Social Media Analytics - CS-EJ5621

Lecture 4

Aalto University
School of Science

FITech
NETWORK UNIVERSITY

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

02.10.2020

Contents of the lecture

- Course practicalities
- Native vs. non-native analytics
- Metrics for analyzing social media data
- Organizing gathered data & cleaning
- Bots and how to detect them
- Hands-on: Number of retweets and favorites in TAGS
- TAGS Analytics

Course practicalities

- **Quiz 3 due by 2359 today**
- **#climatechange NOT #climate AND #change**
- **Focus on a single # *(or a few @)***
- **Twitter developer account**
- **Other technical issues**

Native vs. non-native analytics

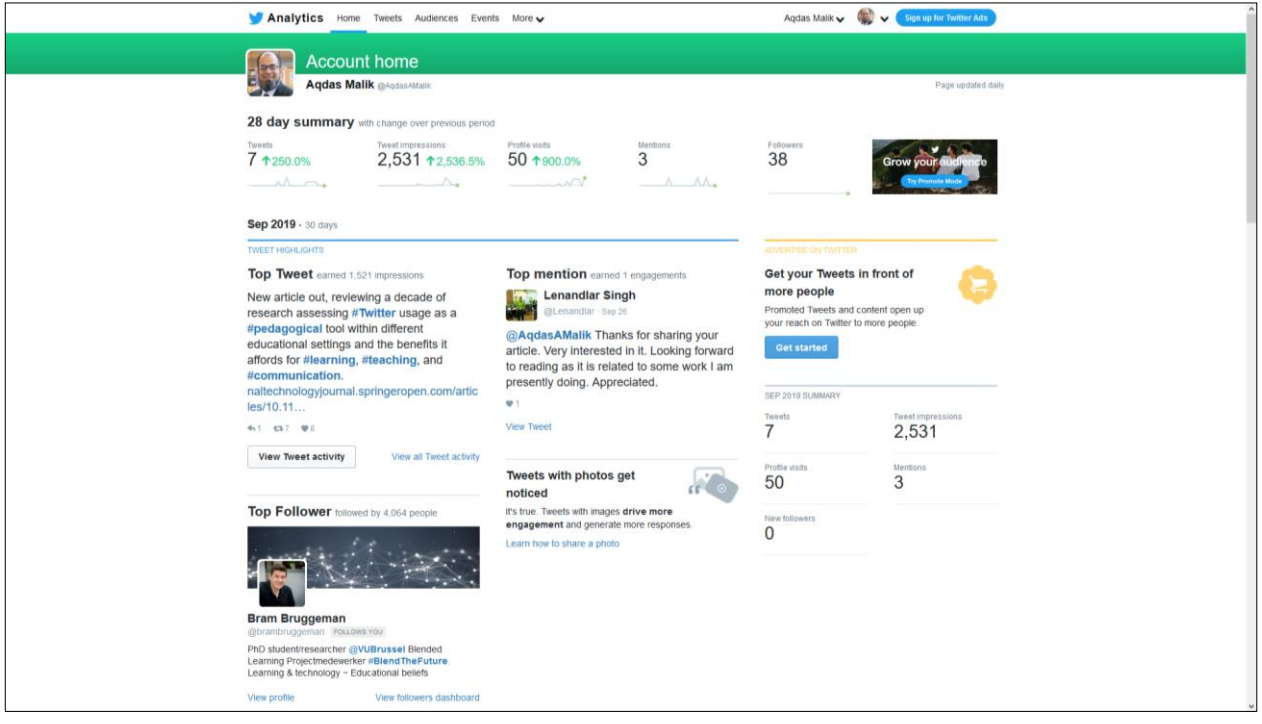
Access and monitor the current and historical performance of your own page, group, or channel through a number of metrics

- Twitter analytics
- Facebook insights
- Instagram insights
- YouTube Analytics

External applications that connect with respective API's to provide additional analysis of your/other content

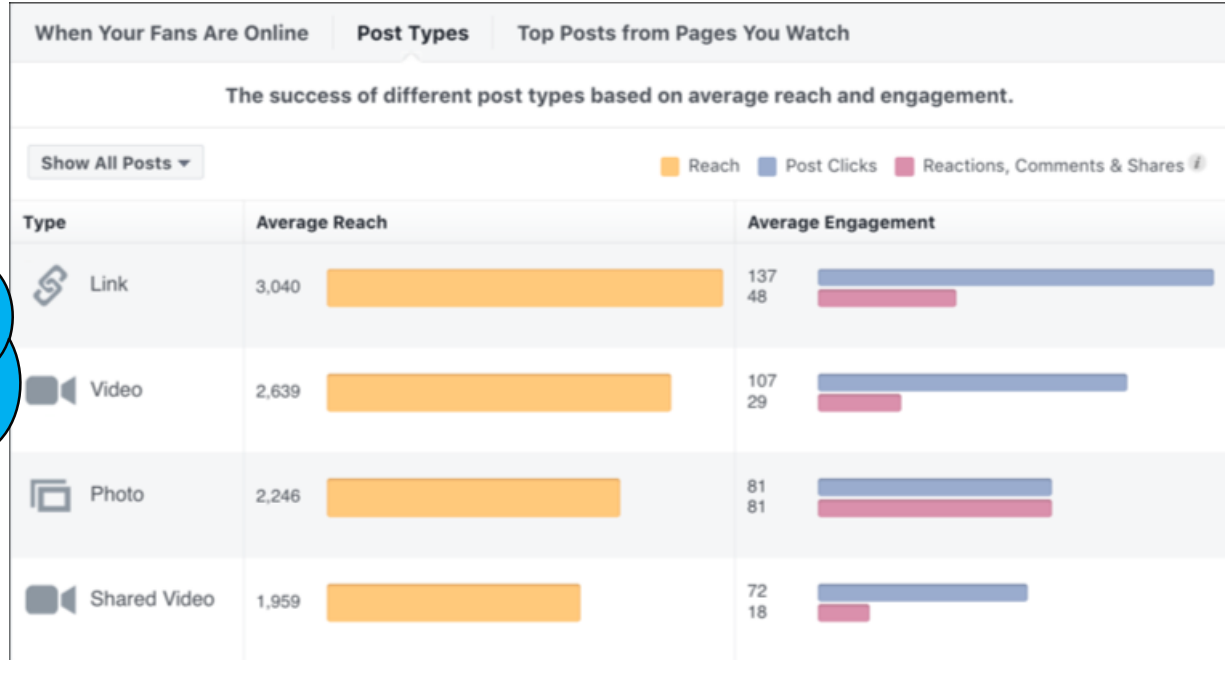
- TAGS
- Chorus
- Cosmos
- Mozdeh
- Paid social media analytics platforms (e.g. Brandwatch, Hootsuite, Keyhole)

Twitter analytics walkthrough



Facebook insights

Performance of different content types

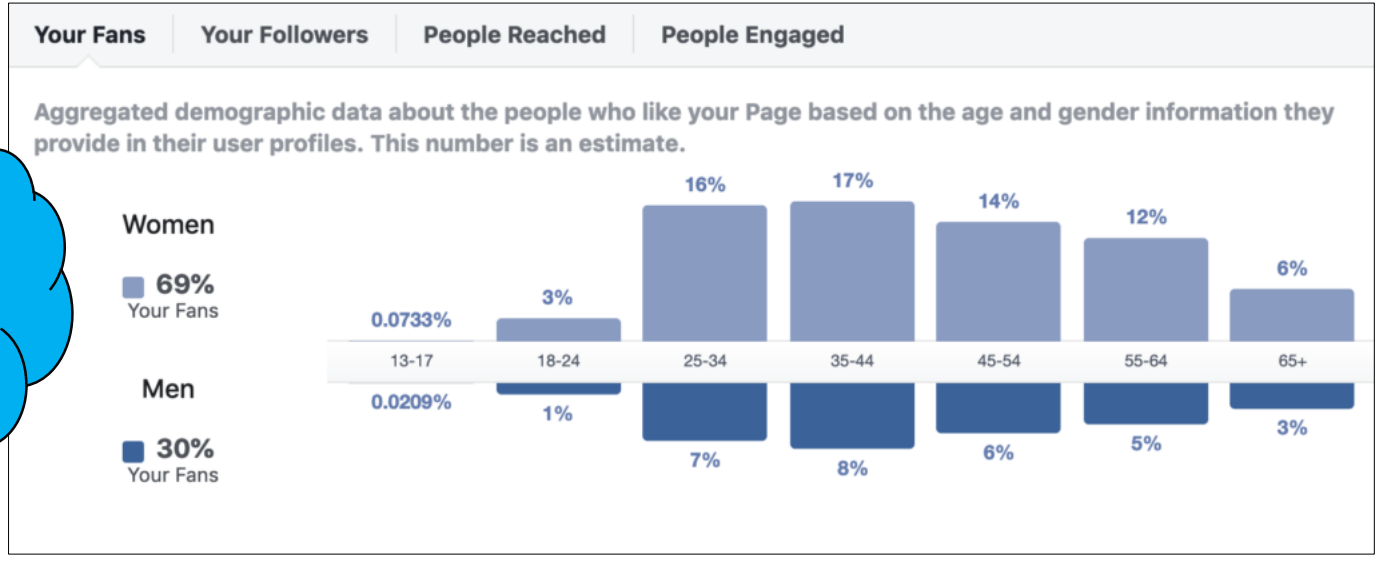


Aalto University
School of Science

source: https://blog.stuller.com/facebook-insights-social-strategy/

Facebook insights

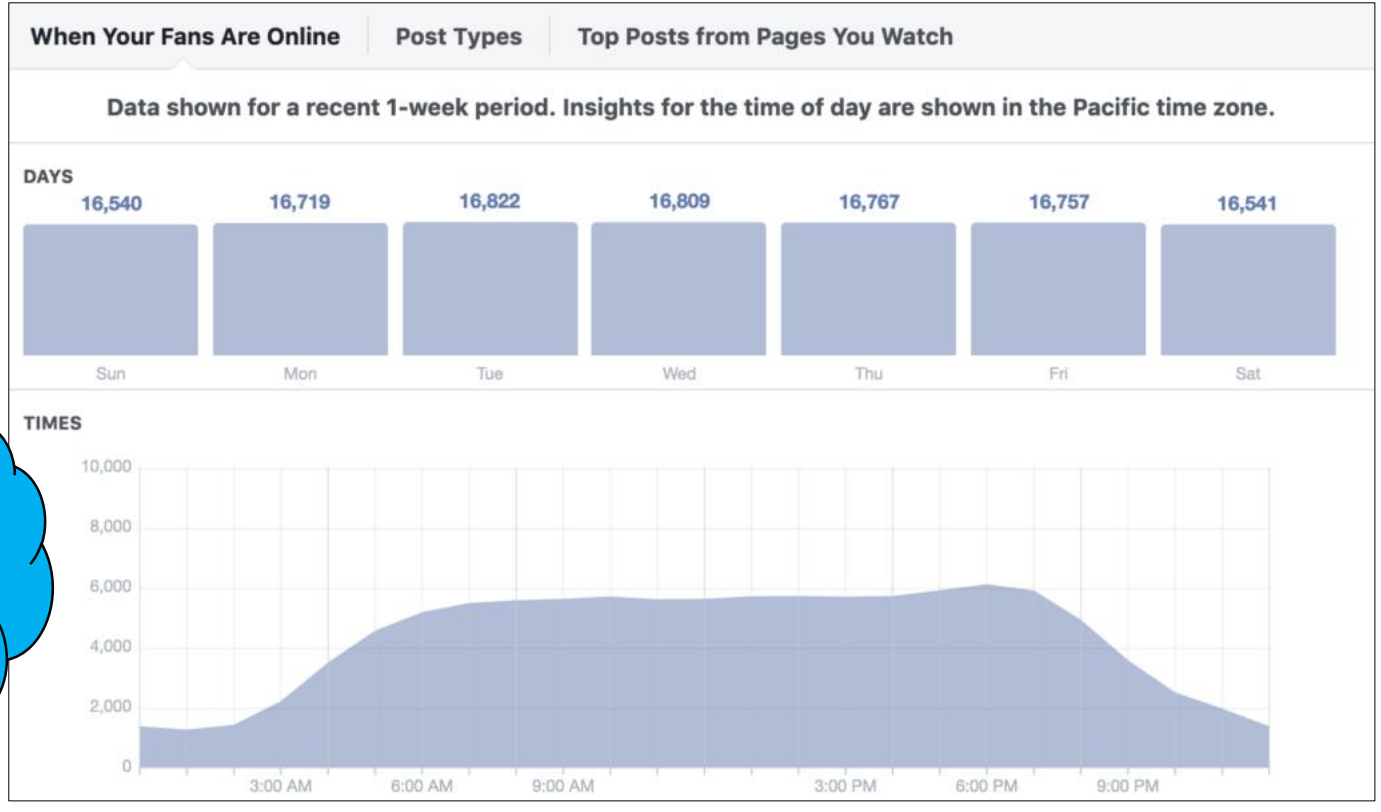
Understand your audience



Aalto University
School of Science

source: https://blog.stuller.com/facebook-insights-social-strategy/

Facebook insights

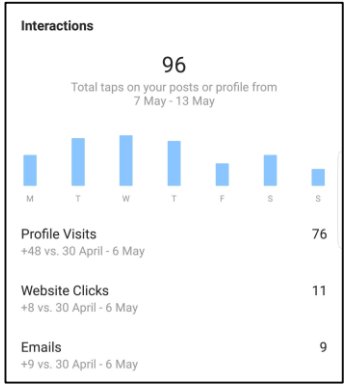


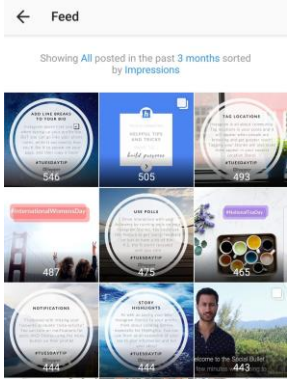
Best times to post

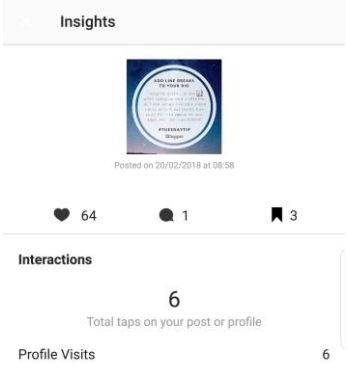
Aalto University
School of Science

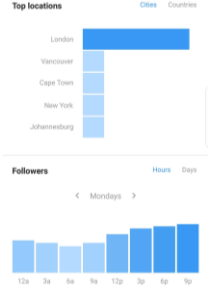
source: https://blog.stuller.com/facebook-insights-social-strategy/

Instagram insights





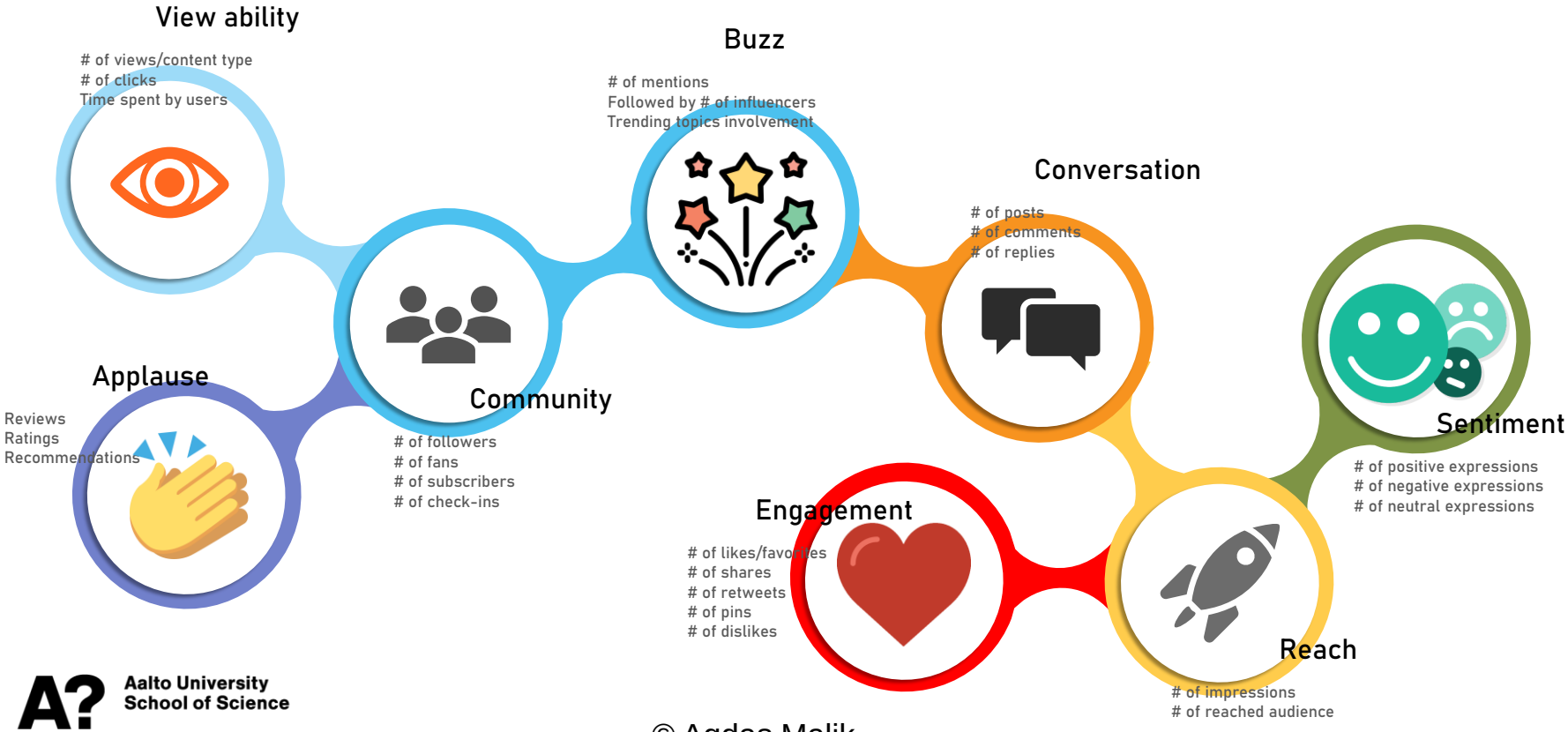




Aalto University
School of Science

source: https://blog.stuller.com/facebook-insights-social-strategy/

Metrics/indices



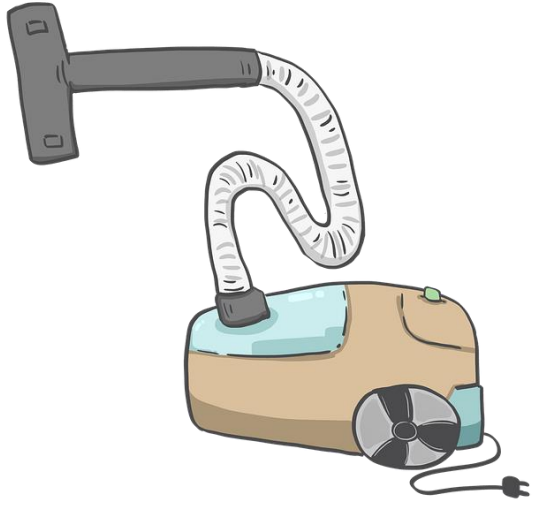
Aalto University
School of Science

Organizing data

- Specify analytical timeframe
- Specify analytical sample
- Do not edit original files
- Convert units/formats (if required)
- Keep track of all the actions
- Only keep the data attributes that are required
- Might need to switch between Google sheets and Excel

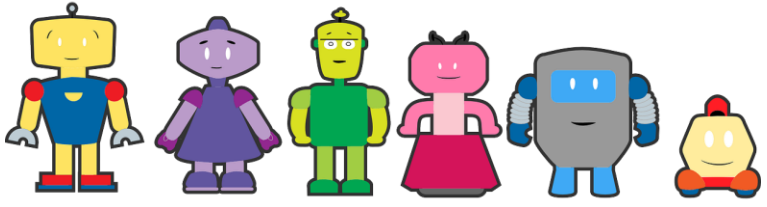
Cleaning data

- Non-English tweets
  `DETECTLANGUAGE, GOOGLETRANSLATE`
- Retweets
- Extracting mentions (@)
- Extracting URL's (https://)
- Extracting hashtags (#)
- Duplicate tweets (mostly bots)
- Only hashtags?
- Deleted tweets
- Suspended accounts
- Irrelevant

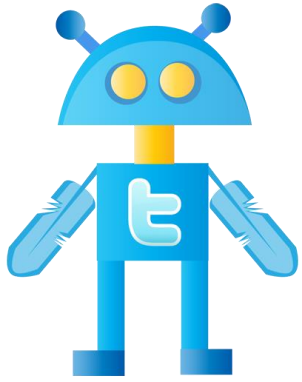


What is a Bot

- An automated software that controls and acts as a social media user account
- Engages in automated tasks (simple, repetitive, high volume)



Social bots (fake accounts)



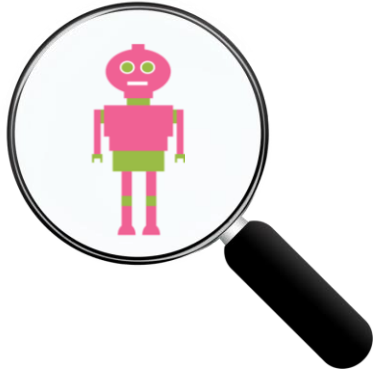
- Automated content creation
- Convincing persona to trick other users within the network as a real and legitimate entity
- Builds trust to spread within the network
- Quite a common phenomenon on Twitter (due to its affordances)
- 5–10 % of all users are bots, 20–25 % of all tweets are auto-generated *
- Aim: gain followers, spread fake news, propagate ideology, hate speech, influence opinions

*source: https://lnu.se/en/meet-linnaeus-university/current/news/2019/new-application-can-detect-bot-posts-on-twitter-in-any-language/

How to detect Twitter bots?

Manual detection

- High followers and following, user names with numbers, active retweeting, low original tweets, repeated content, no bio or profile image



BotoMeter

BotCheck

TweetBotOrNot (R)

Hands-on: Number of retweets and favorites in TAGS *(and analytics)*

- **Add two new columns**
- **Title: `favorites_count` and `retweet_count`**
- **Likewise, other tweet attributes can also be added if needed (https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/tweet-object)**

Note: This action is not retrospective, as the counts will be captured for new data only

Next lecture – 09.10.2020

- **Social media data pre-processing**
- **Social media text analytics**
- **Sentiment analysis**

Thank you