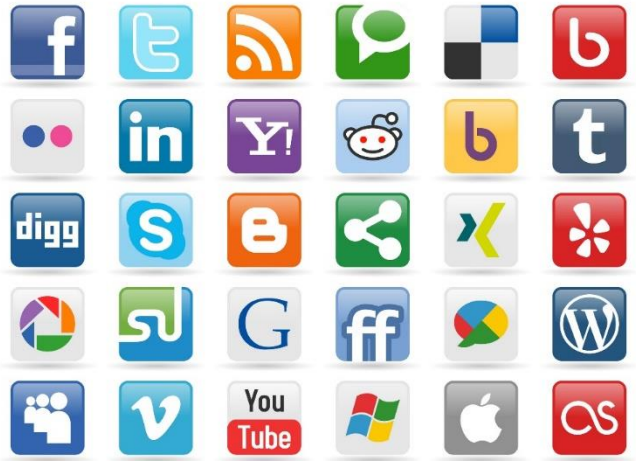


Introduction

Social Media & Network Analytics

Acknowledgement

- Lecture slides based on the ones available from:

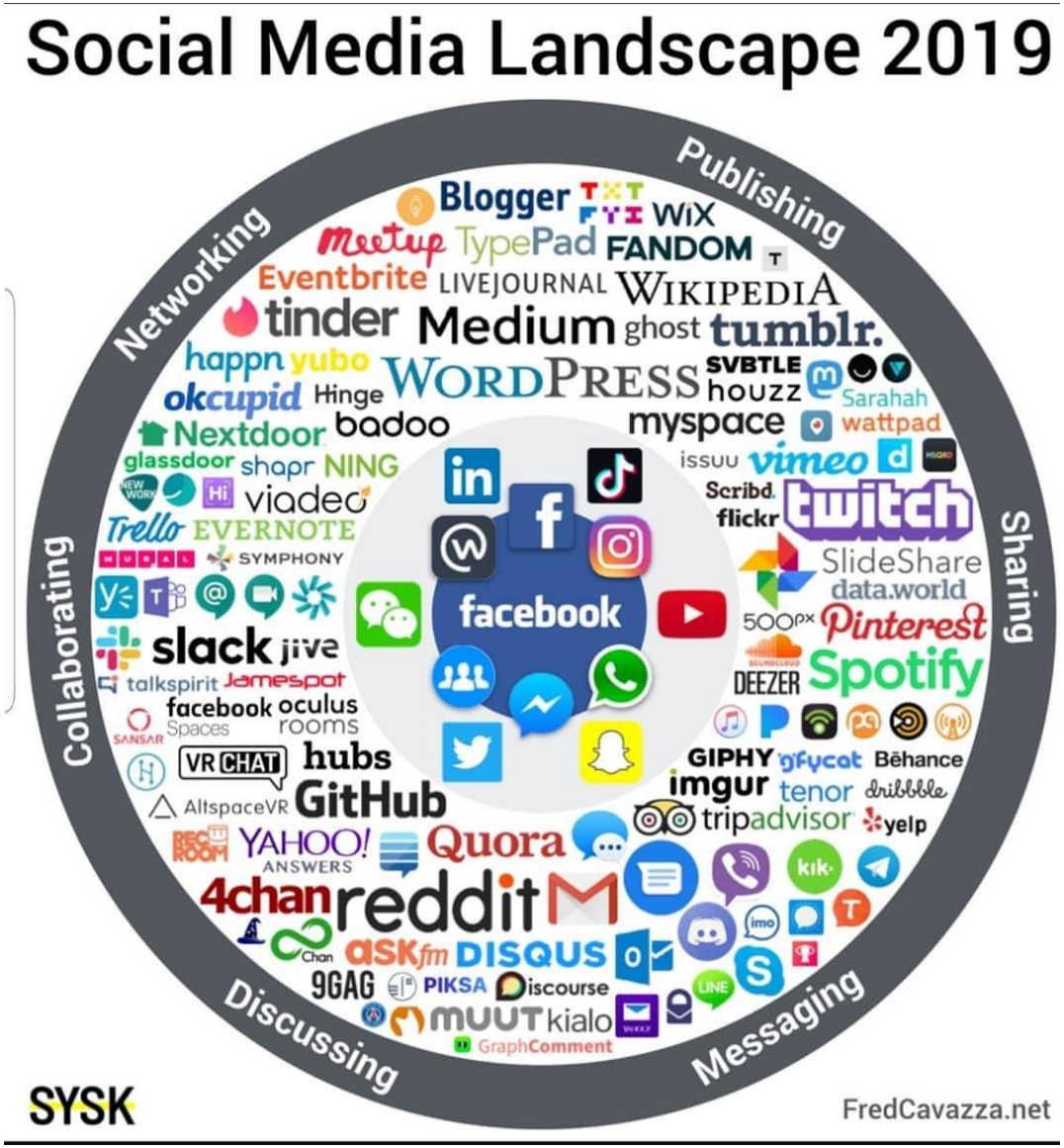
R. Zafarani, M. A. Abbasi, and H. Liu, *Social Media Mining: An Introduction*, Cambridge University Press, 2014.

Free book and slides at
http://socialmediamining.info/

Before we do anything further…

- Please introduce yourself to your neighbour
- What course are you studying?
- What kind of social media and networks do you use?
- Then I want you to use Menti to enter what social network use
  - Type www.menti.com then enter the code on the screen

Social Media & Networks

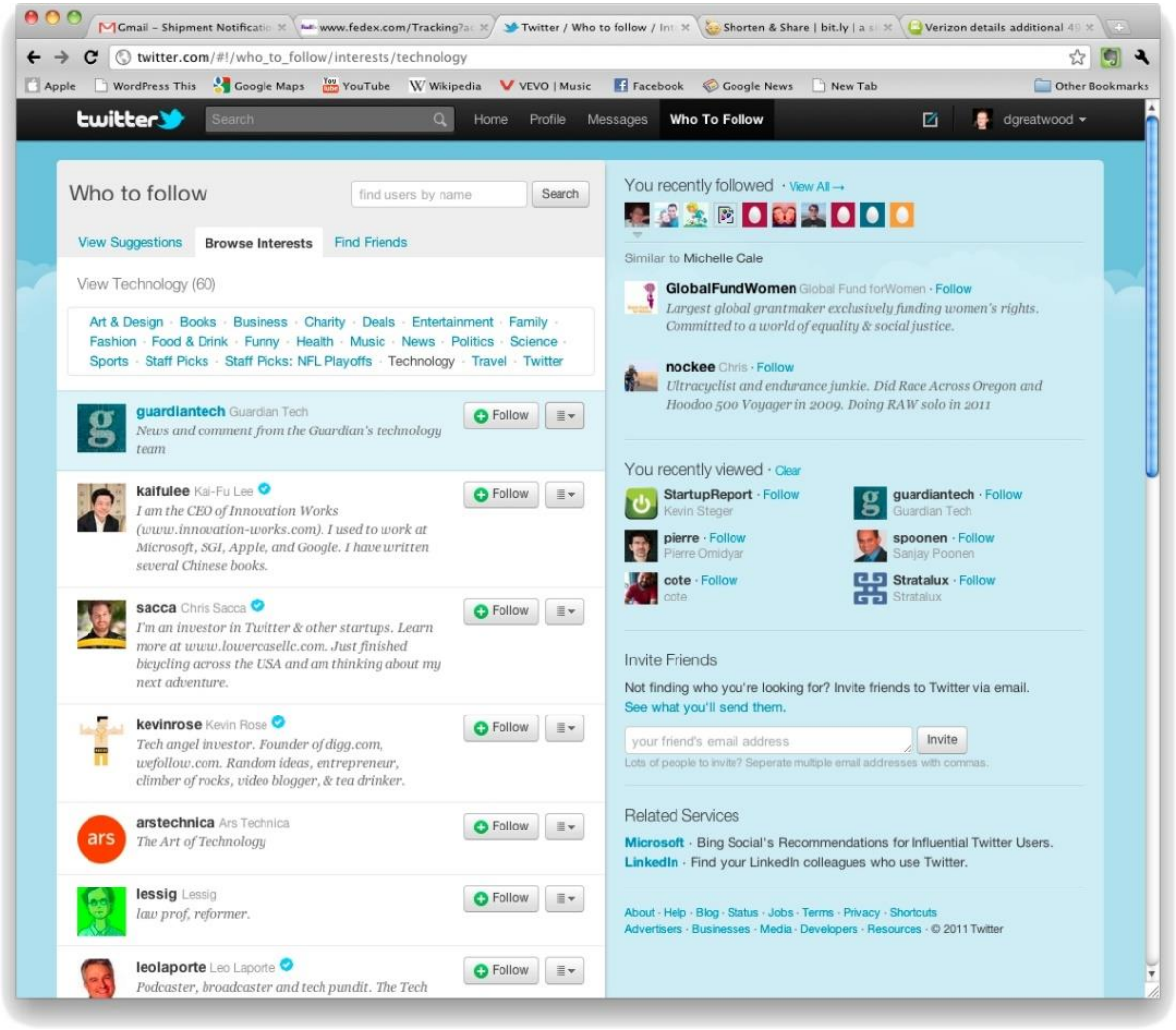


Facebook

- **What data does Facebook collect?**
- **How does Facebook use your data?**

What about Amazon?

Or Twitter?


twitter.com/#!/who_to_follow/interests/technology
twitter
Search
Home
Profile
Messages
Who To Follow
dgreatwood
Who to follow
find users by name
Search
View Suggestions
Browse Interests
Find Friends
View Technology (60)
Art & Design · Books · Business · Charity · Deals · Entertainment · Family · Fashion · Food & Drink · Funny · Health · Music · News · Politics · Science · Sports · Staff Picks · Staff Picks: NFL Playoffs · Technology · Travel · Twitter
guardiantech Guardian Tech
News and comment from the Guardian's technology team
Follow
kaifulee Kai-Fu Lee
I am the CEO of Innovation Works (www.innovation-works.com). I used to work at Microsoft, SGI, Apple, and Google. I have written several Chinese books.
sacca Chris Sacca
I'm an investor in Twitter & other startups. Learn more at www.lowercasellc.com. Just finished bicycling across the USA and am thinking about my next adventure.
kevinrose Kevin Rose
Tech angel investor. Founder of digg.com, wefollow.com. Random ideas, entrepreneur, climber of rocks, video blogger, & tea drinker.
arstechnica Ars Technica
The Art of Technology
lessig Lessig
law prof, reformer.
leolaporte Leo Laporte
Podcaster, broadcaster and tech pundit. The Tech
You recently followed · View All →
Similar to Michelle Cale
GlobalFundWomen Global Fund forWomen · Follow
Largest global grantmaker exclusively funding women's rights. Committed to a world of equality & social justice.
nockee Chris · Follow
Ultracyclist and endurance junkie. Did Race Across Oregon and Hoodoo 500 Voyager in 2009. Doing RAW solo in 2011
You recently viewed · Clear
StartupReport · Follow
Kevin Steger
guardiantech · Follow
Guardian Tech
pierre · Follow
Pierre Omidyar
spoonen · Follow
Sanjay Poonen
cote · Follow
cote
Stratalux · Follow
Stratalux
Invite Friends
Not finding who you're looking for? Invite friends to Twitter via email.
See what you'll send them.
your friend's email address
Invite
Lots of people to invite? Seperate multiple email addresses with commas.
Related Services
Microsoft · Bing Social's Recommendations for Influential Twitter Users.
LinkedIn · Find your LinkedIn colleagues who use Twitter.
About · Help · Blog · Status · Jobs · Terms · Privacy · Shortcuts
Advertisers · Businesses · Media · Developers · Resources · © 2011 Twitter

Course Information

Social Media and Network Analytics

- Course Id:
  - COSC2671
- Masters of Data Science/Analytics/Operations Research
- Classroom and Hours (checked myTimetable for latest):
  - Lectures: 10.08.24, Wednesday 18:30 – 20:30
  - Tutorials: 14.6.*, Thursday 17:30 and 18:30
  - Labs: 14.10.31 and 14.11.37, Tuesday 15:30 and Wednesday 20:30
- Canvas:
  - Everything (slides/homeworks/projects/discussion board etc.)

Teaching Team Introduction

- **Course Coordinator and Lecturers:**
  - Jeffrey Chan (jeffrey.chan@rmit.edu.au)
  - Saiedur Rahaman (last 2 weeks)

- **Tutor and Lab Demonstrators:**
  - Jeffrey Chan
  - Saiedur Rahaman

- **My office Hours (Starting July 30th (week 2)):**
  - Tuesday 4:30 – 5:30pm (after one of the labs), in 14.08.15

Objectives of Our Course

- Use social media and network to understand social aspects of the Web
  - Social Theories + Social media/Network + Analytics
  - Learn to collect, clean, and represent social media/network data
  - Focus on text and network analysis
- Study or ask interesting questions from their analysis
- Learn representative algorithms and tools

Course Information

- Official Prerequisites:
  - COSC 2670 – Practical Data Science or
  - COSC 2531 – Programming Fundamentals
- Expected Knowledge
  - Introductory level in Data Science
  - Machine Learning
  - Able to program and do analysis in Python
- Desirable
  - Data Structures and Algorithms (Algorithms & Analysis)
  - Statistics

Syllabus

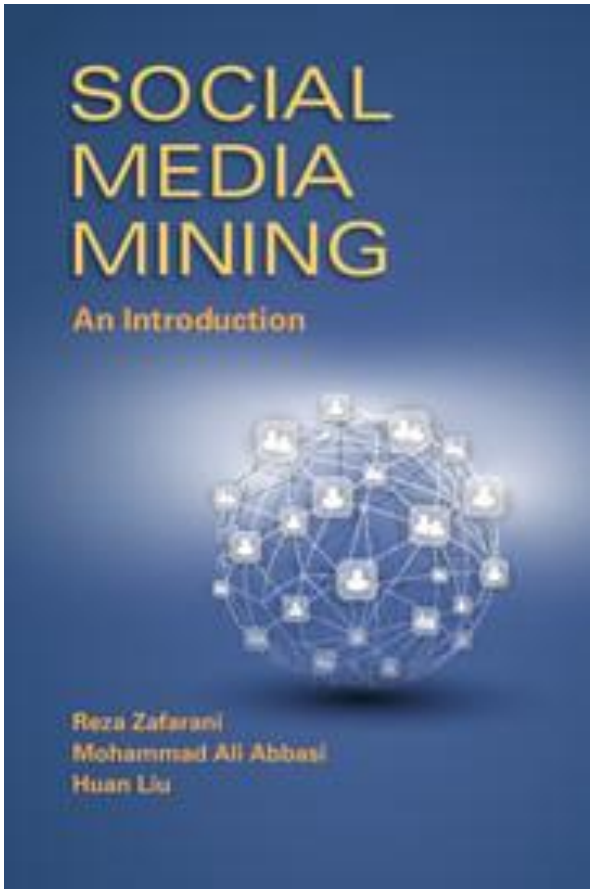
Week	Topic
1	Introduction
2	Introduction to Text Analysis & Machine Learning
3	Sentiment Analysis
4	Sentiment Analysis & Topic Modelling
5	Topic Modelling
6	Event Detection
7	Graph Introduction
8	Social Network Analysis
9	Community Detection
10	Information Diffusion
11	Influence & Homophily
12	Presentations

Text and Reference Books

Primary Reference:

Social Media Mining, Reza Zafarani, Mohammad-Ali Abbasi, Huan Liu, Cambridge University Press 2014,

- Available at http://socialmediamining.info or amazon.com



Other references

- Web Mining
  - "Web Data Mining" by Bing Liu, 2nd Ed 2011 (https://www.cs.uic.edu/~liub/WebMiningBook.html)
- Sentiment Analysis
  - "Sentiment Analysis and Opinion Mining" by Bing Liu, 2012 (https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf)
- Network Analysis
  - *"Networks, Crowds and Markets: Reasoning About a Highly Connected World*" by *David Easley* and *Jon Kleinberg*. (http://www.cs.cornell.edu/home/kleinber/networks-book/)
  - "*Introduction to Social Network Analysis*" by Robert Hanneman and Mark Riddle. (http://faculty.ucr.edu/~hanneman/)
  - *Social Network Analysis: Methods and Applications*" by Stanley Wasserman and Katherine Faust

Course Workload and Evaluation

- Assessment:
  - Projects/Assignments (35%)
    - Two of them, last one has presentation
  - Weekly quizzes (15%) – 11 quizzes, one per week, up to week 12. Count best 10 quizzes, each counted quiz contribute 1.5% of total marks.
    - Quizzes closes on Tuesday 11:59pm, please do it before then
    - No quizzes for this week
  - Exam (50%)
  - Late penalty for assignments:
    - Standard CS, -10% of maximum total each day
  - Special considerations please have documentation

Expectations

- Me -> You
  - This is a elective post-graduate course. It is (hopefully) really interesting and useful for your careers, but you also **need to put in**
  - Try to **attend** classes (but I understand there may be work and other constraints)
  - Interact while in class! Be curious, ask questions.
    - In this class, we have the **"no stupid question"** rule!
  - I don't have a loud voice, so please use common sense and courtesy if need to talk to your neighbour while in class
  - I try to be fair as possible and allow you opportunities to ask me about assessment marks, but please do not haggle about them
  - **Please don't cheat!**
- You -> Me
  - Menti

Communication Channels

Me → You

- Announcements are made regularly on Canvas
  - Please check Canvas regularly
- Emails will be sent out on a need basis
- Lecture announcements

You → me

- Consultation hours, email, before/after class, at labs and tutes if not helping others

Many ↔ Many

- Discussion Board on Canvas

Academic Integrity

- Please see:
  - https://www.rmit.edu.au/students/student-essentials/rights-and-responsibilities/academic-integrity
- You are encouraged to form groups to solve problems; however, when writing or programming, write in your own words or code, and provide your own solutions.
- Reasons:
  - Learn by trying things, making mistakes
  - Fair for everyone
  - Quality control of your degree

Social Media & Networks

Definition

Social Media & Networks are electronic and Internet tools for the purpose of sharing and discussing information and experiences with other human beings in more efficient ways.

Examples of Social Media & Networks

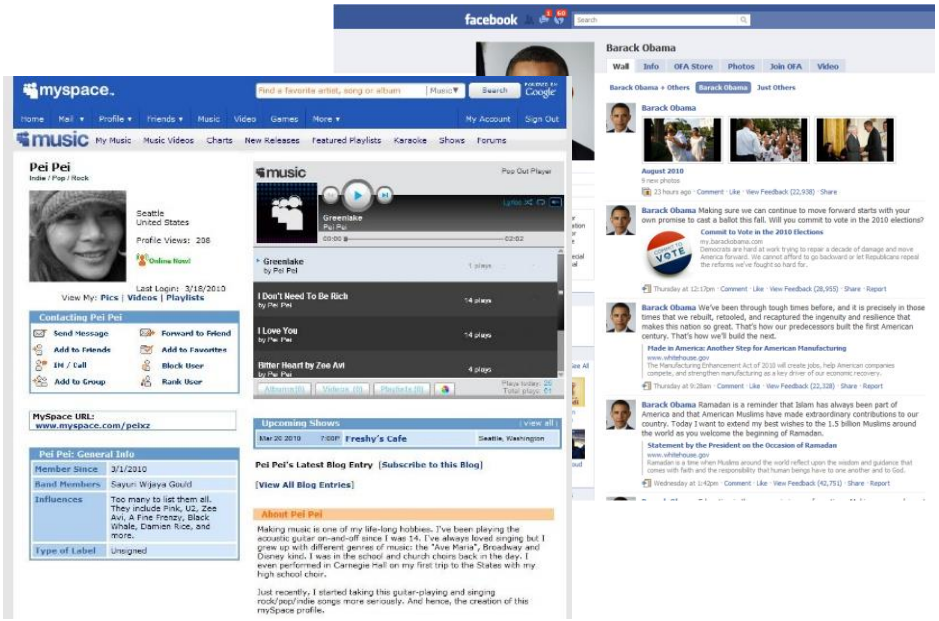
- Online Social Networking
- Publishing
  - Blogging
  - Wiki
- Micro blogging
- Social News
- Social Bookmarking
- Media Sharing
  - Video Sharing
  - Photo Sharing
  - Podcast Sharing
- Opinion, Review, and Ratings Websites
- Answers

Online Social Networking

Online Social Networks are web-based services that allow individuals and communities to connect with real world friends and acquaintances online

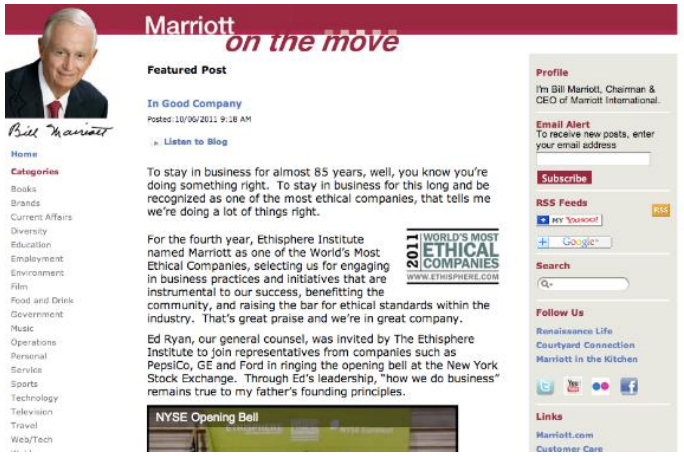
- Interactions
  - Friendship interaction
    - Friends, like, comments, …
  - Media Sharing
  - Sending and receiving messages

- Examples
  - Facebook.com
  - Bebo.com
  - Orkut.com
  - LinkedIn.com

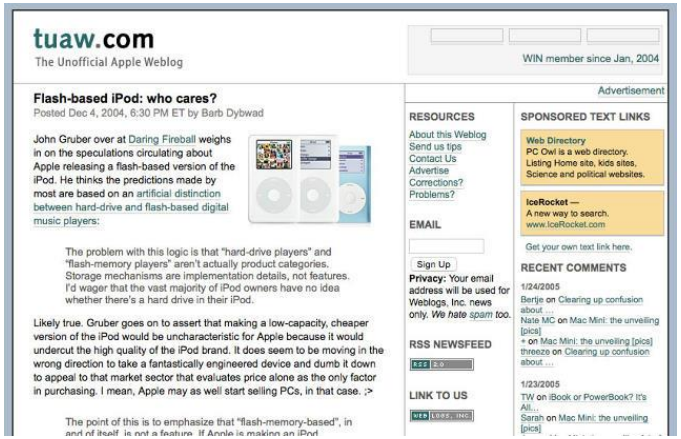


Blogging

A blog is a journal-like website for users, a.k.a. bloggers, to contribute textual and multimedia content, arranged in reverse chronological order



- Maintained both individually or by a community
  - See a tutorial at KDD http://videolectures.net/kdd08_liu_briat/



- Usages:
  - Sharing information and opinions with friends and strangers
  - Disseminating subject-specific content
  - Who is the influential http://videolectures.net/wsdm08_agarwal_iib/

Microblogging

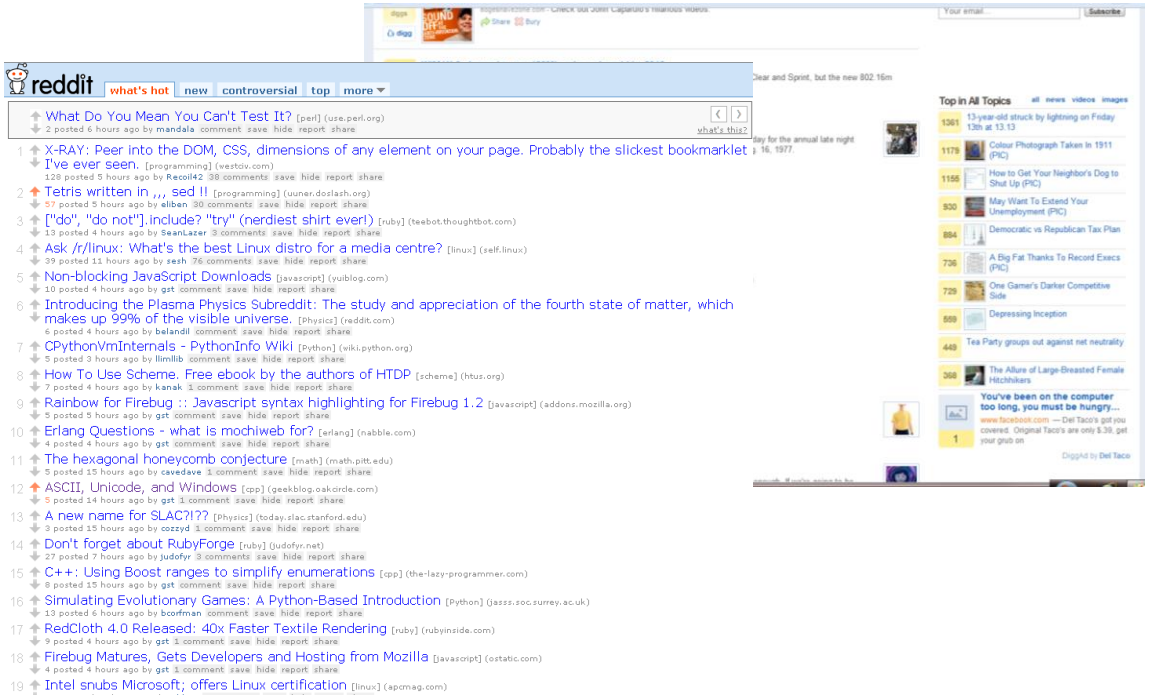
Microblogging can be considered as a counterpart to blogging, but with limited content

- Usage
  - communication medium
  - social interaction
  - citizen journalism
- Service Providers:
  - Twitter

Social News

Social News refers to the sharing and selection of news stories and articles by a community of users.

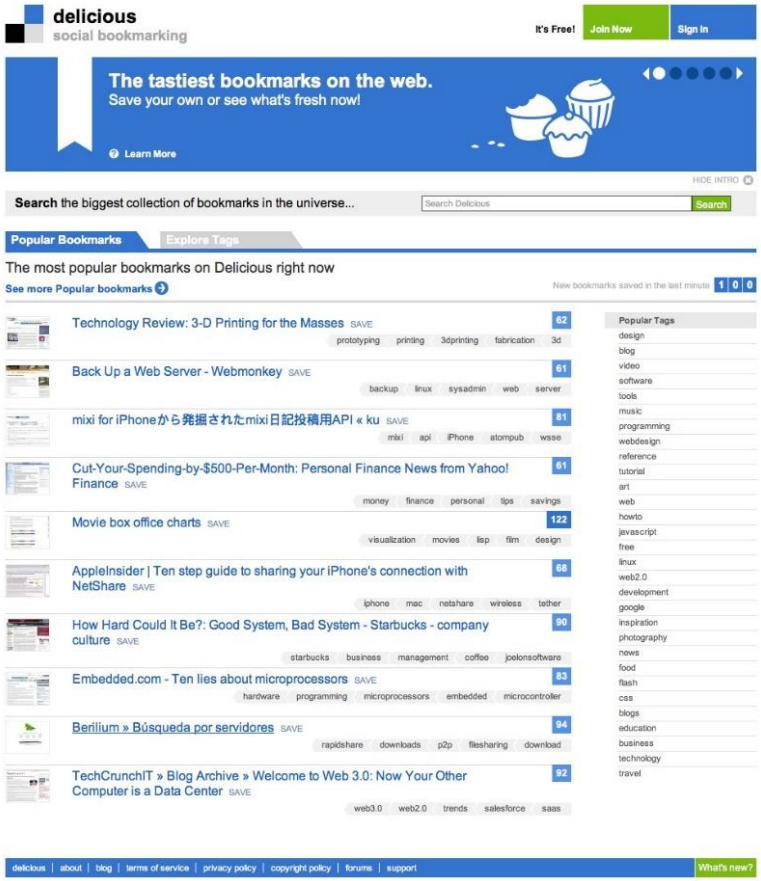
- Users can share articles that they believe would interest the community
- Samples:
  - Digg.com
  - Slashdot
  - Fark
  - Reddit



Social Bookmarking

Social Bookmarking sites allow users to bookmark web content for storage, organization and sharing.

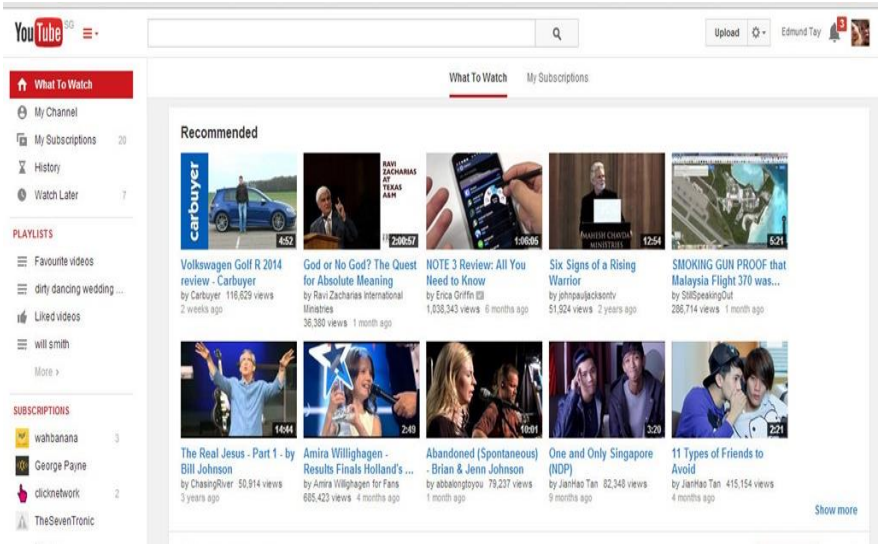
- These bookmarks can be tagged with metadata to categorize and provide context to the shared content, allowing users to organize information making it easy to search and identify relevant information.
- Samples
  - Delicious.com
  - StumbleUpon.com



Media Sharing

Media sharing is an umbrella term that refers to the sharing of a variety of media on the web.

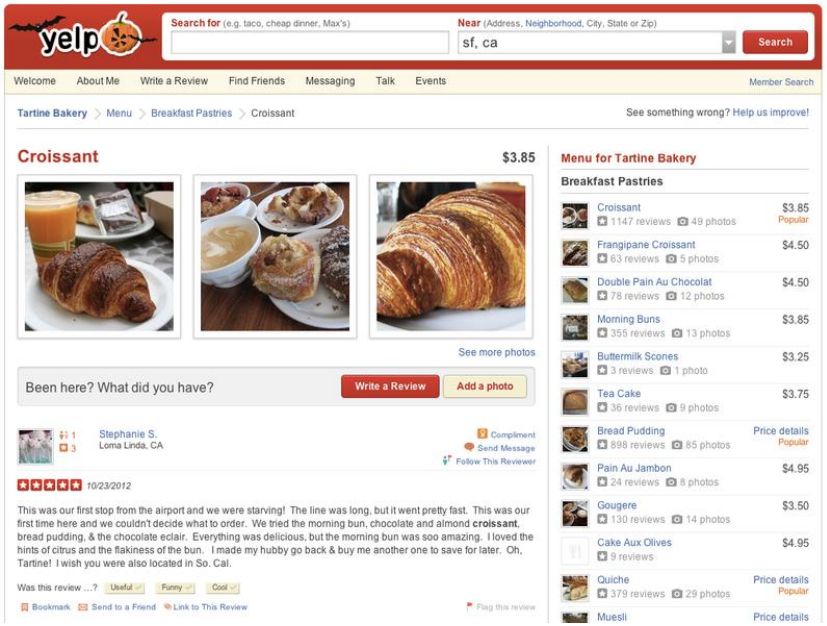
Users share such multimedia content of possible interest to others



- Samples:
  - Video Sharing:
    - YouTube.com
  - Photo Sharing:
    - Flickr.com
  - Document Sharing:
    - Slideshare.com
  - Livecasting:
    - Justin.tv, Ustream.com

Opinion, Review, and Ratings Websites

Opinion, review, and ratings websites are websites whose primary function is to collect and publish user-submitted content in the form of subjective commentary on existing products, services, entertainment, businesses, places, etc. Some commercial sites may serve a secondary purpose as review sites by publishing product reviews submitted by customers.



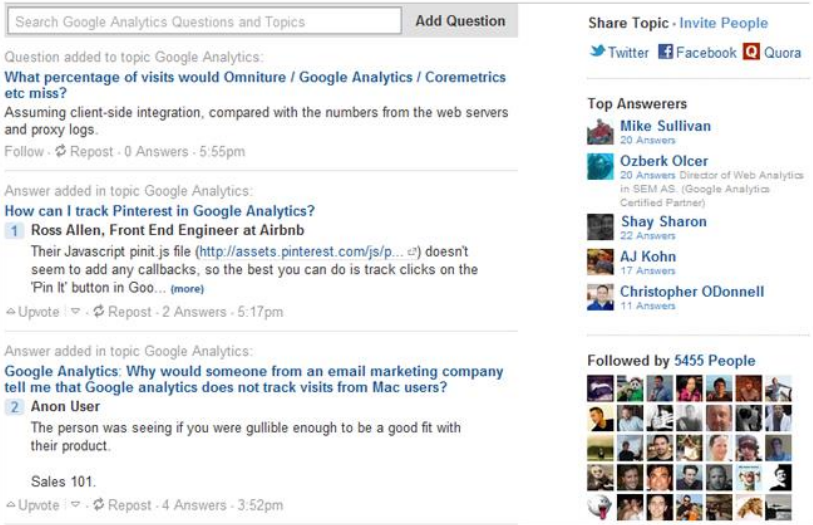
- Examples
  - Cnet.com
  - Epinions.com
  - yelp.com
  - tripadvisor.com

Socially-Provided Answers

In these sites, users who require certain guidance, advice or knowledge can ask questions. Other users from the community can answer these questions based on knowledge acquired from previous experiences, personal opinions or from relevant research.

- Unlike review and opinion sites, which contain self-motivated contribution of opinions, answer sites contain knowledge shared in response to a specific query.

- Samples:
  - WikiAnswers, Yahoo Answers, Quora



Main Characteristics

- Participation
  - social media encourages contributions and feedback from everyone who is interested. It blurs the line between media and audience.
- Openness
  - most social media services are open to feedback and participation. They encourage voting, comments and the sharing of information. There are rarely any barriers to accessing and making use of content – password-protected content is frowned on.
- Conversation
  - whereas traditional media is about “broadcast” (content transmitted or distributed to an audience) social media is better seen as a two-way conversation.
- Community
  - social media allows communities to form quickly and communicate effectively. Communities share common interests, such as a love of photography, a political issue or a favorite TV show.
- Connectedness
  - Most kinds of social media thrive on their connectedness, making use of links to other sites, resources and people.

Social Media & Network Analytics *is the process of representing, analysing, and extracting meaningful patterns and knowledge from social media and network data*

Types of Applications

- What is being discussed about my brand?
  - Is it positive or negative?
  - How are the reviews?
  - Who are the influential people driving news/perception about our brand?
- What is being discussed on my favourite blog sites or social news?
- Can we detect and identify abnormal (social) events, such as uprisings and riots?
- What about traffic incidents so we can bypass them to get to where we need to faster?
- Who are the key people in my communities whose absence will break them?
- …

Taster: Twitter Analytics

Winter Olympics: PyeongChang 2018

- February 26th, 2016, 1-5pm EST, ~4 GB of tweets
- #pyeongchang2018, #Olympics, olympics,
  https://twitter.com/search?q=%23pyeongchang2018&src=typd

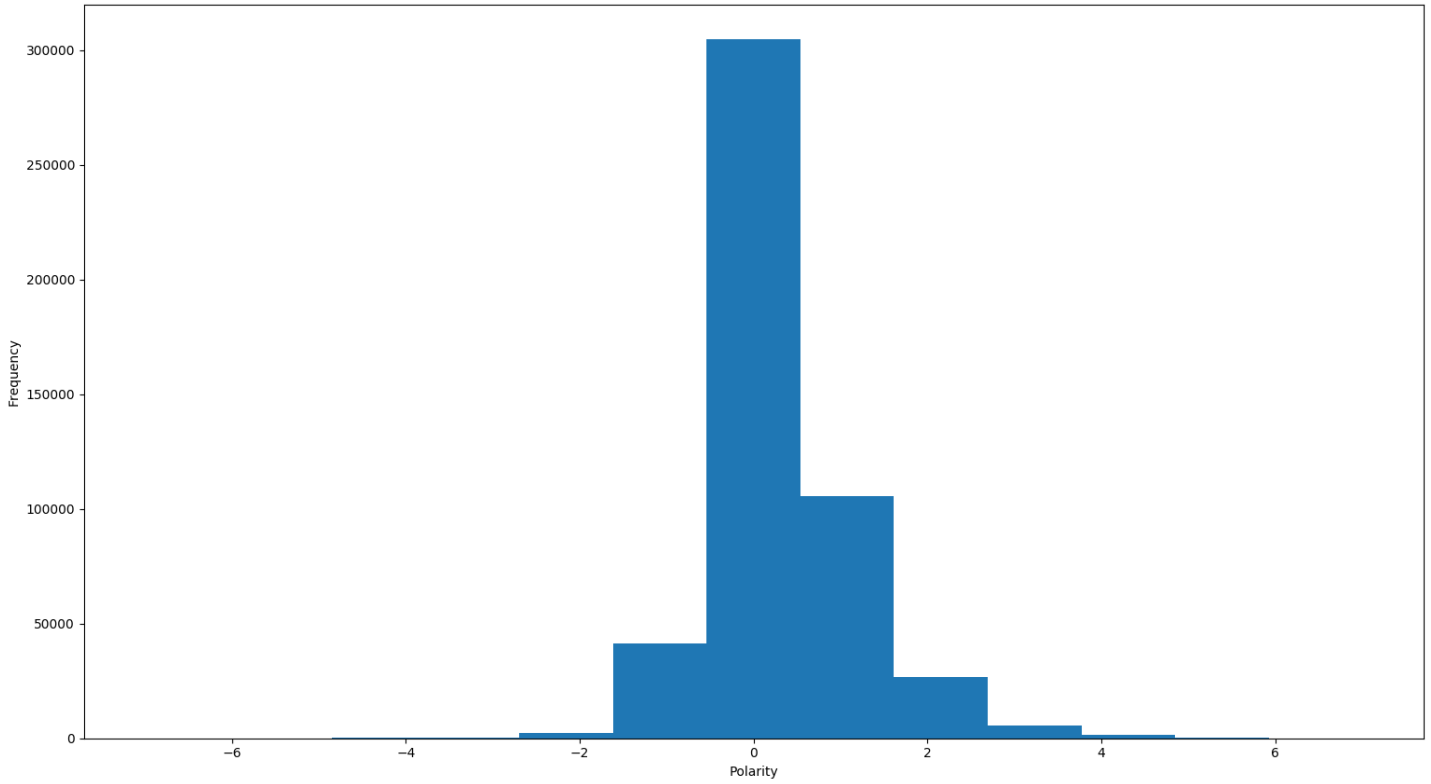
- What type of analysis we would be interested in?
  - Hashtags mentioned?
  - Sentiment analysis of tweets?
  - Text analysis (topics discussed)?
  - Distribution of tweets over time (trends)
- In addition:
  - Communities in Twitter?
  - Influentials?

Most Popular Hashtags in Tweets

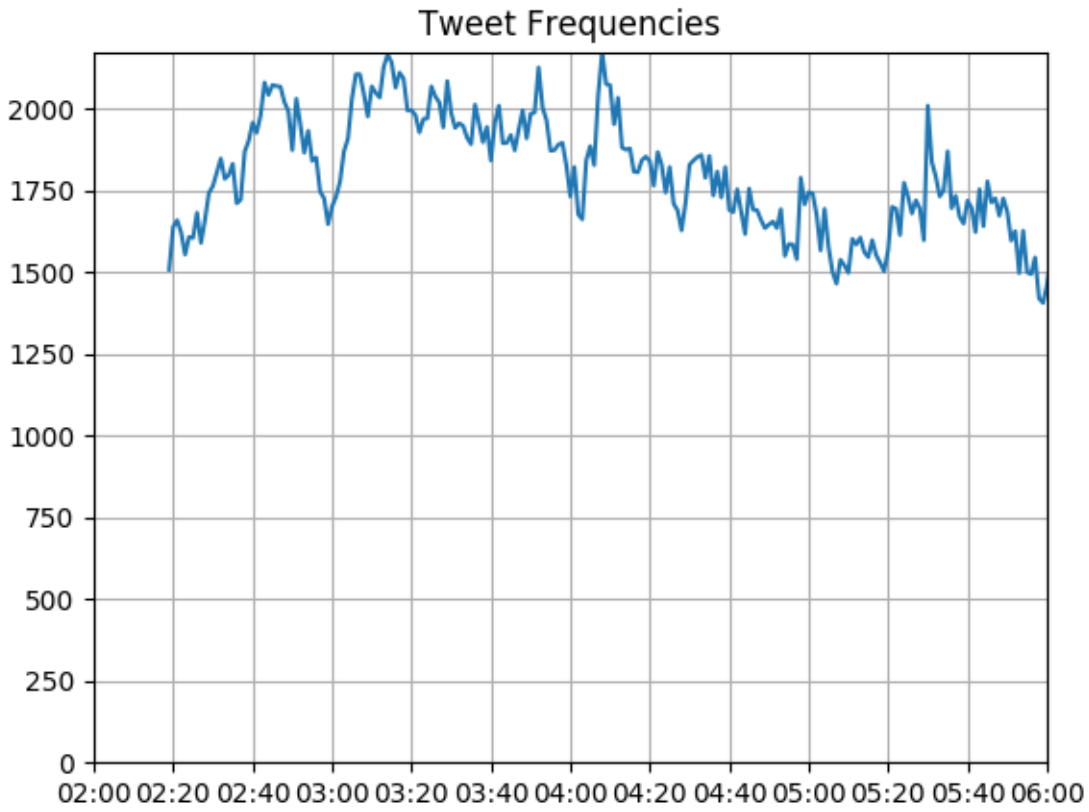
1. closingceremony: 168741
2. pyeongchang2018: 145139
3. olympics: 120974
4. exo: 94857
5. olympics_exo: 87605
6. exol: 71135
7. iheartawards: 49651
8. 엑소: 48791
9. bestfanarmy: 42084
10. exo_olympics: 23077

1. 폐막식: 15241
2. weareone: 12731
3. 2018평창동계올림픽: 11594
4. soohorang: 8085
5. 올림픽: 7284
6. kpop: 7231
7. exo_primetime: 6788
8. winterolympics: 6527
9. shazam: 6393
10. power: 6037

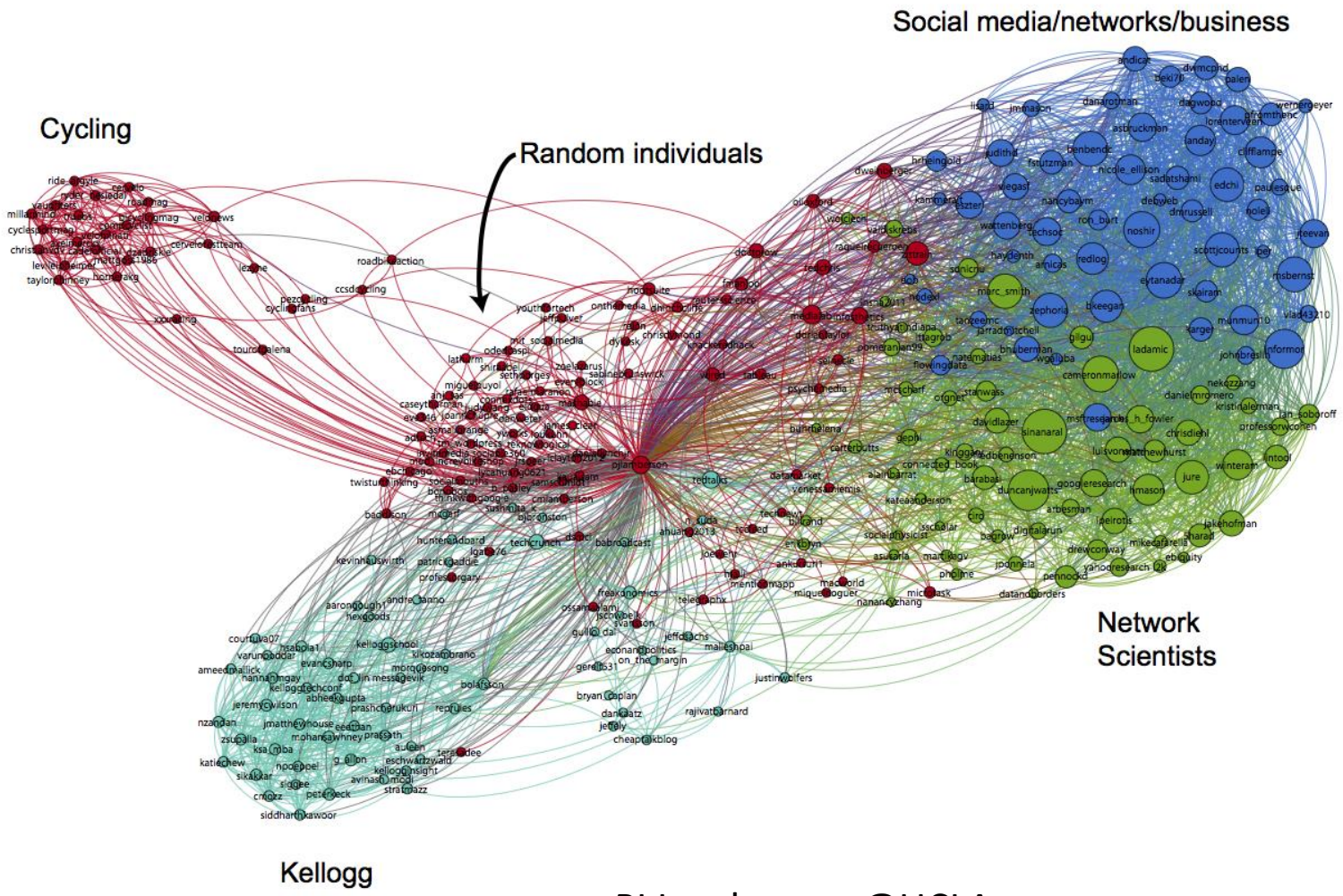
Distribution of Sentiment of tweets



Timeseries

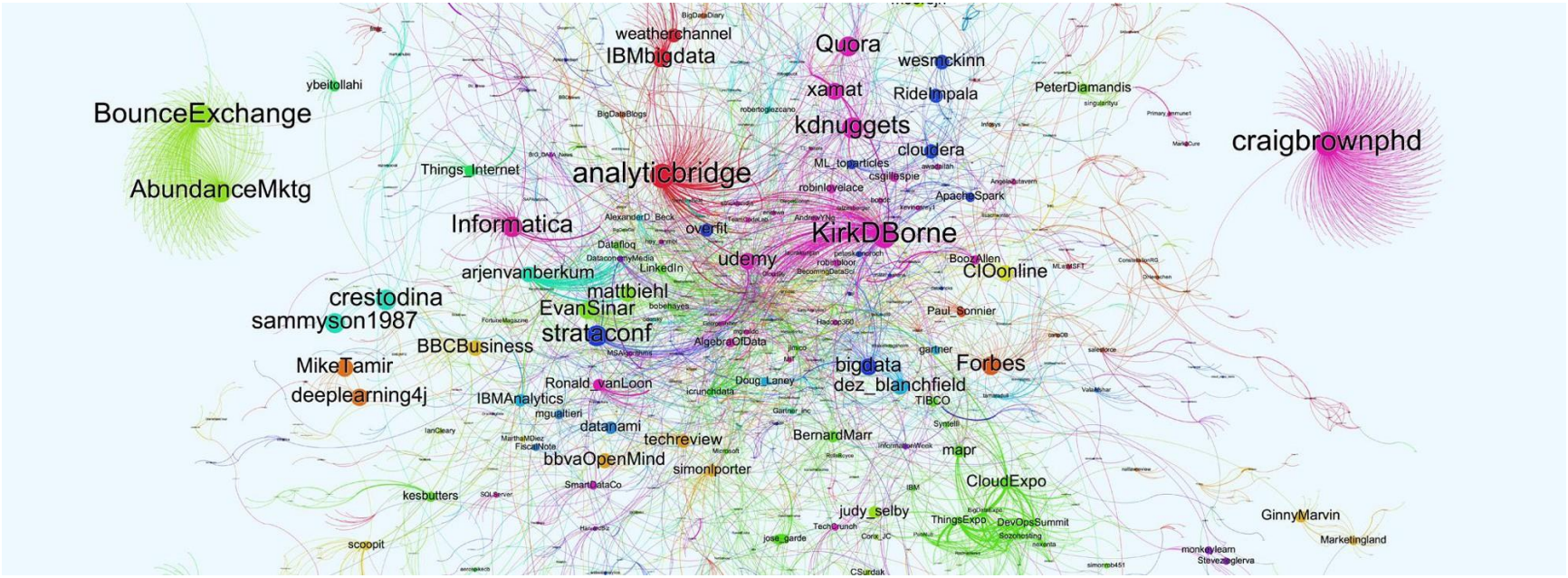

Tweet Frequencies
2000
1750
1500
1250
1000
750
500
250
0
02:00 02:20 02:40 03:00 03:20 03:40 04:00 04:20 04:40 05:00 05:20 05:40 06:00

Communities on Twitter



PJ Lamberson @UCLA
http://social-dynamics.org/twitter-network-data/

Influentials for data science & big data



John Swain
https://medium.com/@swainjo/an-analysis-of-twitter-influencers-in-the-field-of-data-science-big-data-48b0809027ef

Summary

- Introduced Social Media and Networks
  - Types
  - Applications
- Demonstrated some types of analysis for Tweeter and other media/networks

TODO Items

- 3 To-do items for you:
  - Familiarize yourself with Canvas
  - Revise: pre-processing of data and basic data science techniques
    - Load and do basic pre-processing of data
    - Understand and can run basic machine learning algorithms (practical data science)
    - Basic knowledge of statistics
  - Do the prescribed readings and video watching for next week