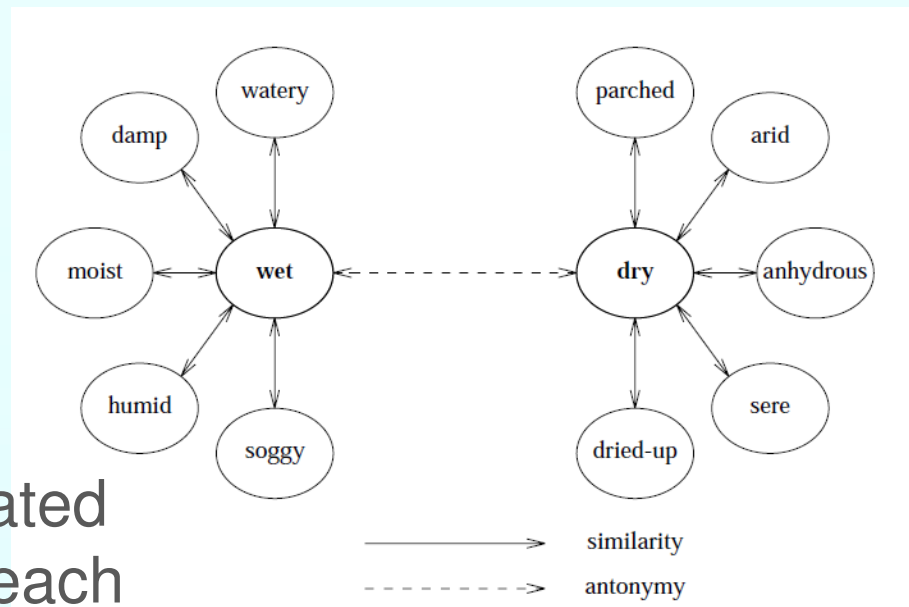


# SentiWordNet

**WordNet** : English nouns, verbs and adjectives are organised into synonym sets, each representing a concept.

**SentiWordNet** : add a value to the polarity. *For example, watery has a higher value than damp*

A SentiWordNet accumulated value will be assigned to each tweet to decide the polarity



# Key Phrases

## Examples of

positive key phrases : like it, interesting, =)

Negative key phrases : irritating, fallen behind, =(

Contextual sentiment: not not good => good

Negator : not, doesn't, won't etc

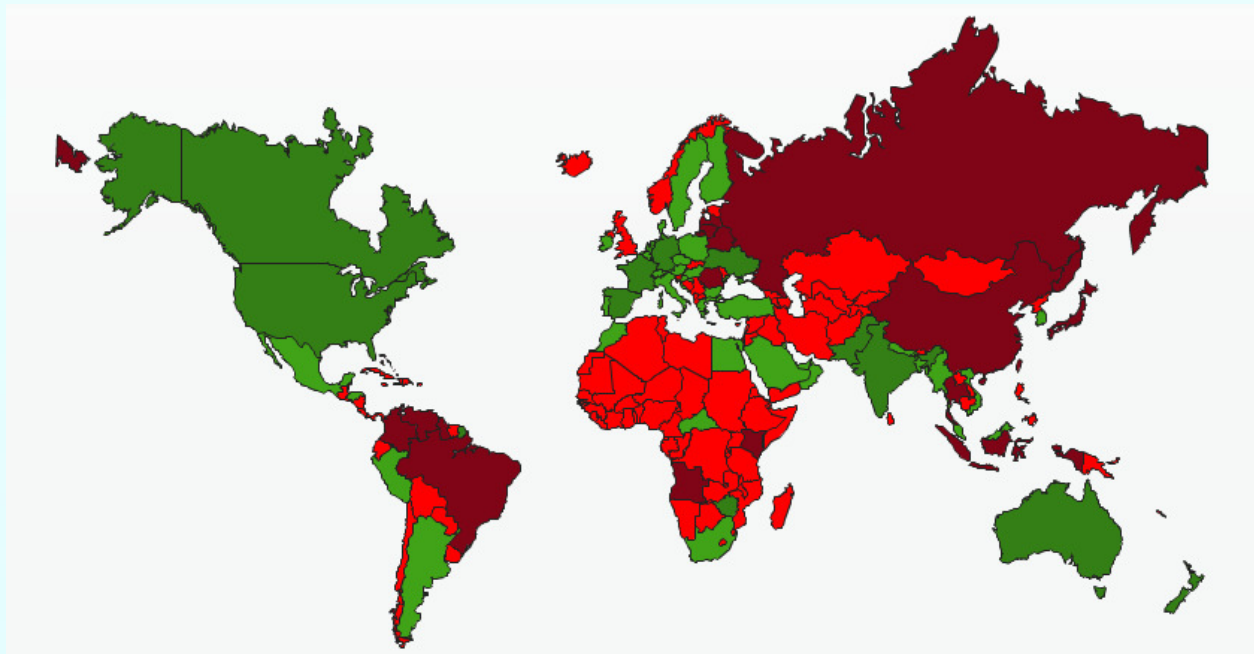
Content-word negator : eliminate, only

# Geographical Analysis

## Geographical Analysis

Using IBM ILOG Elixir visualisation components (heat map and vector map) on world map

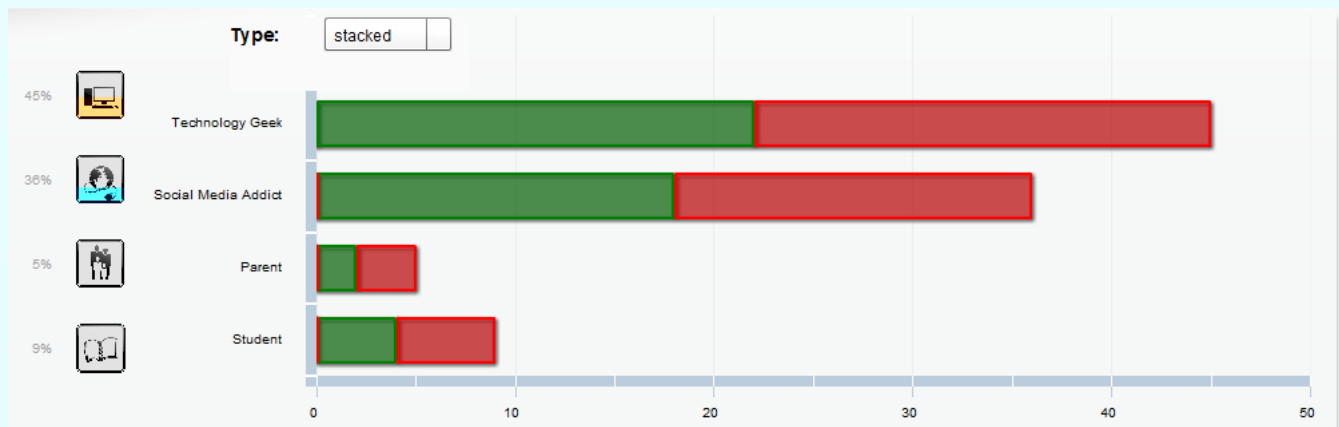
Combine with sentiment analysis



# User Category Analysis

Using biodata from twitter user account and key phrases matching to do user categorisation

Combine with influential and sentiment analysis





# Verizon iPhone Case Study

The Verizon iPhone4 launch in US was used in assessing the sentiment analysis and influential study algorithms (30 Nov 2010 to 22 Feb 2011 [12 weeks]).

**Keyword:** Verizon

**Event:** News of Verizon iPhone launch

**Event Timeline:**

**7 Jan 2011** : Verizon says that it will make an announcement on 11 Jan. Many speculated that it is related to iPhone launch.

**11 Jan 2011** : Verizon announces that iPhone 4 will be launched in early Feb.

**10 Feb 2011** : iPhone 4 available from Verizon.

# Influential Analysis

**MINING OPINIONS**  
FROM SOCIAL MEDIA

verizon

*Influential Analysis* *Sentiment Analysis* *Geographical Analysis* *Category Analysis*

**Influential Details**

**WSJ**

UserName: WSJ  
Name: Wall Street Journal  
Location: New York, NY  
BioData: Breaking news, features, and business coverage from WSJ.com.

Influential Index: 42  
Domain: #verizon

Top 10 Influentials		Influential Index
WSJ	Wall Street Journal	42
	Verizon	21
	Verizon Wireless	16
	TIPb	9
	pod2g	6
	iClarified	6
	AppAdvice.com	6
	jimmy fallon	5
	USA TODAY Top News	5

Date And Time	Content	Sentiment
2011-01-11 23:53:57	We're live blogging the 11 am ET Verizon announcement. What plans does it have for the iPhone? <a href="http://on.wsj.com/f1OaeF">http://on.wsj.com/f1OaeF</a>	positive

Copyright © 2010 by Nanyang Polytechnic

**Top 3 influencers:**

**Wall Street Journal followed by 2 Verizon official twitter accounts**

Verizon

*TechCrunch Review: The Name's iPhone. Verizon iPhone <http://tcn.ch/f6KLZa> (via @techcrunch)*

Verizon Wireless

*MacWorld? Verizon iPhone Review #6: The only difference that matters is the network itself.*

*<http://bit.ly/gt81ce> #VZWiphone*

# Sentiment Analysis

Week 6 & 7  
(4 Jan- 17 Jan)

Week 10 & 11  
(1 Feb- 14 Feb)

**MINING OPINIONS**  
FROM SOCIAL MEDIA

verizon

Influential Analysis  
Sentiment Analysis  
Geographical Analysis  
Category Analysis

**Week 11**  
(8 Feb 2011 to 14 Feb 2011)

Positive		Negative	
User Name	Content	User Name	Content
S_DUNC	RT @MVD59: Hoooolll! RT @hugejackass: Dude finally got the Verizon iPhone. It's wonderful. Team iPhone whatup? #getatme	ItsMeScottG	apparently the Verizon iPhone4 has the absolutely the same battery life as the AT&T iPhone4, it has no new (cont) <a href="http://tl.gd/8qbr3c">http://tl.gd/8qbr3c</a>
davidejorgensen	Verizon iPhone 4 has CDMA/GSM hybrid chipset. As expected, future models will likely work across carriers World Wide. <a href="http://is.gd/odwZpb">http://is.gd/odwZpb</a>	lamBrownLegacy	Follow me DM me your #s. I got the new iPhone4 for verizon and i can't get my old contacts back.
lanihzklassen	@shidsterlove Verizon is emailing coupons for free iPhone 4s!!! <a href="http://is.gd/Hg3JA5?=-mze4">http://is.gd/Hg3JA5?=-mze4</a>	PETTY279	RT @richi: Verizon iPhone 5 release date causes "poor" Apple sales? - bloggers count pinhead-balancing angels <a href="http://richij.com/itbw">http://richij.com/itbw</a>

Week	Positive	Negative
Week 7	~100	~100
Week 8	~100	~100
Week 9	~100	~100
Week 10	~600	~450
Week 11	~600	~450
Week 12	~100	~100

Content	User	Date Time
My wife with his new Verizon iPhone	JeremyCasts	2011-02-22 09:25:43
...ing a kid with a new puppy- so		
...rapport124 i got a verizon	TMNFVM	2011-02-22 06:55:36
... iPhone I will have :)	jenk14034	2011-02-22 05:55:49

Best 20 Records >>

Copyright © 2010 by Nanyang Polytechnic

# User Category Analysis

Bubble graph (

- centre of the bubble (above the separator, positive sentiment)
- volume of the bubble (percentage of users in the category. the bigger the bubble, the more the users)



Sentiment bar chart in stacked (absolute percentage of each category) or 100% (relative percentage in 100% for the category)



THANK YOU!

twitter 

Microsoft  
Exchange

facebook

[lo\\_siaw\\_ling@nyp.gov.sg](mailto:lo_siaw_ling@nyp.gov.sg)