# Social Media Analysis

Ashish Kumar Jha

# Agenda

**Introduction to course**

**Introduction to Social Media**

**Online opinion formation**

**Network structures**

# Reading material

- **Primary textbook**

– Miranda, S. M. (2019). Social Analytics: Network and Text Methods with NodeXL and R. Prospect Press.

– The below material is available online at no cost

- Silge, J., & Robinson, D. (2017). Text mining with R: A tidy approach. " O'Reilly Media, Inc.".
- Gephi Tutorials – Available at https://gephi.org/users/

# Other Recommended Books

- For deeper mathematical insights into graph theory and network structures in non media settings
  - Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2018). Analyzing social networks. Sage.
- For deeper insights into human cognition and brand management, buzz management and virality on social media from marketing perspective
  - Wendy W. Moe ; David A. Schweidel, Social Media Intelligence, (2014) Cambridge University Press
- A comprehensive codebook in Python for mining various other social media sites
  - Russell, M. A. (2013). Mining the social web: data mining Facebook, Twitter, LinkedIn, Google+, GitHub, and more. " O'Reilly Media, Inc.".
- A codebook for working with Twitter and applying various text analytics techniques
  - Danneman, N., & Heimann, R. (2014). Social media mining with R. Packt Publishing Ltd.

# Assessment

**Individual Assignment 40%**

- This would be a take home assignment that would be delivered individually. Students would be assigned 1 reading each. They would have to critique the reading and provide a short written report on the same.

**Discussion Participation 20%**

- Participation in online discussion board to enable peer and out-of class learning

**Group project 40%**

- A group project shall be deliverable by students working in groups of 3-4. The groups would be formed by the students. Group project would be delivered within 3 weeks from end of the course.

# Individual Assignment

## Submission

- Each student will have to upload on blackboard a critique of about 800 words of the assigned article. The critique should talk about the following
  - A brief (100 word) summary of the purpose of the article.
  - A brief summary of the data and the method utilized in the article as well as the possible use cases of this method.
  - What are the real life implications of the findings discussed in the article (beyond the implications discussed in the article itself)
  - Do you see any challenges in implementation of those techniques presented in the article?

**Deadline: Submission for each Individual assignment would be 11 Pm the day before the session**

# Readings for Critique

| Session | Article Title | No. of Students |
|---|---|---|
| 2 | Increasing the ROI of Social Media Marketing | 6 |
| 2 | Competent Jerks, Lovable fools and formation of social networks | 6 |
| 3 | Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach | 6 |
| 3 | Harshness-aware sentiment mining framework for product review | 6 |
| 4 | Coupling Topic Modelling in Opinion Mining for Social Media Analysis | 6 |
| 4 | Tweeting on COVID-19 pandemic in South Africa: LDA-based topic modelling approach | 6 |
| 5 | A Social Networks Approach to Public Relations on Twitter: Social Mediators and Mediated Public Relations | 7 |
| 5 | Communities of shared interests and cognitive bridges: The case of the anti-vaccination movement on Twitter | 7 |
| 6 | SENSEMAKING IN SOCIAL MEDIA CRISIS COMMUNICATION – A CASE STUDY ON THE BRUSSELS BOMBINGS IN 2016 | 7 |
| 6 | An Analysis of Online Classes Tweets Using Gephi: Inputs for Online Learning | 7 |

# Group Project

**Teams of 6 members each by January 30, 2023. Team details would be posted on blackboard**

**Project Submission Deadline: -** 11: 59 Pm on Friday 17th March 2023

**Submission**

- In spirit of the course i.e. social media analysis, the submissions would be in video format. Teams would have to upload 15 min video (As 3 x 5 min) explaining their problem statement, data, analysis, results, implications for business. Students are free to be creative in development, design and presentation of the video. High level of creativity would be suitably appreciated and marked. There will not be any report or ppt submission.
- All the data and analysis files, code files etc. would need to be separately uploaded as a zip folder on the Blackboard page.

**There will be a component of peer evaluation as well for the group project where each team will evaluate other team's projects**

# Discussion

- Discussion on a discussion thread on the blackboard page for the module.
- Students would be encouraged to share their thoughts, ideas, comments related to topics being discussed in class on the forum.
- Students could also share related news, articles, insights from scientific or popular press outlets. Only sharing external links would not be entertained.
- When sharing external links, it should contain the students' commentary or inputs on its relevance.
- Students are also encouraged to engage with each other's posts and discuss their implications in relation to the concepts being discussed in class.

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Trinity Business School

# Any Questions ??

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Trinity Business School

# Session 1: Introduction to Social Media Analysis

# Social Media analysis

- **What is social media**

- **Why is it necessary**

- **What are implications of social media**
  - In society
  - In business

- **How can you use it**

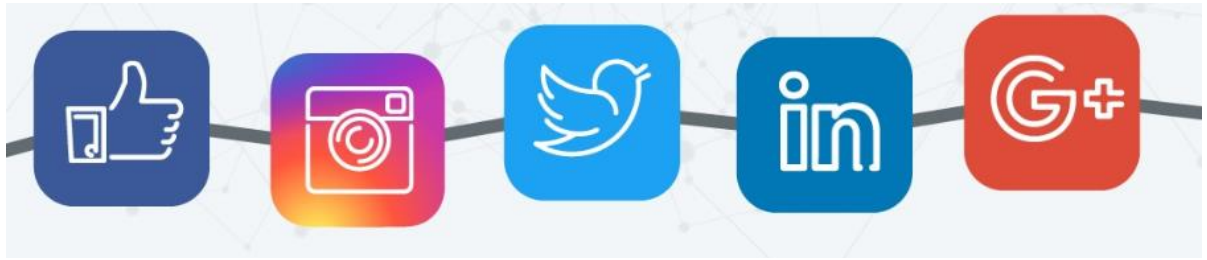
# Social Media analysis

- What is social media
- Social media as social network
- Understanding Networks
- Analytics on Social Media texts

# Social Media

- **Social media refers to online tools and services which allow an exchange of ideas, information, videos, pictures, and graphics — just about anything you can name.**

- **Examples**



- **Social Media analysis is a special case of social network analysis**
  - Social media is a special case of social networks

# Social Media

- Bookmarking Sites and Social News Sites (Digg)
- Blogs and Microblogs (Twitter, Tumblr)
- Social Networking Sites (Facebook, Google+)
- Shopping Sites (Amazon)
- Multimedia Sharing (YouTube, Flickr)
- Virtual Worlds (World of Warcraft, Second Life)

# Social Media

- **Positives and negatives**
  - Catastrophe
  - Natural Disasters
  - Ease of communication
  - Addiction
  - Social unrest
  - False news

# Social Media Marketing

- Tools and strategies that were cutting-edge just a few years ago are fast becoming obsolete, and new approaches are appearing

- Marketers must manage new challenges such as the inclusion of other departments or employees in marketing actions

- Competitive nature of the firm's primary industry moderates the effect of social media marketing and that in the hotel industry, firm reputation impacts the effectiveness of social media efforts

# Social Media Marketing

- **Facebook, MySpace, and Twitter attract more than 90% of young adults and teens and represent over a quarter of all Internet traffic**

- **Pros and Cons of using social media as a marketing tool**
  - Dave Carroll's United Breaks Guitars video case
- **Four primary customer engagement strategies for social platforms:**

– relationship management

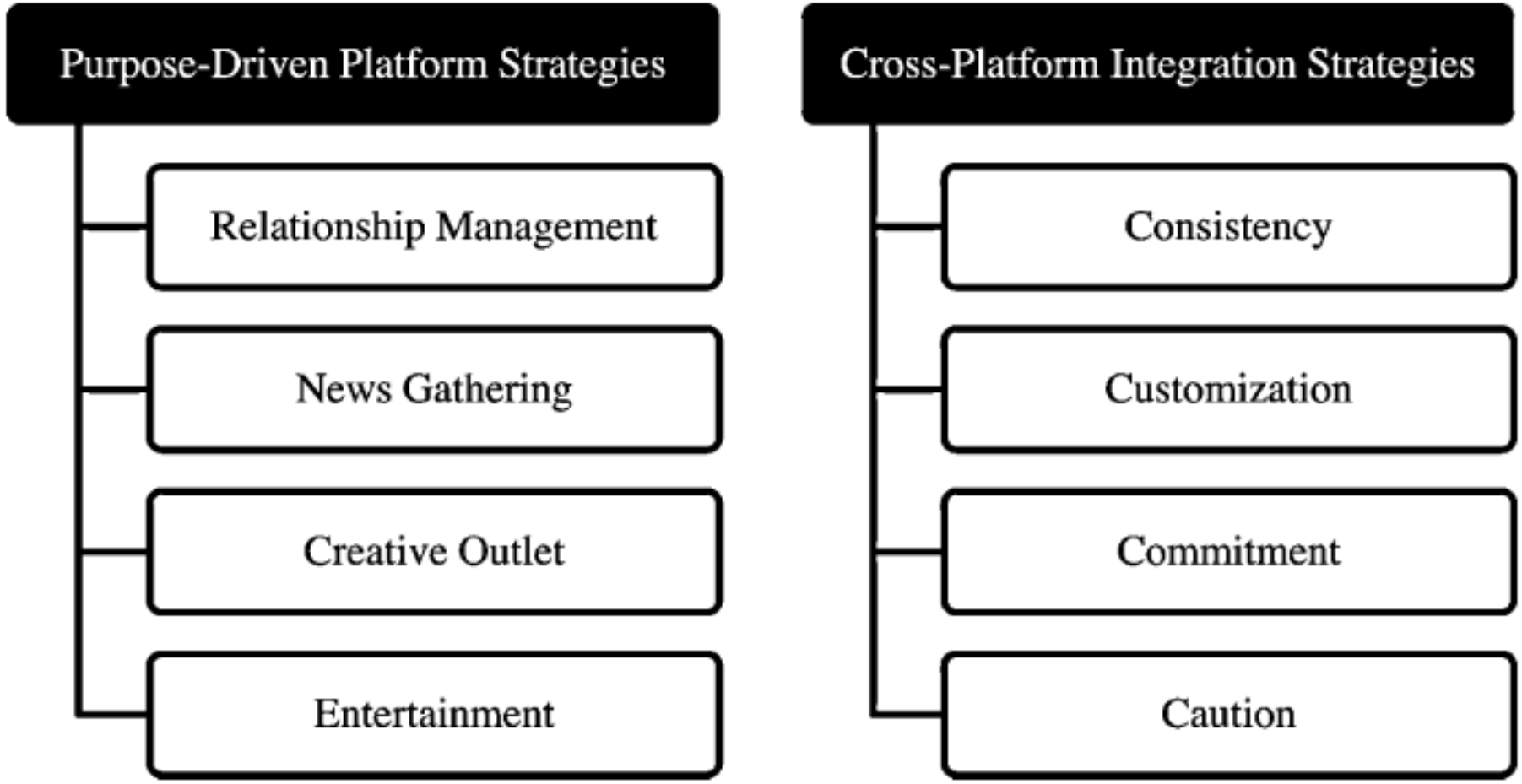
– news gathering

– Creativity

– entertainment

# Social Media Marketing

- **Word of mouth vs. social media**
  - Public forums vs. private conversations
  - On the record
- **What is required**
  - Understand what is important to measure
  - Understand how people behave on social media
  - Impact of social media user behavior for organizations
  - Identify right metrics and networks


Purpose-Driven Platform Strategies
Relationship Management
News Gathering
Creative Outlet
Entertainment
Cross-Platform Integration Strategies
Consistency
Customization
Commitment
Caution

# Social Media Monitoring

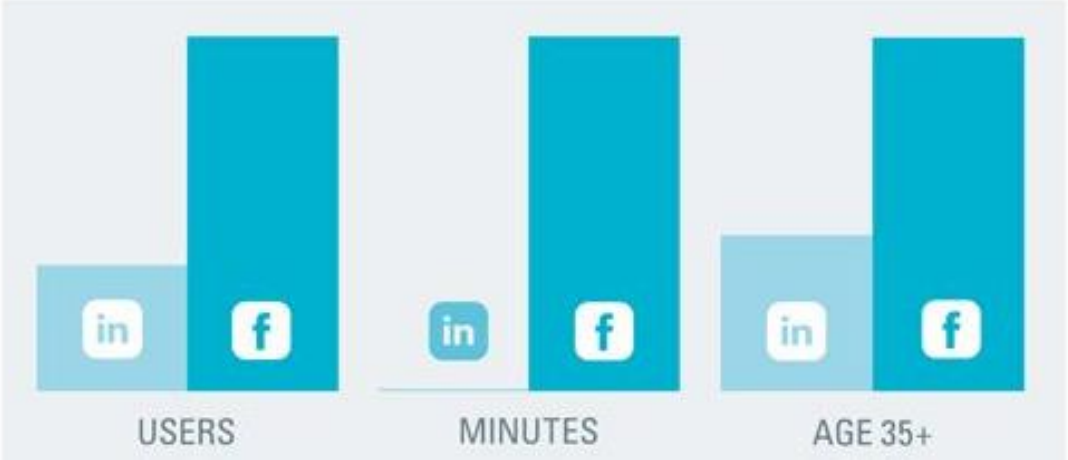
The success of different post types based on average reach and engagement.

Show All Posts ▾ Reach Post Clicks Reactions, Comments & Shares

| Type | Average Reach | Average Engagement |
|---|---|---|
| Video | 102,438 | 5,490<br>1,241 |
| Photo | 72,243 | 5,662<br>1,861 |
| Link | 19,148 | 1,076<br>207 |
| Status | 10,623 | 887<br>262 |

## Differences In Usage

People do not shut off business thinking when on Facebook. The best reason to consider Facebook over LinkedIn for B2B marketing is that business people spend more time on it. Business decision makers are typically older and there are almost twice as many on Facebook...



# Social media monitoring tools- Corporate


# Social media monitoring tools- Corporate

- Brandwatch
- Sproutsocial
- ZohoSocial
- SparkToro
- Buffer Analyze
- Google Analytics

# Challenges with basic monitoring

- **Posting comments is a voluntary act and therefore posted opinions may not necessarily represent the opinions of the majority**
  - extreme opinions are more likely to be expressed than more moderate ones.
- **Extreme opinions are oversampled**
  - Comments that are either extremely positive or extremely negative tend to be selected and scrutinized
- **Presence of Bias**
  - For example, a brand manager who has a hunch that a product is overpriced may tend to focus on the comments that talk about price or interpret a comment
- **Not scalable**

# Analytic metrics and techniques

- **Text mining**
  - Word count, Word cloud
  - Sentiment mining
  - Co-occurence of words
  - Topic extraction
- **Volume metrics**
  - # of comments # no. of likes etc.
- **Valence metrics**
  - Average sentiment, ratio of positive or negative comments etc.

# Analytic metrics and techniques

Table 1.1. Taxonomy for the use of volume and valence metrics

| | | Volume | |
|---|---|---|---|
| | | Low | High |
| Valence | Negative | Nuisance | Red Flag |
| | Positive | Potential Success | Viral Success |

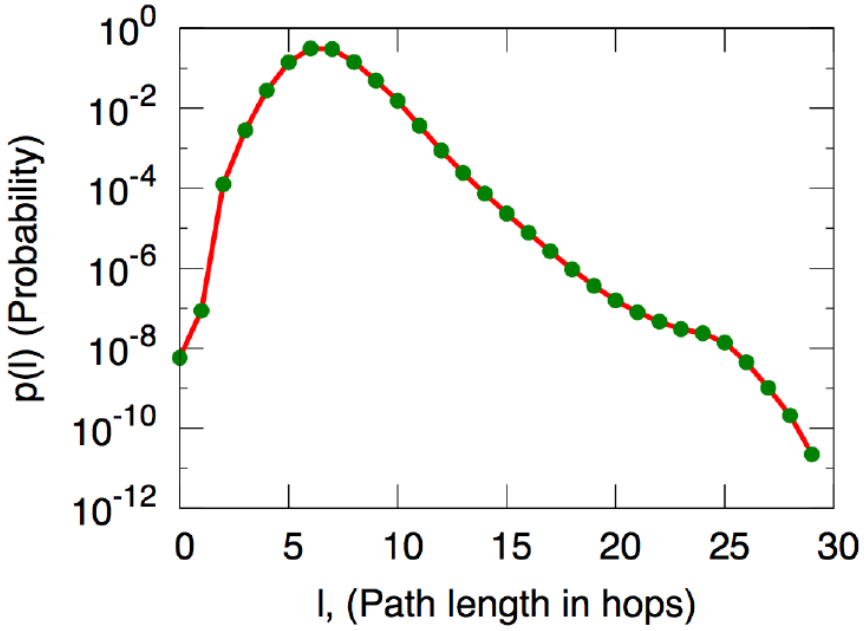
Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Trinity Business School

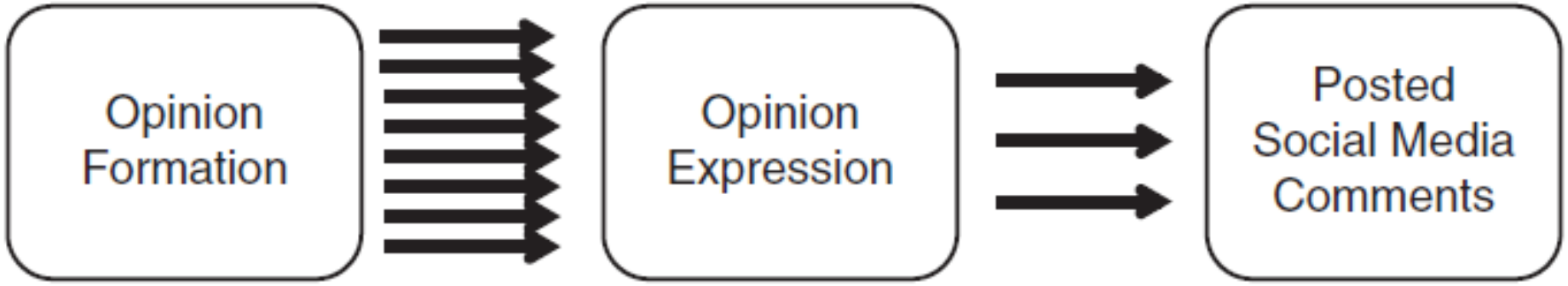
# Theories for social media

# Small World Phenomenon

- Social networks tend to have very short paths between essentially arbitrary pairs of people
- The world looks “small" when you think of how short a path of friends it takes to get from you to almost anyone else
- Erdos Number

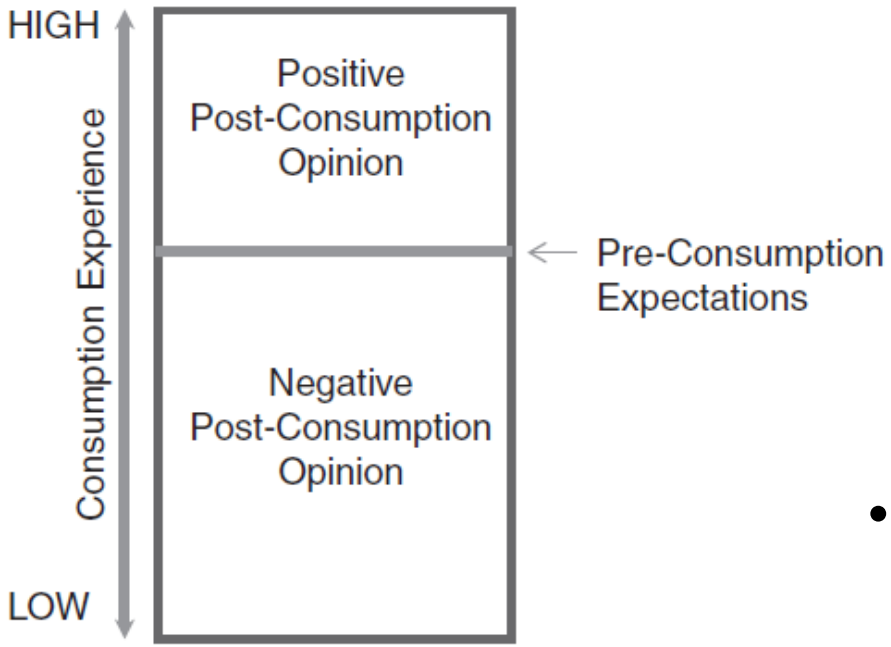


# Opinion formation




- Opinion formation vs expression
  - posted social media comment is an outcome of both the opinion formation and opinion expression stages
  - our opinions are also influenced by a host of other factors such as advertising , preconceived expectations , and social influence
- Utility functions
  - the value to the user of the product, service, or experience into its component parts

# Opinion formation factors




- Expertise influence online opinions
  - Those who lack experience or knowledge will tend to voice opinions that are based on public perceptions
  - more experienced and knowledgeable consumers will focus more on the actual performance of the product.
- Expectation vs. experience

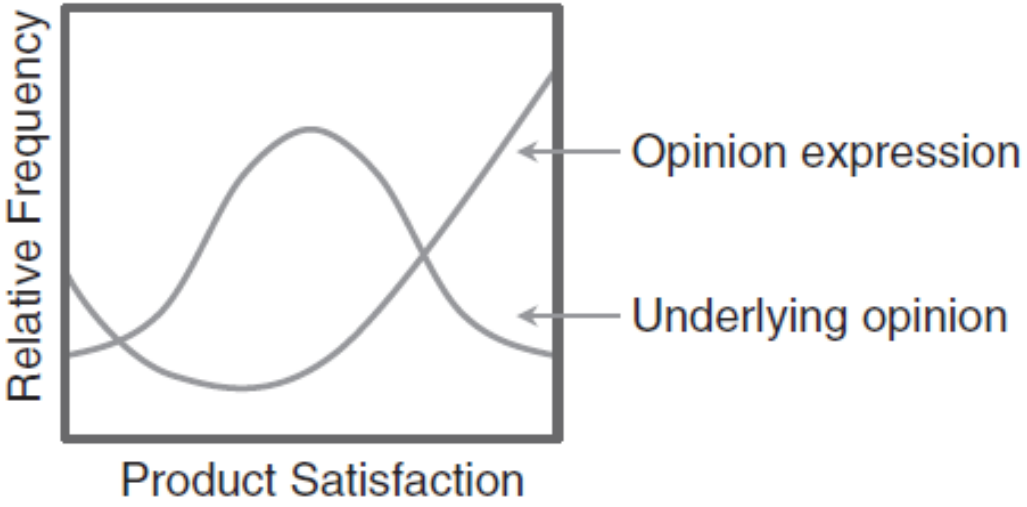
  A consumer's post-consumption utility is a result of reconciling the pre-consumption expectations with the actual experience

# Population Effect: Information Cascade

- **Restaurant A and Restaurant B with information asymmetry**
  - You plan information A with private information
  - Restaurant B has crowd. Where do you go?
  - Herding behavior
- **Information Cascade**
  - They can be wrong
  - Based on very little information
  - They are fragile

# Why do we share opinion

- **Posters vs lurkers**
  - Lurkers see social media as a source of information
  - Posters, on the other hand, turn to social media to express themselves
  - Understand *Opinion Leaders*
- ***Reasons for posting***
  - *Product-involvement*
  - *self-enhancement*
  - *altruistic* motivation
- **Extremity Bias**



# Implications for Social media Intelligence

- *Online opinions reflect the vocal minority*
  - voices represented on social media are not necessarily representative of the greater population
- *Opinion variety is exaggerated*
  - If there is no added value to contributing your opinion, why should you spend the time and effort to write and post it?
- *Opinion variety attracts experts and negative opinions*
  - as online opinions evolve and opinion variety inevitably increases
- *Opinions are more extreme*
- *Moderates are underrepresented and their numbers are decreasing*

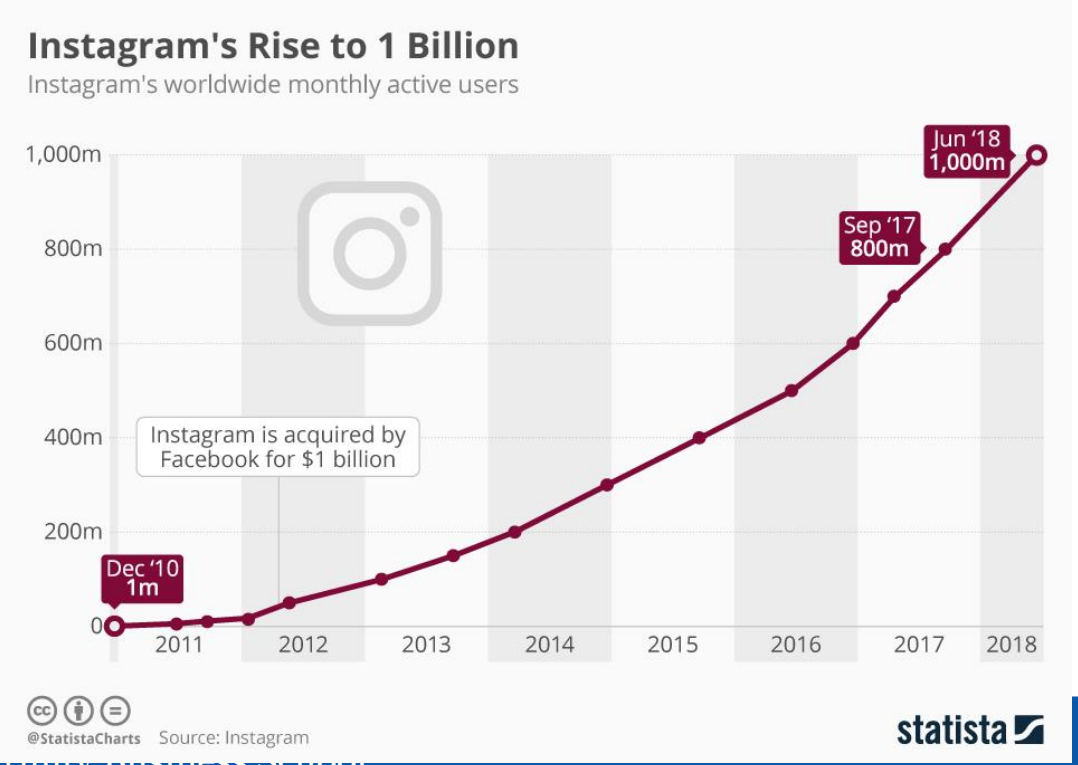
Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Trinity Business School

# The science behind the networks on social media

# Network Economy

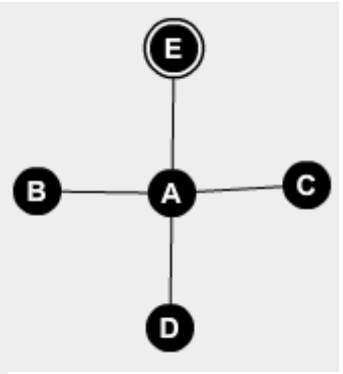
- Networks → Networked Economy ????
- Network Effect

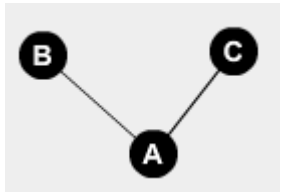


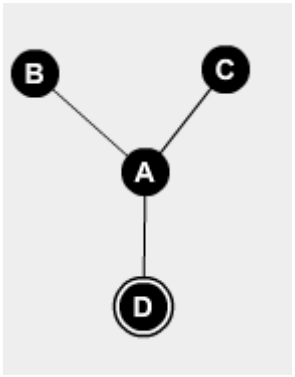

# Network Analytics

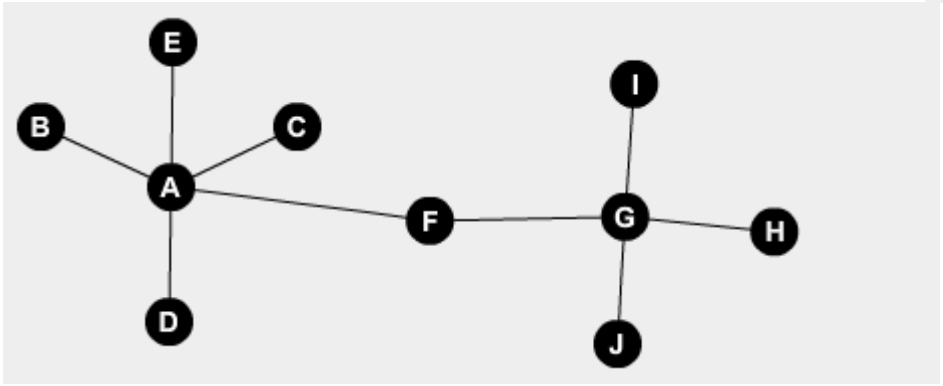
- Strength of weak ties

- Broader field of analysis → Deeper insights



# Social Network Analysis

- A social network is a social structure made up of individuals (or organizations) called "nodes", which are tied (connected) by one or more specific types of interdependency, such as friendship, kinship, common interest, financial exchange, dislike, sexual relationships, or relationships of beliefs, knowledge or prestige.

- Social network analysis views social relationships in terms of network theory consisting of nodes and ties (also called edges, links, or connections). Nodes are the individual actors within the networks, and ties are the relationships between the actors. The resulting graph-based structures are often very complex.

# SNA software

- **Popular Academic software**
  - UCINet
  - Pajek
  - Gephi
  - Statnet
- **Non-Academic Software**
  - iGraph
  - Tulip
  - Rapidnet

# Social Network Terms

- *Whole networks* (also known as *complete networks*), all of the ties containing specified relations in a defined population
- *Personal networks* (also known as *egocentric networks*), the ties that specified people have, such as their "personal communities"
- *Snowball network* refers to the idea that the alters identified in an egocentric survey then become egos themselves and are able in turn to nominate additional alters
- *Hybrid networks* refers to idea in which egos in complete networks can nominate alters otherwise not listed who are then available for all subsequent egos to see

# Social Network

- **Levels of Analysis**

– Dyads- Pairwise relationship O(n^2)

- Fundamental unit of network data collection
- Has highest frequency
- Kind of question- Are employees with offices next to each other likely to develop stronger friendships??

– Node – smallest unit O(n)

- Smallest unit irrespective of network
- Node level network properties are aggregation of dyad level computations
- Are employees more central in organization more likely to leave??

# Social Network

- **Levels of Analysis**
  - Group level- highest level of analysis O(n^0)
  - Analyzes properties of all collective nodes as a network
  - Are more connected firms likely to be more innovative??
- **There are some other levels of analysis that are seldom used**
  - Triads
  - subgroups
- **At each level of analysis, the node could be individual or collectivity**
  - Board members vs. boards
  - Employees vs. firms

# Networks and Relations

- **How an individual becomes acquaintance ?**
  - Milgram experiment- 6 degrees of freedom
- **Relation due to state**
  - Similarity
    - Location – city, country
    - Participation – club/ school
    - Attribute- gender, age
  - Relation
    - Kinship – family
    - Other- Friend/ Boss

# Networks and Relations

- **Relation due to state**

– Cognition

  - Affective- Liking/ hating
  - Perceptual- Knows

- **Relation due to event**

– Interaction- Sold to/ helped/ fought

– Flows- information flow/ product flow/ money flow

# Network Mode

- **Mode refers to distinct set of entities belonging to similar data collection structure**
  - i.e. mode is the collection of distinct nodes in a network
  - A single mode network will have all nodes from a single background e.g. all nodes are employees of a corporation
  - A two mode (or more) network will have entities from different background e.g. one set of nodes belong to corporation employees and other to government agencies

# Network Studies

| Level | Network as independent variable | Network as dependent/ outcome variable |
|---|---|---|
| Dyad Level | Friendship between farmers to predict which farmers sow same kind of seeds | Similarity of interests to see who becomes friends |
| Node level | Centrality in organization to see who is chosen for promotion | Use extraversion as a variable to predict who becomes a central figure in friendship network |
| Group Level | Connectedness of a network of individua's to predict which group solves problems faster | Type of organizational culture to predict structure of trust network |

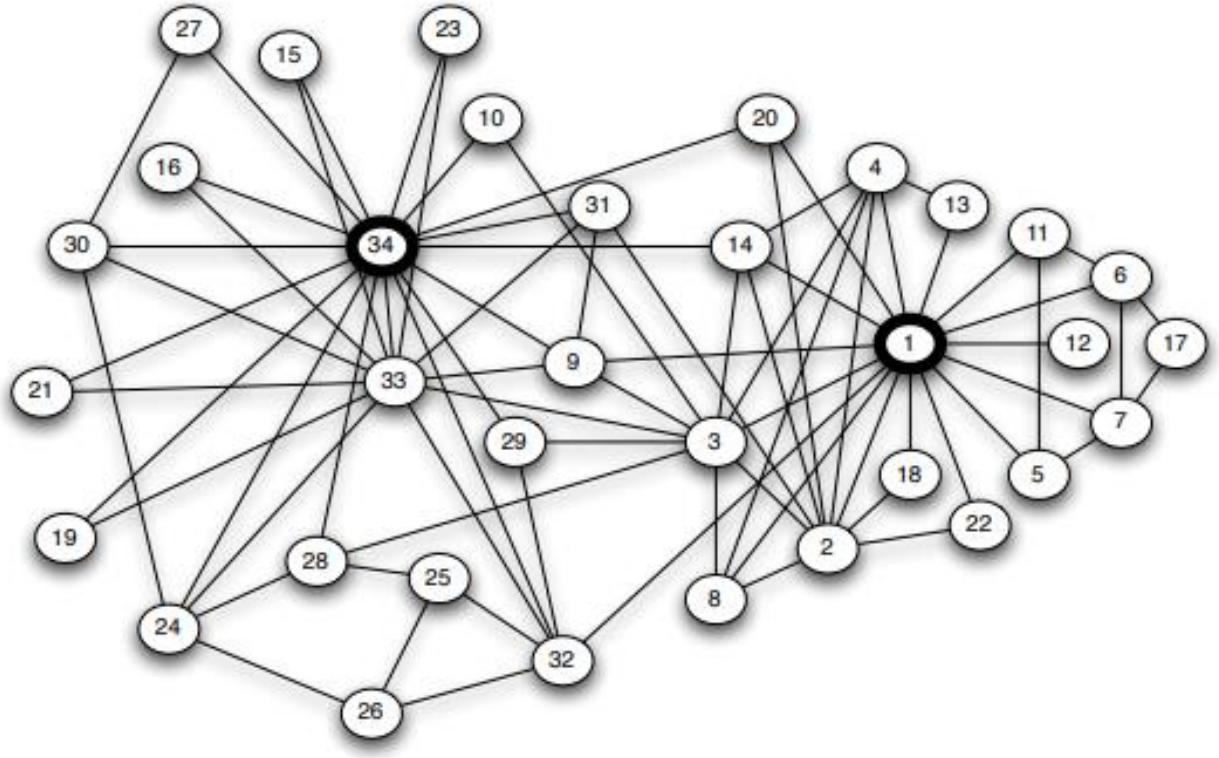



Figure 1.1: The social network of friendships within a 34-person karate club [421].

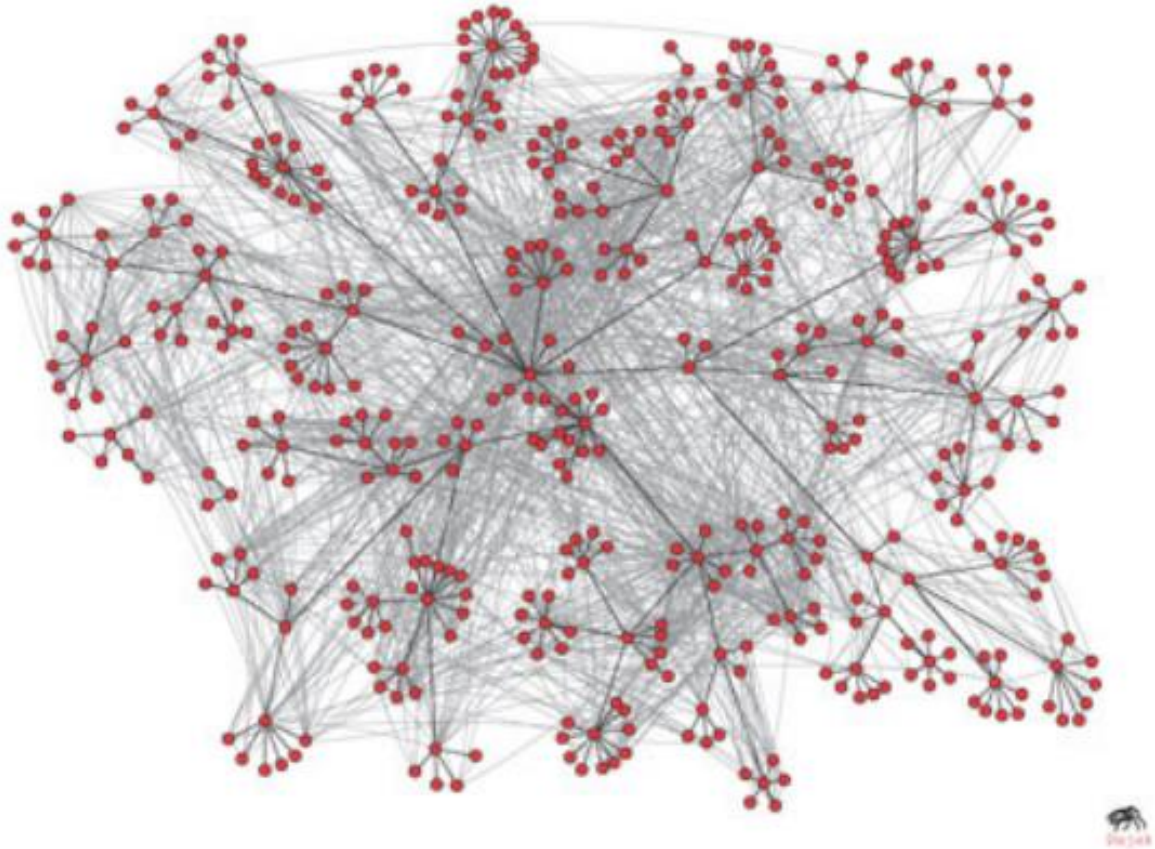


Figure 1.2: Social networks based on communication and interaction can also be constructed from the traces left by on-line data. In this case, the pattern of e-mail communication among 436 employees of Hewlett Packard Research Lab is superimposed on the official organizational hierarchy [6]. (Image from http://www-personal.umich.edu/ ladamic/img/hplabsemailhierarchy.jpg)

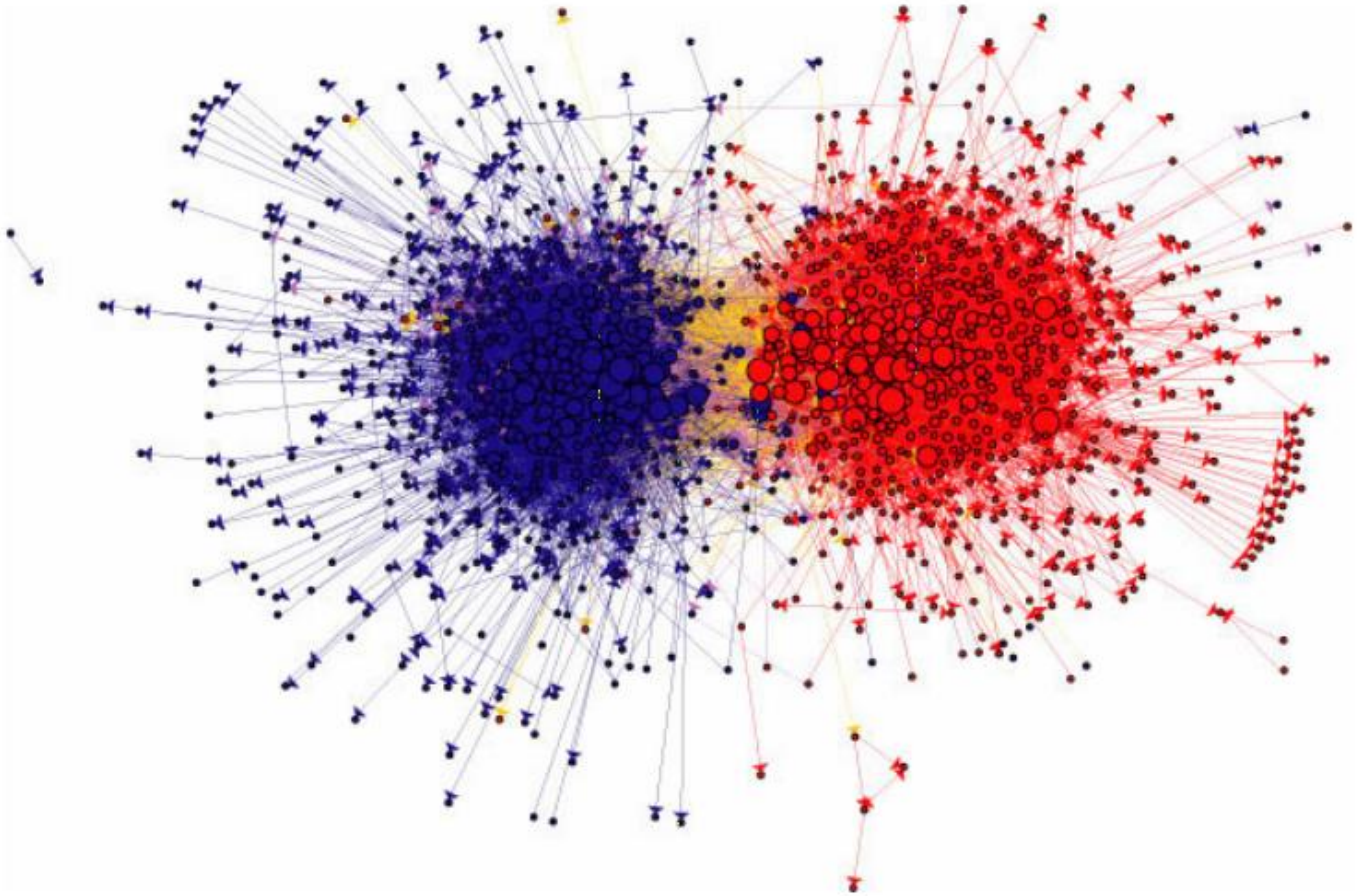


Figure 1.4: The links among Web pages can reveal densely-knit communities and prominent sites. In this case, the network structure of political blogs prior to the 2004 U.S. Presidential election reveals two natural and well-separated clusters [5]. (Image from http://www-personal.umich.edu/ ladamic/img/politicalblogs.jpg)

# Fundamental Economic concepts

- **Game Theory**
  - a group of people must simultaneously choose how to act, knowing that the outcome will depend on the joint decisions made by all of them
  - Prisoner's Dilemma
  - Nash Equilibrium
- **Braess's Paradox- adding resources to a transportation network can in fact create incentives that seriously undermine its efficiency, in a phenomenon known as Braess's Paradox**
- **Auctions and payoff**

# Network dynamics: Population

- **large population over time, shows a recurring pattern**
- **people influence each other's behaviour**

– Conformity bias - individuals with no a priori desire to conform to what others are doing — will nonetheless copy the behavior of others

– Groupthink

– Bandwagon effect

– Backfire effect

– Congruence bias

# Network dynamics: Structural

- **Network structure provides important further insights into how such kinds of influence take place**
- **Cascading effect**

– you care more about aligning your own behavior with the behavior of your immediate neighbours in the social network, rather than with the population as a whole