Social Media Analytics - CS-EJ5621

Lecture 6

Aqdas Malik (aqdas.a.malik@aalto.fi)

16.10.2020

Aalto University
School of Science

FITech
NETWORK UNIVERSITY

Course practicalities

- **Missing quizzes**

- **Quiz 4 due today at 2359**

- **Final lecture next week**

Agenda

- **Content analysis**
- **Sentiment analysis**
- **Social network analysis**

Content analysis - Definition

- **Systematic procedure devised to examine the content of recorded information (Walizer & Wienir, 1975)**

- **A method for studying and analyzing communication in a systematic, objective, and quantitative manner (Kerlinger, 1986)**

Content analysis as a method

- A research method that investigates different forms of content

- Determines the presence of certain words, themes, or concepts within the given data (e.g. text, videos, photos)

- Quantify and analyze the presence, meanings and relationships of such certain words, themes, or concepts

- Draw inferences about the messages within the content, the writer, the audience, and to some extent the culture and time of surrounding the content

Why content analysis?

- **Reduces and gives insights to large amount of unstructured data**

- **Patterns of human behaviors and attitudes can be studied indirectly through analyzing their communication**

- **Helps in determining the presence of certain words or concepts within multiple content forms (written, spoken, or visual communication)**

- **Multiple channels: Newspapers, magazines, novels, ads, textbooks, songs, movies, social media**

Content analysis of social media

- **Original posts**

- **Comments**

- **Photos**

- **Videos**

- **URLs**

- **Hashtags**

Uses of content analysis

- **Describing communication content**

- **Testing hypothesis of message characteristics**

- **Comparing media content with real world**

- **Assessing the image of a particular group in society**

- **Revealing international differences in communication content**

Describing communication content

- **Identify "what actually exists" in the content**

- **Catalogs the characteristics of a given body of communication content**

- **Example: What type of images are used by Aalto University on Twitter?**

A? **Aalto University**
@AaltoUniversity

Prior to EAIE conference in Helsinki, the conference participants from 35 countries visited Otaniemi campus during Aalto University Partner Day & Campus Experience Day ⭐ #EAIE2019 #aaltocampus #aaltouniversity

5:58 PM · Sep 24, 2019 · Sprout Social

2 Likes

Testing hypothesis of message characteristics

- Examines the relationship between two variables

- Example: Are females more likely to participate in anti-smoking campaigns on Facebook than males?

  Variable A: Female/Male
  Variable B: Commenting (Yes/No)
  Variable C: Liking (Yes/No)
  Variable D: Sharing (Yes/No)

Comparing media content with real world

- Reality checks of portrayal of certain groups, traits, or phenomenon in content vs. the standard taken from real life

- Example: Number of Asians in job advertisements compared to the US population %

Assessing the image of a particular group in society

- **Explore the image and perception of certain minorities or groups**

- **Helps in assessing the policy or document social trends**

- **Example: Representation of Latino/Hispanic immigrants on the Democrats social media pages**

The Democrats ✔ @TheDemocrats · Apr 2
In the spirit of #FactCheckingDay, we've got one of our own:

Trump rescinded DACA and has struck down every bipartisan agreement from Congress. Because of him, Dreamers like Carlos are at risk of being deported and live in uncertainty every single day.
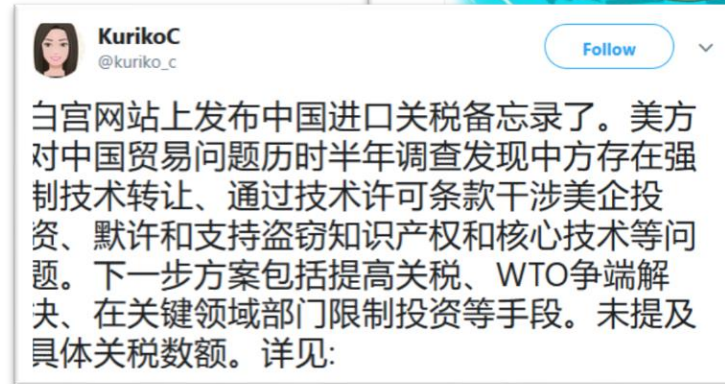
My name is
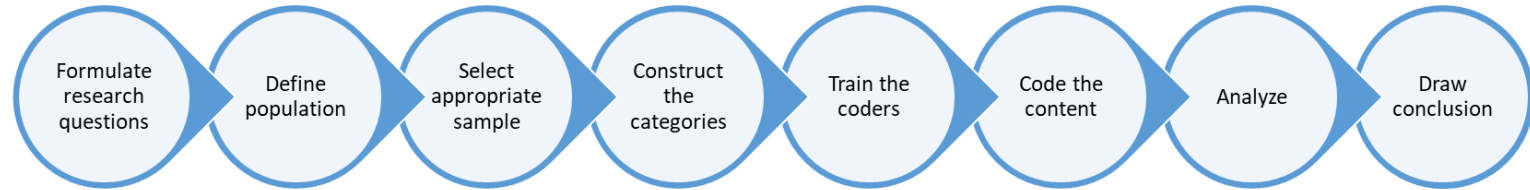Carlos Adolfo Gonzalez.

💬 60   🔁 153   ♡ 222   ✉

Revealing international differences in communication content

- **Reaction of people or groups across different regions on a specific topic?**

- **Example: Donald Trump's tariff plans on China**

Content analysis process

Formulate research questions → Define population → Select appropriate sample → Construct the categories → Train the coders → Code the content → Analyze → Draw conclusion

Pros and cons

- **Easy & inexpensive**

- **Does not involve the use of people as subjects**

- **No special equipment/system required**

- **Unobtrusive: observer has no effect upon the material being collected**

- **Qualitative and quantitative analysis**

- **Historical and cultural insights over time**

- **Limited to communication analysis**

- **Descriptive and not explanatory**

- **Tedious and time-consuming**

- **Subject to increased error**

- **Analysis may not truly reflect reality**

- **Reliability issues**

Content analysis example

Category	Definition	Example	Theme count (%)
Distribution of Tweets by Theme[a]			
Promotional	Sale, coupon, voucher, discount, offer, review or other activities promoting Juul in general	10% off on #juul and #juulpods plus 15% off on all #vapes and #juices	89 (7.9%)
Flavors	Tweet discussing or promoting different flavors	I am falling in love with the #mango #juul #pod. This is absolutely refreshing	78 (6.9%)
Features	Tweet highlighting or promoting design related aspects of Juul (eg, size, usb charging etc.)	Just bought a super charging kit for my #juul as a backup ;) this thing charges quick	43 (3.8%)
Advocacy	Tweet glorifying or signifying Juul use by attributing it as not harmful, attractive, cool, or trendy	Fully charged and a new pod. @myfriend lets meet at the #Lisbongrounds to rock and roll with #juul.	319 (28.2%)
Tips and queries	Tweet mentioning tips, trick, hacks, and support related to Juul	Is it possible to get the cheap refills for #juul as the original ones are too expensive? And how to refill?	91 (8.0%)
Smoking cessation	Using Juul to quit smoking or other tobacco products	#Juuling has been great so far. It's been 3 weeks since I smoked a cigarette, and the credit goes to #juul	42 (3.7%)
Fun	Joke, humor, meme, or funny tweet	A new rhyme in our club: Lets juul and be cool at the school. (emojis)	37 (3.3%)
Prevention	Preventive measures, legislations, policy, regulations, educational programs	Please mark your calendar for our latest webinar on #juuling among #teenagers and #FDA. August 16, 12-2 pm Eastern hosted by Abraham Wellford and Jenny Krispock. Register: https://t.co/Njy64645sde2	92 (8.1%)
General health and safety	Health consequences, concerns, dangers, threats, or risks in general	It is quite risky to engage in #juuling as we are still not sure about the long term side effects of #vaping products. Stay safe	43 (3.8%)
Youth health and safety	Health consequences, concerns, dangers, threats, or risks among underage/youth in schools/colleges	Just met a panel of principals from @Mills_county_ schools and all of them synonymously agreed that #juul and #vaporizer is likely to cause health epidemic among our new generation.	183 (16.2%)
Illicit substance use	Mention of an illicit substance use with Juul	My sister who is still 16 recently started using #CBD with her #juul. Is this dangerous?	17 (1.5%)
Discouraging use	Tweet discouraging or rejecting the use and adoption of Juul	#Juul is one of the most hated, unattractive, and dangerous piece of technology I have ever seen. Stay away from #juuling	46 (4.1%)
Research	Research / scientific information / publication / statistics	The latest results from our survey reveal that #juuling among youth has increased gradually since its launch in 2015. https://t.co/ee9434dfgh6	13 (1.1%)
Miscellaneous	Tweet that cannot be classified in any of the above categories	I was sitting at the beach and heard guys playing volley and having juul.	40 (3.5%)

Note.
a: Non-exclusive category

Malik, A., Li, Y., Karbasian, H., Hamari, J., & Johri, A. (2019). Live, love, Juul: User and content analysis of Twitter posts about Juul. American journal of health behavior, 43(2), 326-336.

Content analysis methodologies

- **Manual analysis**

- **Word clouds**

- **N-Gram**

- **Linguistic Inquiry and Word Count (LIWC)**

- **Topic modelling (Latent Dirichlet Allocation - LDA)**

Sentiment analysis (opinion mining/emotion AI)

- Process to determine the emotional tone behind the content
- Helps in understanding the expressed opinions
- Gauge public opinions, brand reputation, customer experience
- Weighted sentiment scores is assigned
- Positive, neutral, negative

Sentiment analysis (manual)

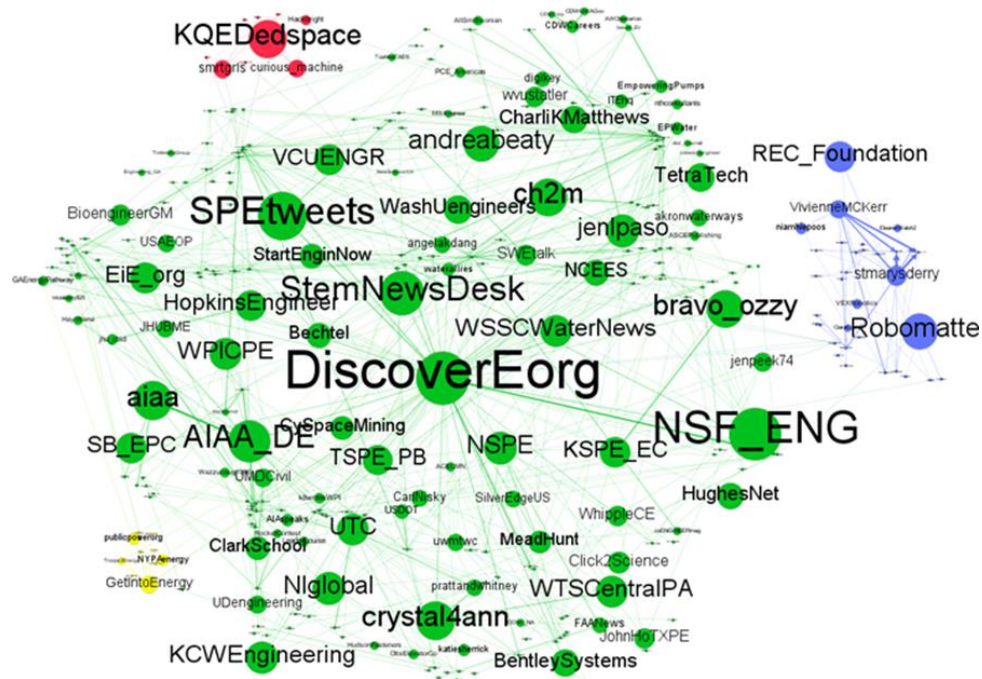# Sentiment analysis (automated)

- **SentiStrength**

- **ETCML**

- **Mozdeh**

- **SentimentViz**

Social Network Analysis
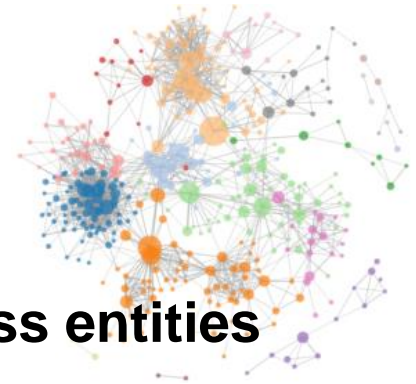
- **Method to analyze the social connections across entities**

- **Helps in understanding how individual, organizations, and political actors interrelate and interact**

- **Who influences whom?**

- **Unable to detect individual behaviors**

- **Tools: Gephi, NodeXL, Pajek, R, UCINet**

#NationalEngineersWeek

Gephi tutorials

https://seinecle.github.io/gephi-tutorials/

https://www.youtube.com/watch?v=2FqM4gKeNO4

MineMyText demo

- **Natural language pre-processing**

- **Topic modelling**

- **Sentiment analysis**

- **Visualization**
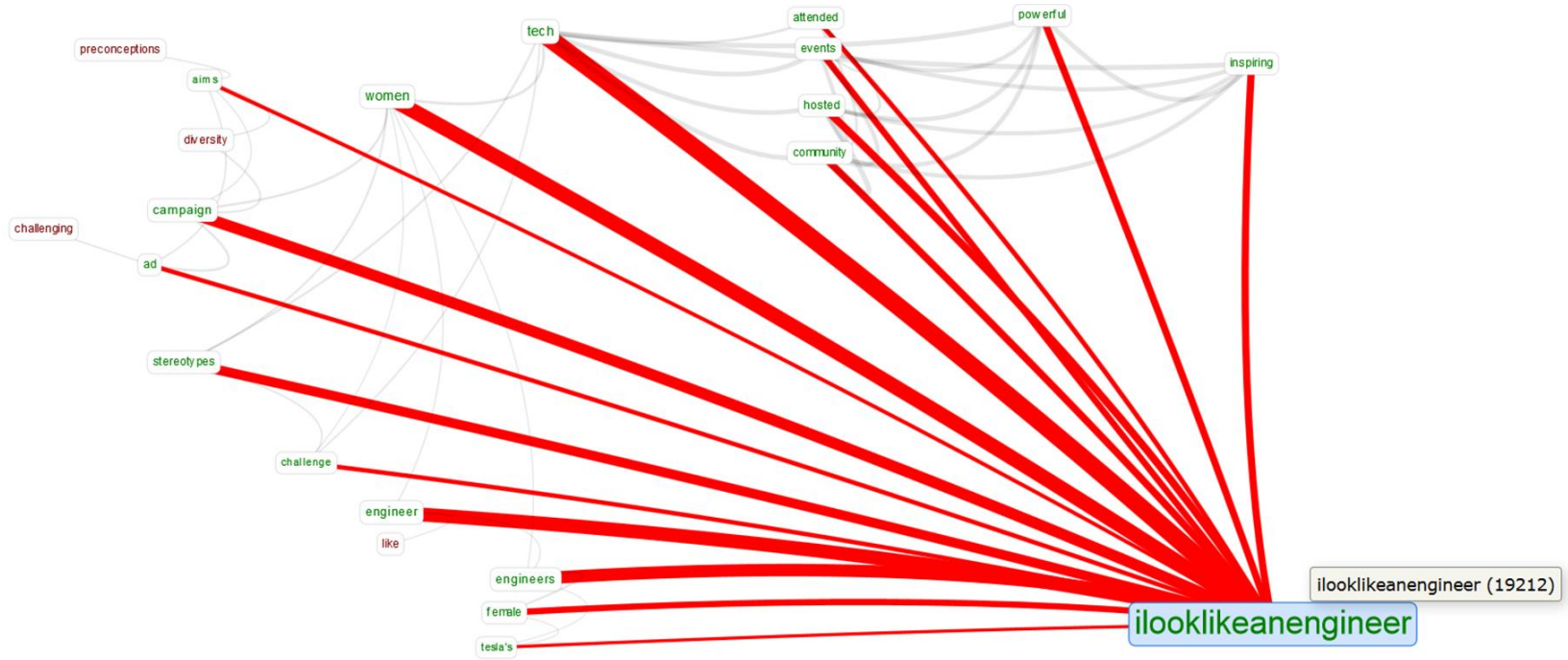

- **https://app.minemytext.com**

Netlytic.org

- **Data capture, analysis, and visualization**
- **Facebook, Twitter, YouTube, and Instagram**
- **Text analysis (keywords)**
- **Discussion themes**
- **Geo-coded data**
- **SNA**

voyant-tools.org/

- **Suite of text-analysis tools**
- **Word cloud**
- **Number of words**
- **Word trends**
- **Keyword in context**
- **Highly interactive**

Quick overview of keywords and relationship

NodeXL.com

- **Network overview, discovery, and exploration for Excel**

- **Data extraction – Facebook, Twitter, Youtube, Instagram**

- **Network graph visualization tool**

- **Most frequent hashtags, domains, URLs, words, word Pairs, mentioned users, replied-to, and most frequent tweeters**

NodeXL - #iOS11
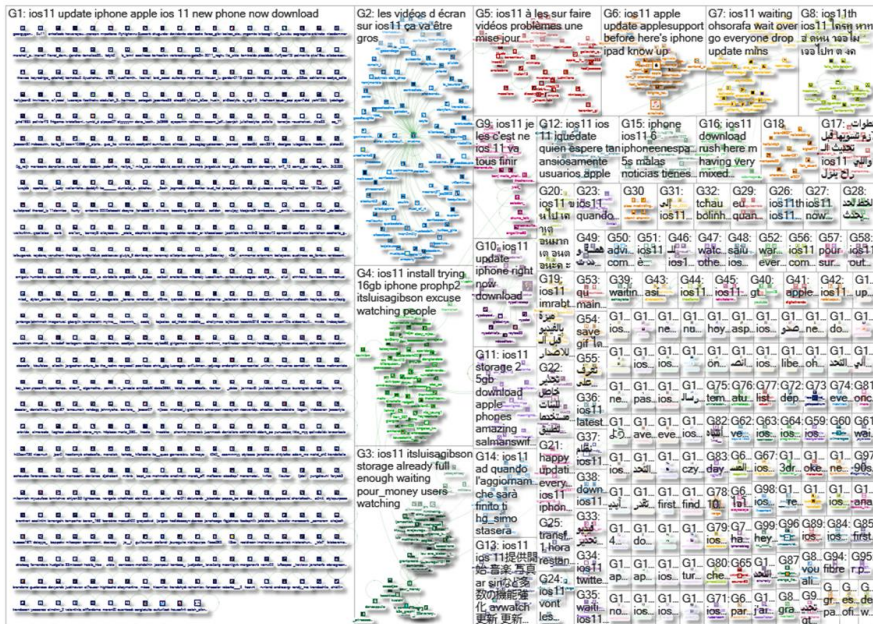
- 1,262 Twitter users

- Top hashtags: ios11, apple, ios11th, iphone, ios, ios9, ios10, ipad, watchos4

- 352 words: Positive sentiment

- 103 Words: Negative sentiment

- Top influencers: @prophp2, @mvximx, @itsluisagibson, @mayadorable, @devierganael

- @shawnmfanboy101, @marcoasenseo, @0ydui, @daralingx, @apple__fr_

Next lecture – 23.10.2020

- **User analysis**

- **Links analysis**

- **Temporal analysis**

- **Hashtags/words analysis**

- **Photo/video analysis**

Thank you