MEANING INTERPLAY BETWEEN EMOJIS AND LINGUISTIC TEXT

Ye Tian

Work done during postdoc at Université Paris Diderot

- We use social media not only to share information, but also to share and to express emotions.
- We can express them using language, but also nonverbal elements such as interjections (“wow”, “hahaha”) and emojis.
- Today I’m going to talk about how Facebook reactions and emojis reflect users’ emotional states.

Emoji factoids

- Emoticons, such as ";)", are representations of facial expressions using punctuation symbols.
  - First used by the computer scientist Scott Fahlman in 1982.
- The word emoji does not derive from emotion
- Loan word from Japan where they originated
  - comes from *e* 'picture' + *moji* 'letter, character'.
- graphic symbols representing facial expressions (e.g. smiling), gestures (e.g. thumbs up), objects (e.g. vehicles) and even actions (e.g. dancing).
- Emojis have gained popularity rapidly in smart-phone texts, emails and social media.
- On certain platforms (e.g. Instagram), over half of all online messages contain emojis in some countries (e.g.Finland and France), and this trend is going up worldwide (Dimson, 2015).

Font effects – from unicode.org/emoji

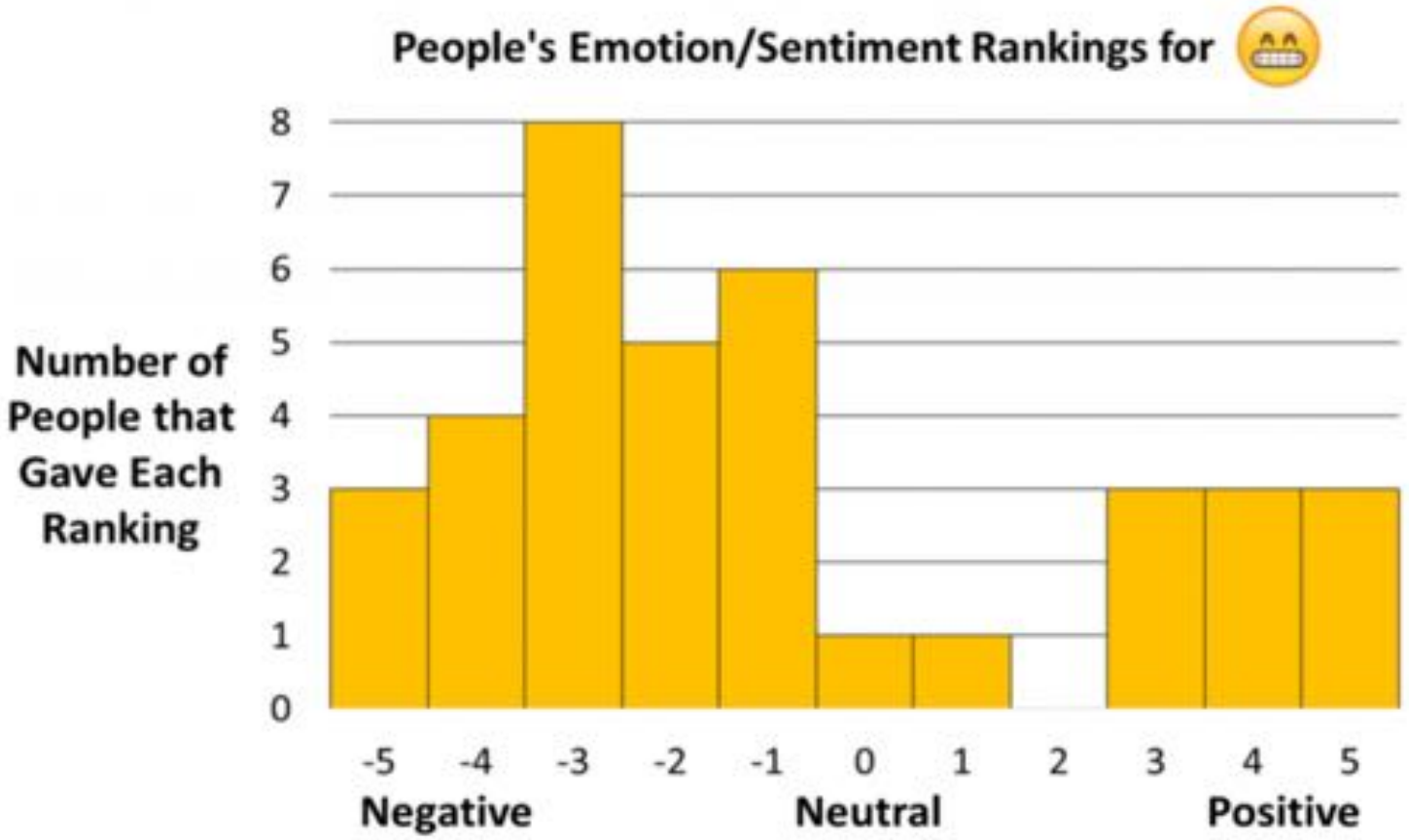
Code	Brow.	Chart	Apple	Goog[d]	Twtr.	One	FBM	Wind.	Sams.	GMail
U+1F600										
U+1F601										
U+1F602										
U+1F923			—				—			—
U+1F603										

Font effects

Dancer

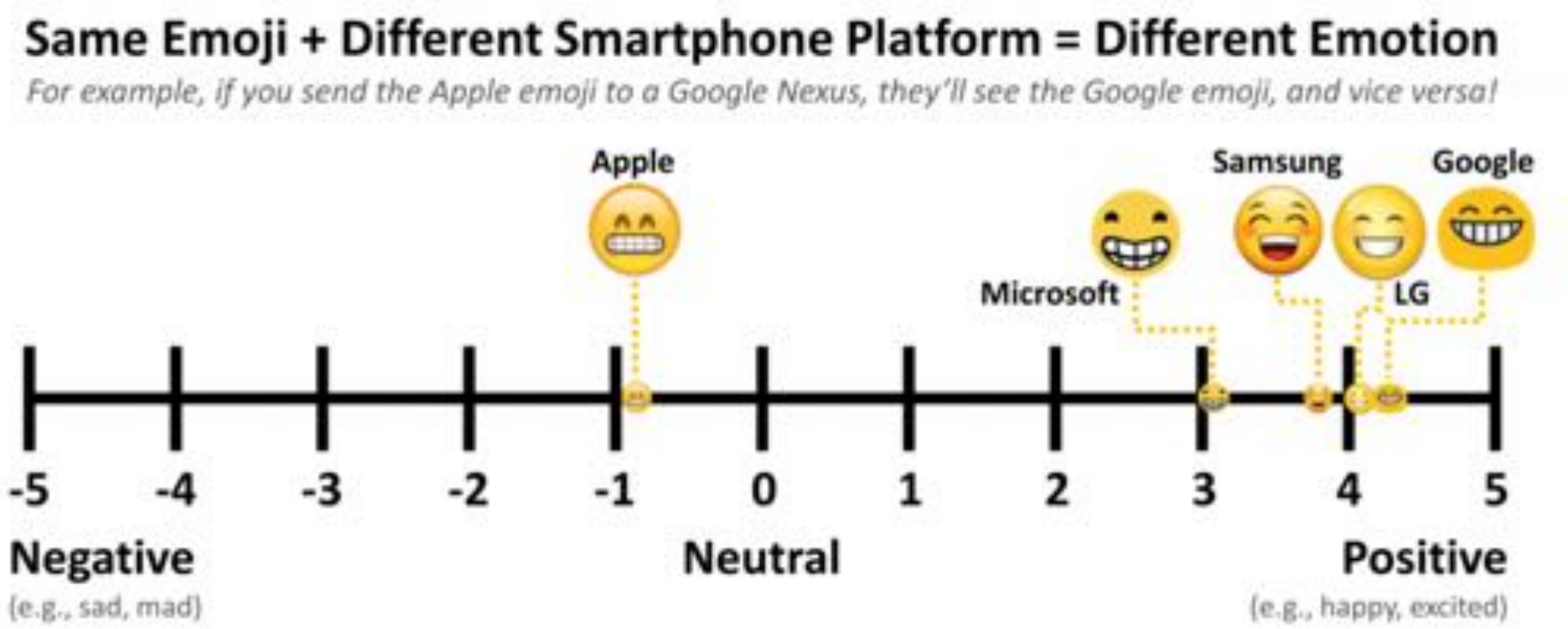
Differences can affect emotional readings

http://grouplens.org/blog/investigating-the-potential-for-miscommunication-using-emoji/



Differences can affect emotional readings

http://grouplens.org/blog/investigating-the-potential-for-miscommunication-using-emoji/



Misinterpretation caused by emoji fonts

http://grouplens.org/blog/investigating-the-potential-for-miscommunication-using-emoji/

Abby using a Google Nexus, texting Bill:

Bill using an iPhone, texting Abby:

Sometimes fonts change

Apple, old and new

Microsoft went the other way

Research on emojis - sociology

- Sociological research is interested in how people with different demographic profiles (age, gender) use emoticons and emojis, how it affects people's relationships and how it fits the cultural context.
- It has been found that, consistent with the stereotype of the emotional female and less emotional male, emoticons and emojis are used more by females (especially young females) than males
- And their usage is important for managing relationships (Huffaker & Calvert, 2005;

Sugiyama, 2015; Wolf, 2000; Kelly & Watts, 2015).

Research on emojis – computer science/NLP

- Research in computer science has primarily focused on using emoticons and emojis as a cue for automatically analysing the sentiment of short messages, commonly tweets, driven partly by the ease of obtaining data and commercial
- Positive emojis are used more frequently than negative ones (Novak et al., 2015).
- The polarity of emoticons and emojis is relatively well correlated with the perceived emotional polarity of the entire text,
  - but is poorly correlated with the perceived emotional polarity of the accompanying linguistic text alone (Boia et al., 2013).
- Using emoticons and emojis as a cue for sentiment analysis of tweets improves the accuracy compared to using the linguistic text alone, to a level between 60% to 75% (Hu et al., 2013; Zhao et al., 2012).
- Emojis tend to be a better indicator for an overall negative tweet than a positive one.
- Although the polarity of emojis frequently mismatch the polarity of the linguistic text or even the entire message, little has been done to analyze the nature of these mismatches.

Giving emojis sentiment scores

Novak, P. K., Smailović, J., Sluban, B., & Mozetič, I. (2015). Sentiment of emojis. *PloS one*, *10*(12), e0144296.

- Authors collected 1.6 million tweets across 13 European languages
- Approximately 4% of the tweets contained emoji
- 83 annotators gave ratings of positive, neutral or genitive : {1, 0, -1}
- 751 emoji were used more than 5 times and given a score
- The resulting emoji score ranged between -0.6 and 0.9 with median 0.3

Research on emojis – linguistics

- The assumption in current research in NLP is that emojis express the speaker's emotional state, and they can be seen as an independent channel of communication from that of the linguistic text.
- This assumption has been challenged by some studies in linguistics.
- Baron (2009) points out that just like linguistic words, the meaning of emoticons and emojis is often underspecified; they can mean different things in different context.
- Dresner & Herring (2010) argues that emoticons often not iconic of actual facial expressions and gestures. The meaning of emoticons is pragmatic: they are indicators of the illocutionary force of the textual utterance that they accompany.
  - They "neither contribute to the propositional content (the locution) of the language used, nor are they just an extralinguistic communication channel indicating emotion" (Dresner & Herring, 2010)[pp.255].

Questions

- At a social-level, the use of emoji is shaped by cultural and demographic factors.
- Computationally, emojis are an effective but far-from-perfect cues for the senti-
- ment of messages.
- We take the linguistic perspective that emojis are not simply signs of emotion.
- Emojis and linguistic text can interact with each other.
- The interactive relationship between the two channels can explain why the polarity of emojis and text can deviate from each other.

Questions

- (Dresner & Herring, 2010) argues that because very often emojis are not used to express emotional states, they do not map onto conventional facial expressions.
- One example is the use of the smiley face as a means to "downgrade a complaint to a simple assertion".
- The argument is that the producer is not actually happy uttering this, and that she would not actually smile when inserting a smiley face in text.
- This leads to the implication that emojis and natural non-verbal acts have different functions in communication.
- However, we know what smiles and laughters in dialogues often are not used to communicate the speaker's positive emotions, but to induce an affiliative attitude (for politeness) (Mazzocconi, Tian and Ginzburg, 2016).

Questions

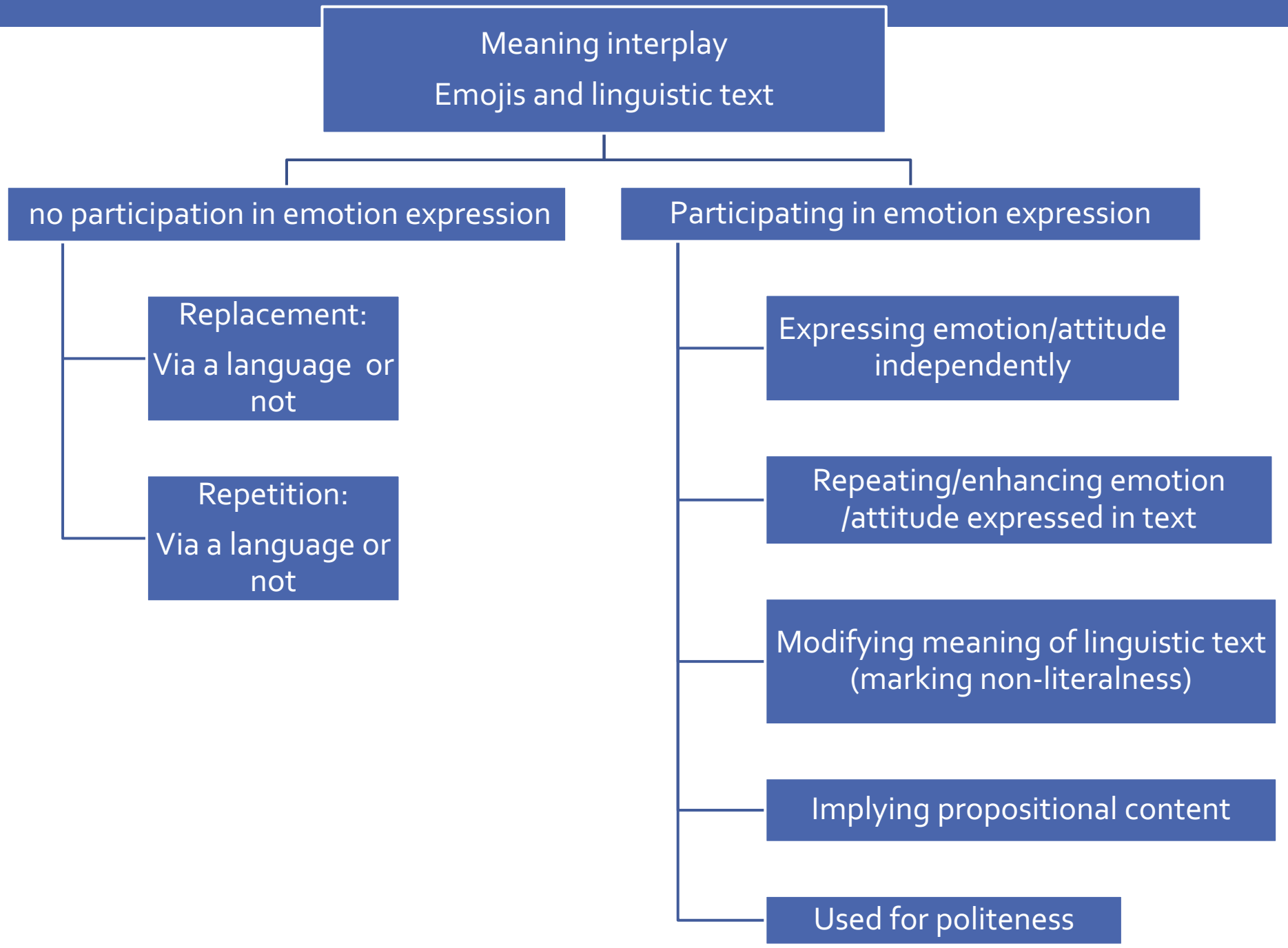
- Dresner & Herring (2010) proposes that emojis do not contribute to propositional content.
- However, there are cases where emoticons alone being used to answer a question or to acknowledge an assertion. In these cases, emoticons clearly imply propositional content.

Our proposal

- We argue that emojis and the linguistic text can modify the meaning of each other. The overall communicated meaning is not a simple sum of the two channels.
- We hypothesize that facial emojis and natural acts such as smiling and laughter have similar functions, though the distribution and frequency may be different.
- For example:
  - Could you please not do that hehe. / Could you please not do that ☺
  - Could you please not do that? Hahahahahaha/ Could you please not do that

Our proposal

- Emojis can be being iconic of a natural referent (e.g. laughter, fireworks) or being iconic of a conventional sign (e.g. winking face, ok hand gesture).
- When interacting with text, they can contribute to meaning in different ways:


Meaning interplay
Emojis and linguistic text
no participation in emotion expression
Participating in emotion expression
Replacement:
Via a language or not
Repetition:
Via a language or not
Expressing emotion/attitude independently
Repeating/enhancing emotion /attitude expressed in text
Modifying meaning of linguistic text (marking non-literalness)
Implying propositional content
Used for politeness

- Semantically compositional with linguistic text:
  - Replacement:

    I want to have a I am drinking a .

    Emoji of a tank to mean "thanks" (mediated via English)

  - Repetition:

- Semantically uncompositional with linguistic text:
  - Expressing emotions or attitudes "independently"

(Facebook update from survivor of the Florida gay club shooting 2016-06-12): I am safely home and hoping everyone gets home safely as well .

  - Repeating/enhancing emotions or attitudes expressed in text

This would probably be really good .

  - Modifying meaning of linguistic text (marking non-literalness)

I bet you are enjoying your revision .

  - Implying propositional content

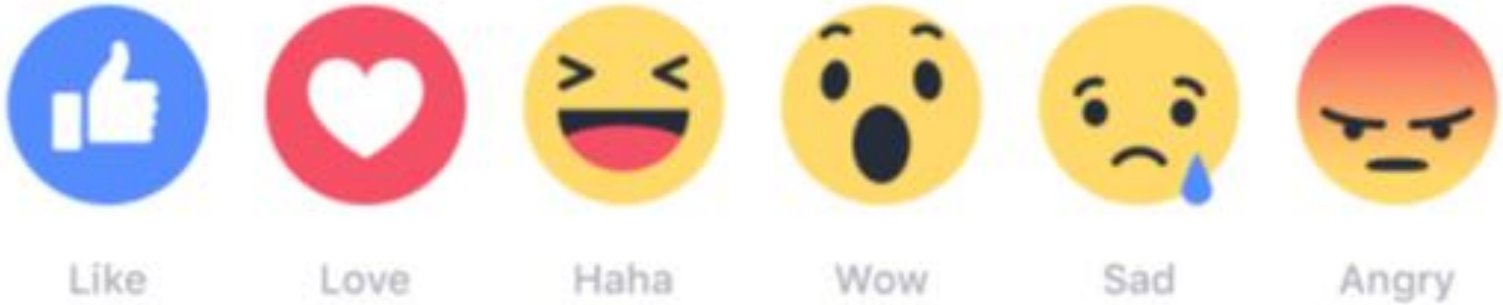
-: Do you fancy a coffee? –B .

  - Used for politeness

SOCIAL MEDIA

Emotions and emojis

Facebook reactions

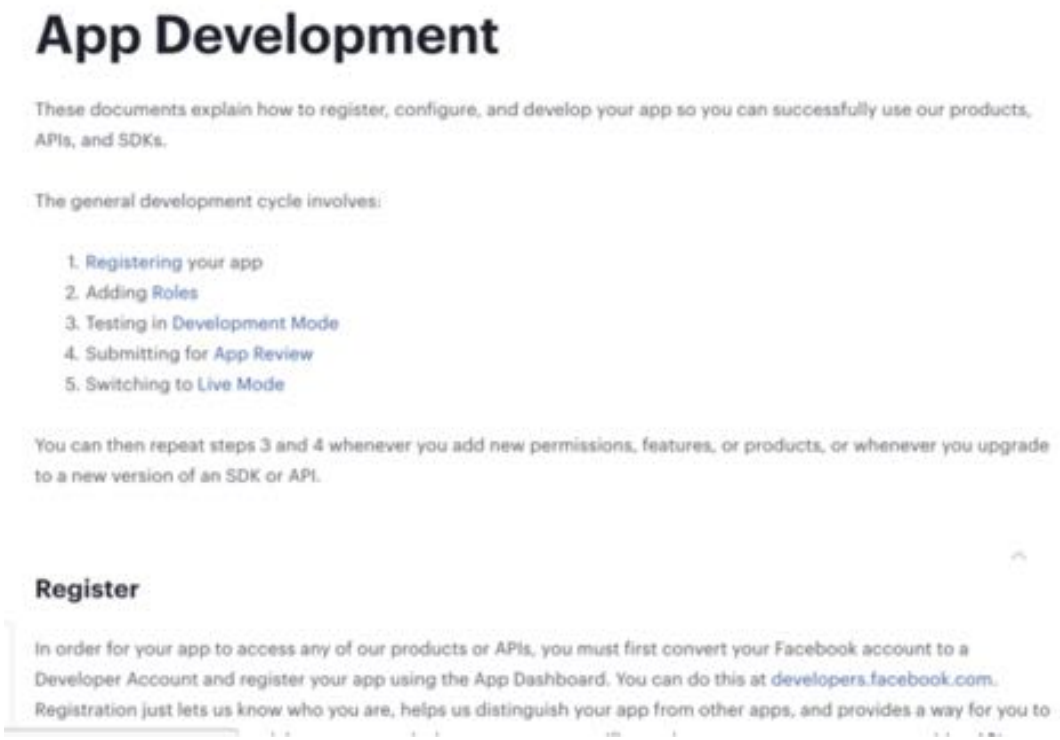


- Current study: treating facebook reactions to posts as the overall sentiment of the user. Then look at the emoji profiles for each reaction, to evaluate emoji sentiment calculation.
- Collected reactions data from 21,000 posts on media facebook pages (e.g. BBC, CNN, Le Monde) from four countries: UK, US, France and Germany.
- "Like" is the default reaction; it accounted for 80% of the 57 million reactions

Getting FB posts and reactions

Get facebook APP ID and secret

1. Login to Facebook
2. Go to https://developers.facebook.com/docs/apps/register
3. click 'Create Developer Account' button
4. Create a new Facebook App
5. Choose platform
6. Choose a name
7. Click on app and get the App ID
8. In the Dashboard can also get App Secret - requires reentry of your FB password



App Development

These documents explain how to register, configure, and develop your app so you can successfully use our products, APIs, and SDKs.

The general development cycle involves:

1. Registering your app
2. Adding Roles
3. Testing in Development Mode
4. Submitting for App Review
5. Switching to Live Mode

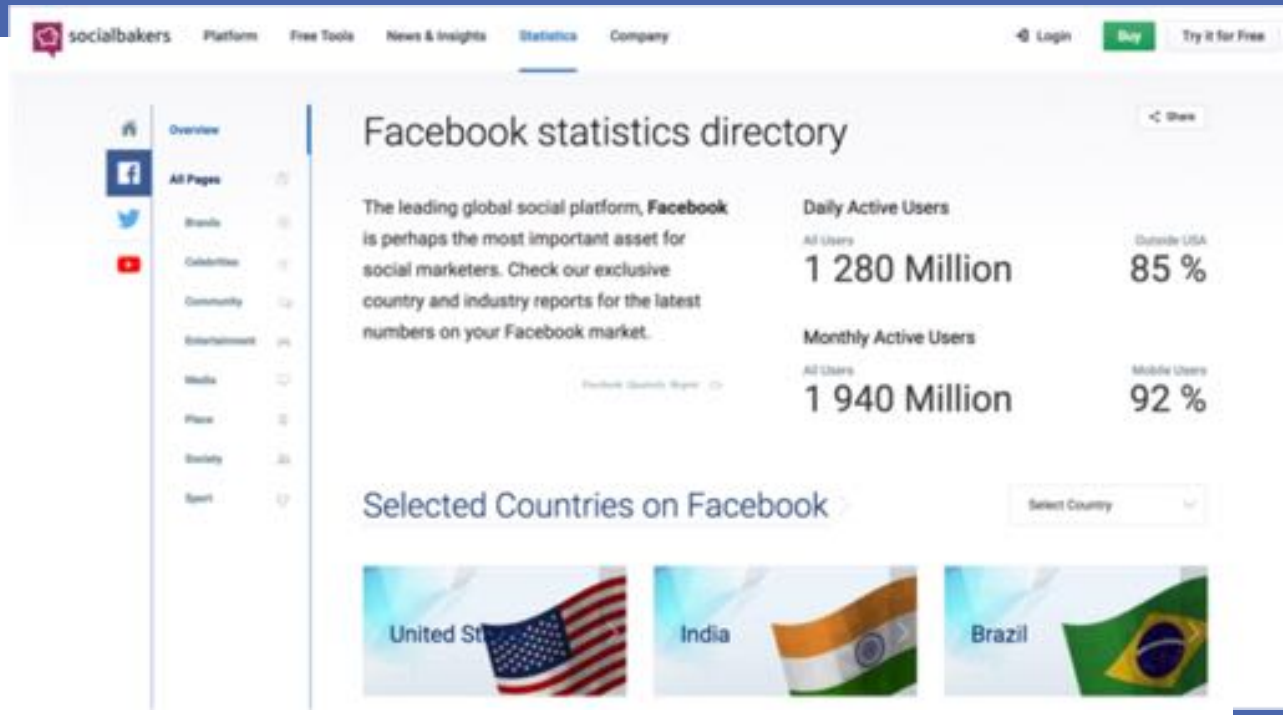
You can then repeat steps 3 and 4 whenever you add new permissions, features, or products, or whenever you upgrade to a new version of an SDK or API.

Register

In order for your app to access any of our products or APIs, you must first convert your Facebook account to a Developer Account and register your app using the App Dashboard. You can do this at developers.facebook.com. Registration just lets us know who you are, helps us distinguish your app from other apps, and provides a way for you to

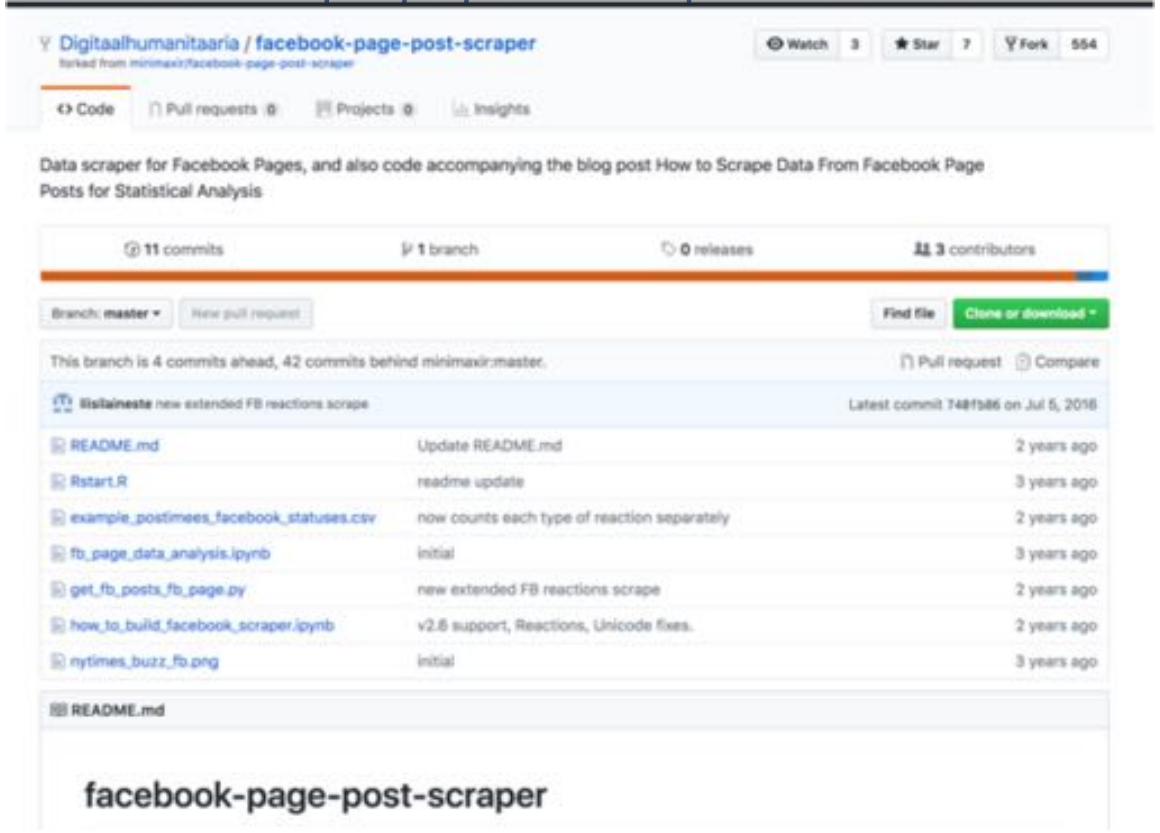
Getting FB public page ID

- Now find a public page where you want to crawl the data (helpful to check https://www.socialbakers.com/statistics/facebook/)
- And get the Facebook page ID:
- http://jsocialfeed.gardainformatica.it/facebook-page-id
- E.g. bbc news page URL is
- https://www.facebook.com/bbcnews/
- Paste it into "Facebook Page URL", to get FB page ID (something like 228735667216)

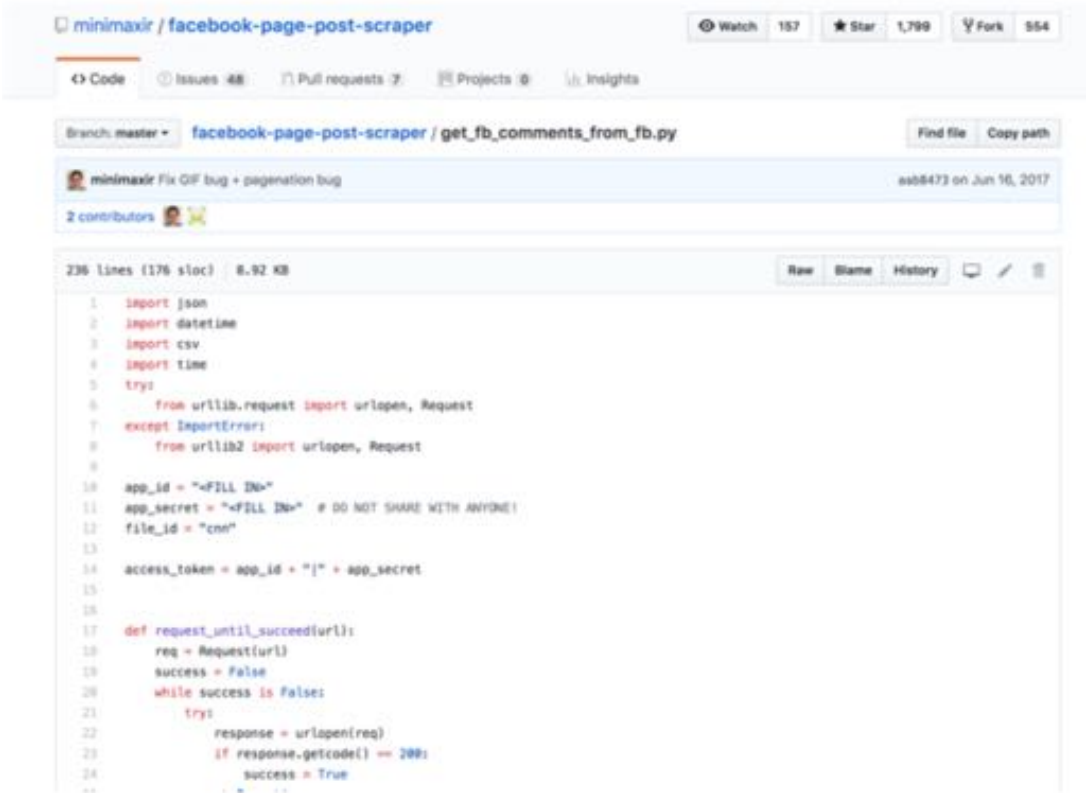


Getting FB posts, reactions and comments

- Find a scraper (or write your own!) e.g. https://github.com/Digitaalhumanitaaria/facebook-page-post-scraper

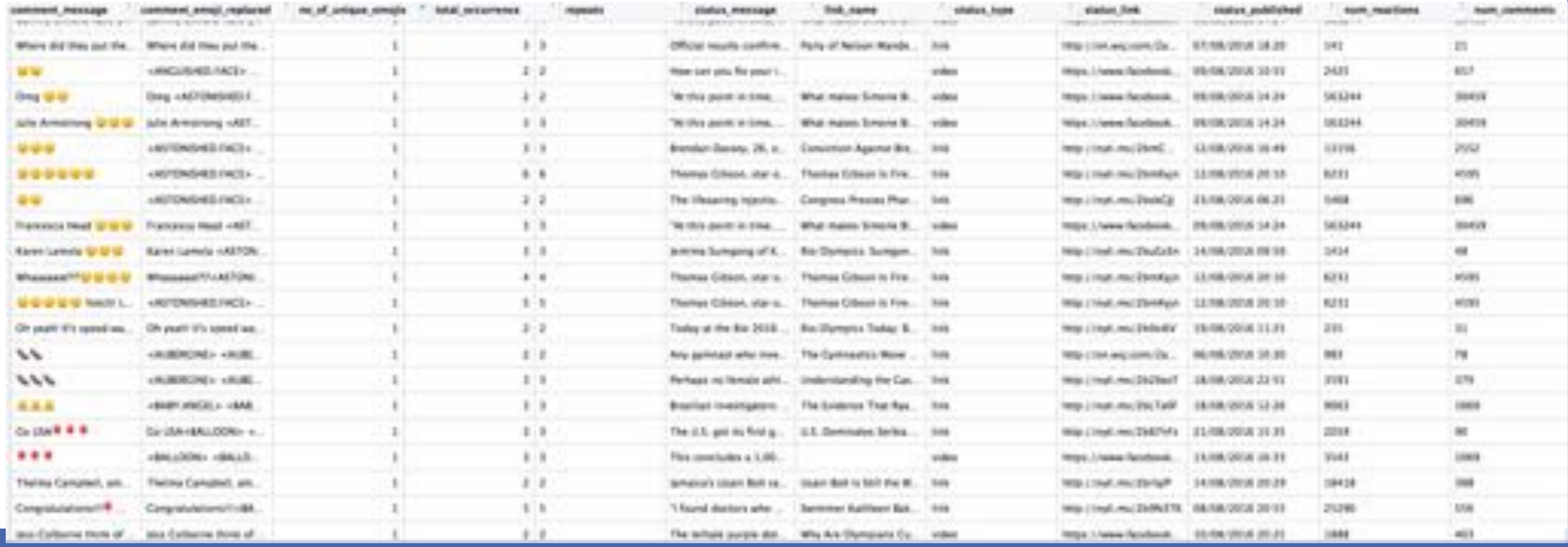


- https://github.com/minimaxir/facebook-page-post-scraper



Facebook comments

- Also collected 8 million comments to these posts.
- 6% of the comments contained emojis – higher rate than in tweets based on data from Novak et. al.



Le Monde
27 min ·

Une partie du pays est plongée dans un brouillard de particules fines et la situation va perdurer. Il s'agit du plus long et du plus intense pic de pollution hivernal depuis dix ans. Les mesures de fond, elles, se font

La France désarmée face à la pollution de l'air

LEMONDE.FR

175 27 commentaires 71 partages

J'aime Commenter Partager

[redacted] 😂😂😂😂

J'aime · Répondre · 17 · 4 h

[redacted] 😡

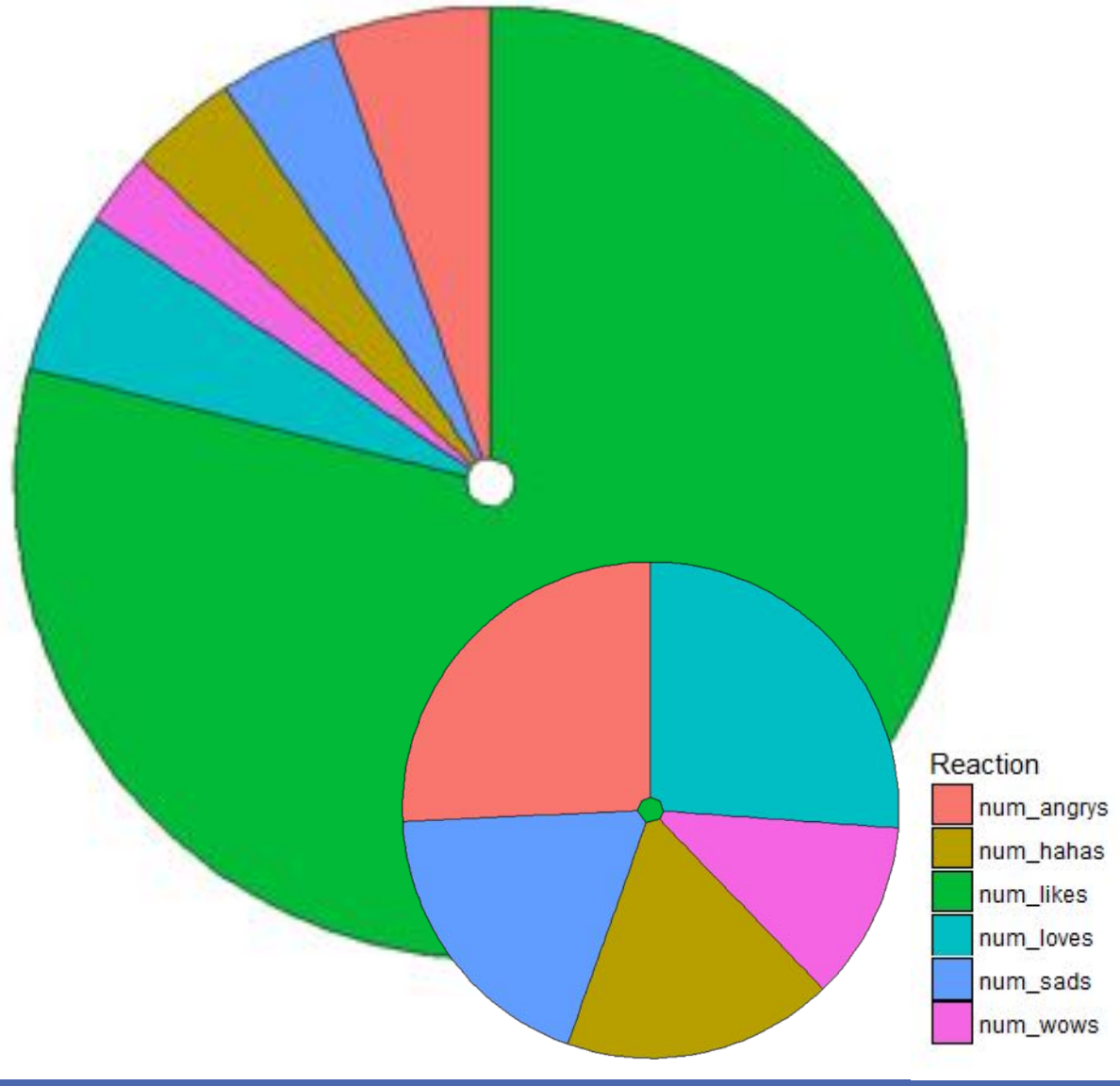
J'aime · Répondre · 7 · 8 h

[redacted] 😂😂😂😂😂😂😂😂 [redacted] 😮😮😮😮

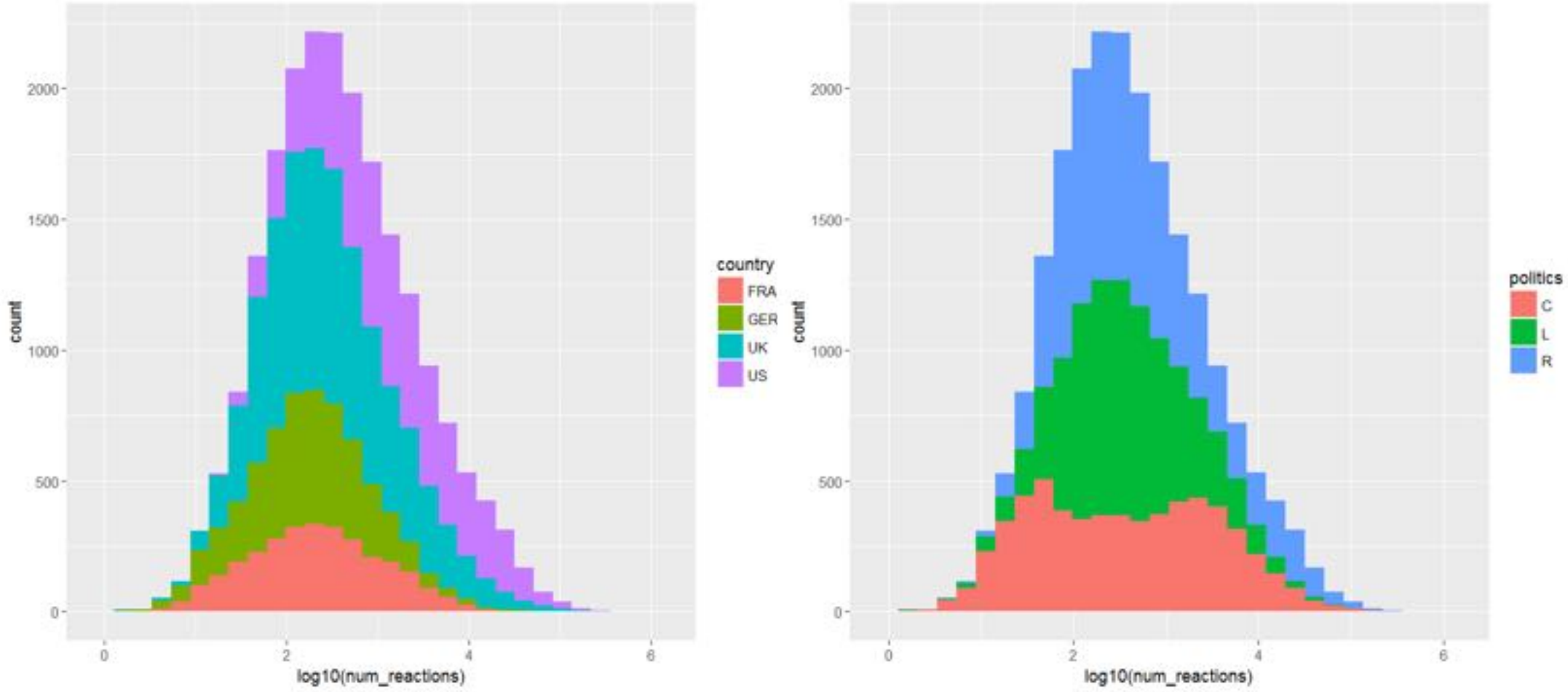
J'aime · Répondre · 1 · 7 h · Modifié

Results - Reactions

- Overall 57,444,404 reactions, 15,273,365 sharing actions, 8,463,602 comments, Likes >>> Loves > Angrys > Sads = Hahas >Wows
- Comments to reaction ratio: 0.15
- Share to reaction ratio: 0.27
- Slight but statistically significant difference in distributions by countries ($X^2(15) = 554810$, $p<2.2e-16$)
  - Angry: highest in France (9%), lowest in UK 3%)
  - Love: highest in US (6%), lowest in Germany (2%)
  - Haha: highest in Germany (6%), lowest in UK (3%)
  - No difference in Sads or Wows

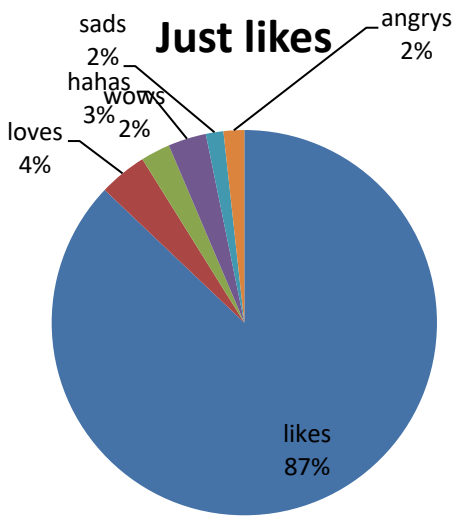


US most active, right-wing most active (caveat apply ☺)



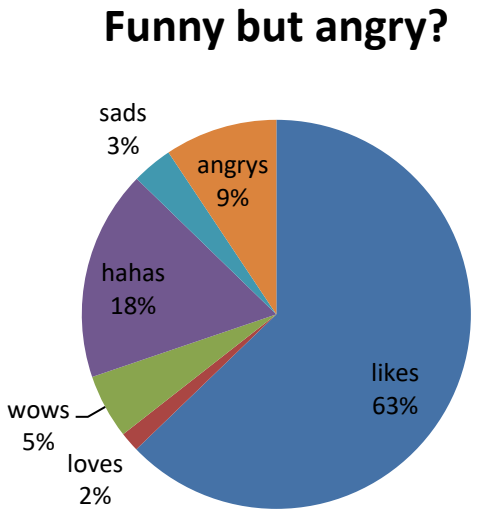
Results - Reactions

- K-means clustering gives four clusters profiles of reactions.
- People are most likely to share the post when reacted with "anger", and least likely to share with just "likes".
- Statistically significant differences in proportions across clusters ($X^2(15)=185$, $p<2.2e-16$)



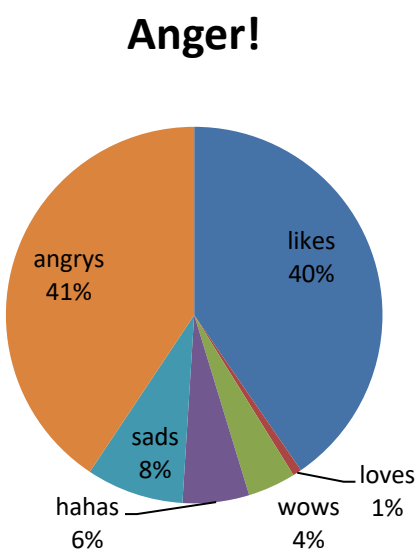
size: 4828

Share/ Reaction: 0.16



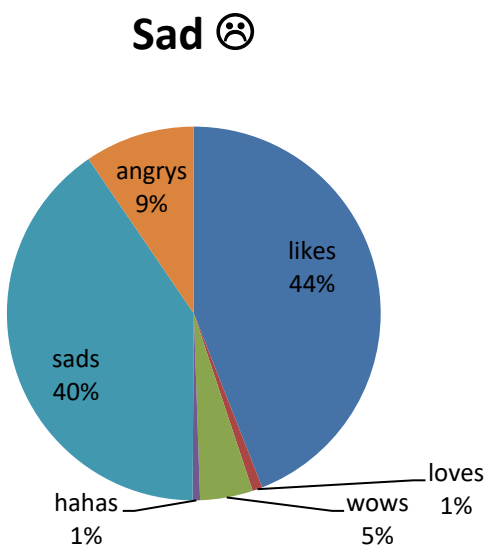
size: 2088

0.24



size: 943

0.33

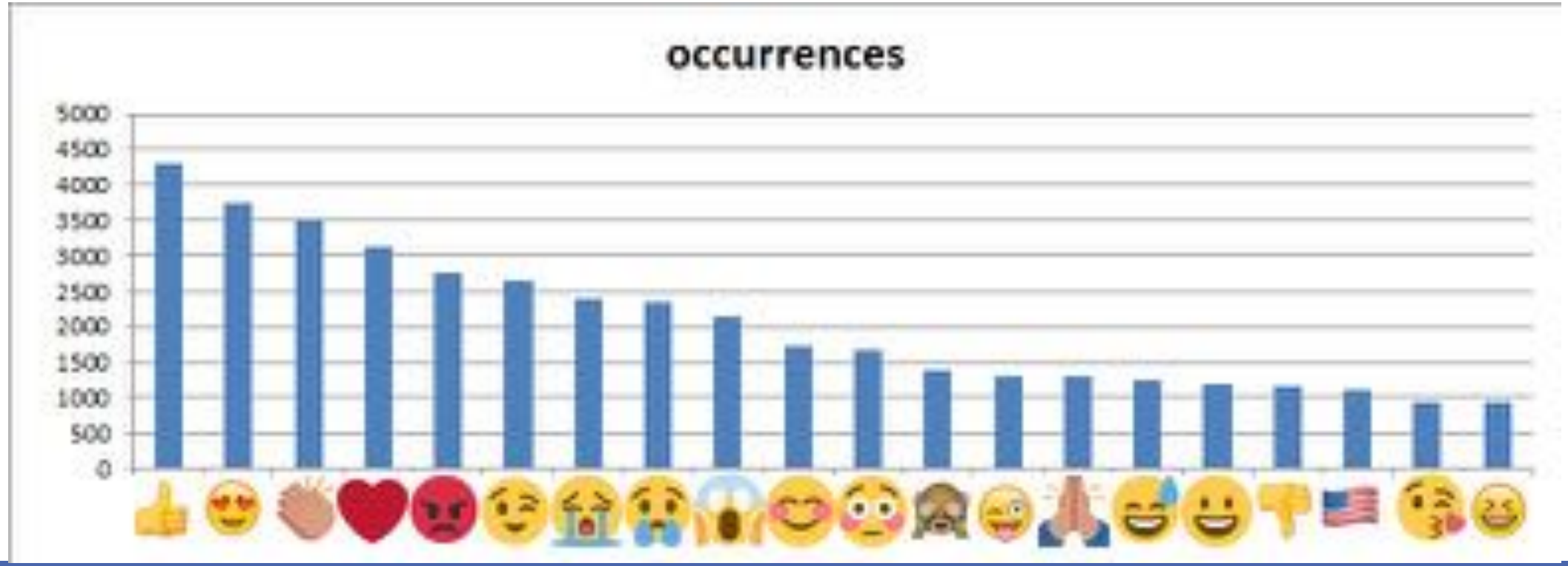


size: 658

0.24

Results - emoji

- We sampled 100,000 comments that contained emoji, and analysed distributions of emoji and their sentiment.
- Overall, the most frequent emoji were the following: the distribution does not have a Zipfian distribution, unlike words in natural language.



Emoji in comments to news posts *different* from general emoji use

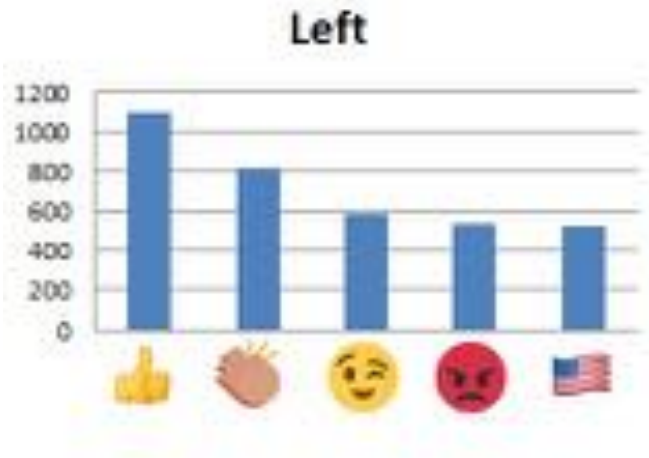
- Data from emojitracker.com, tracks twitter emoji. Laughing with tears No.1 by far.

Results – Emoji by country

- However, different countries use different emoji:

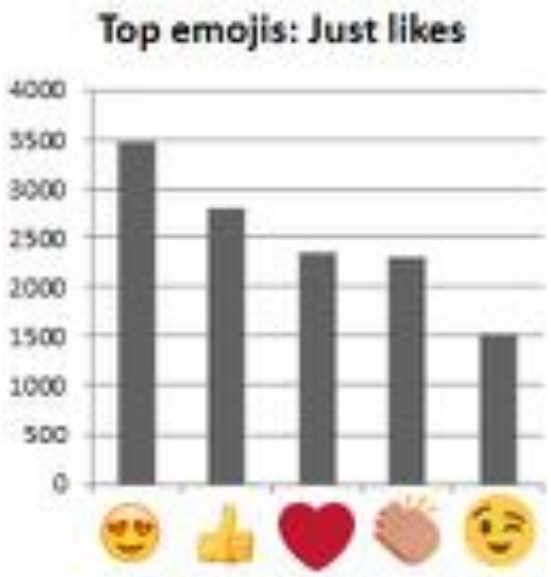
Results – Emoji by politics

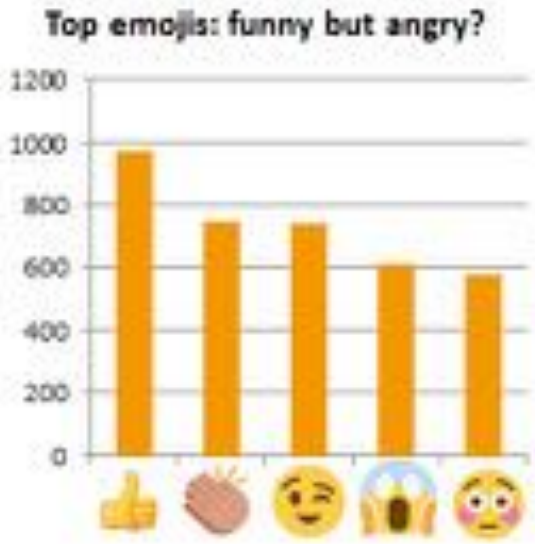
- The distribution of emoji is also different by political stance:

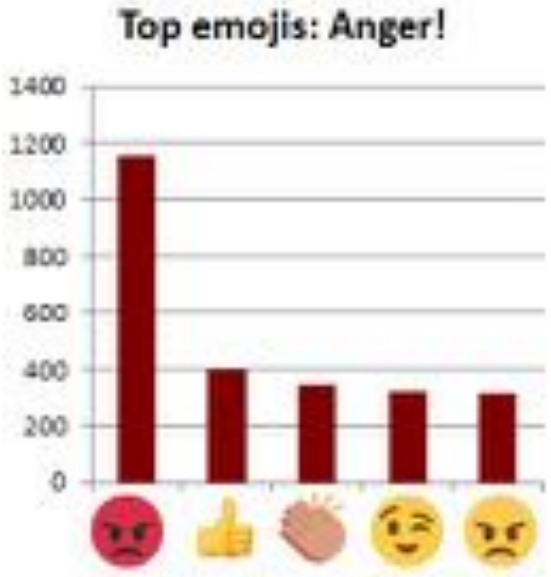


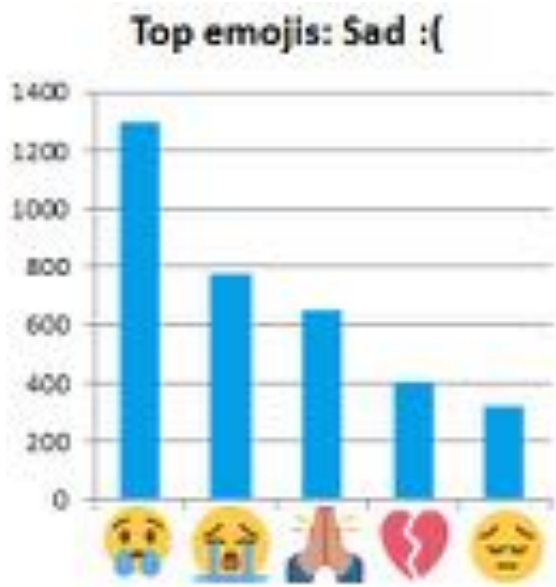
Results – Emoji vs. Reactions

- Are distributions of emoji different in different reaction profiles? Yes!





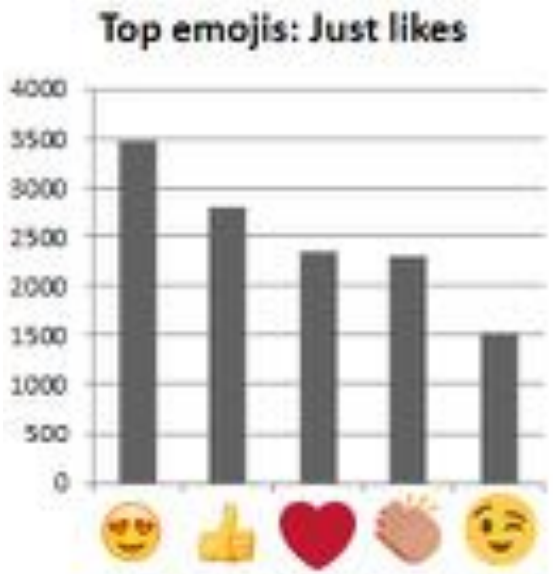


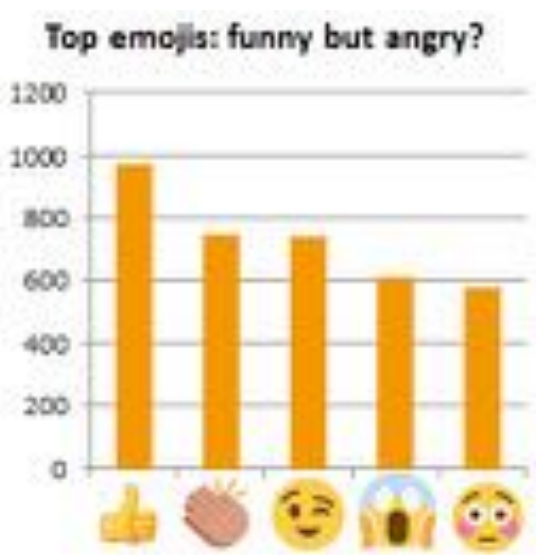


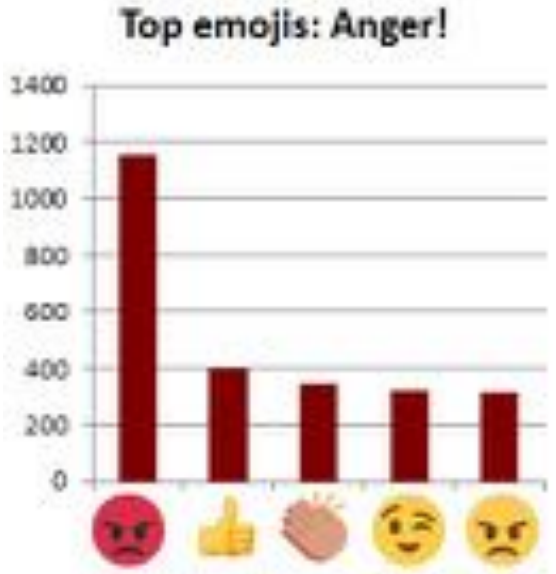
Results – emoji sentiment

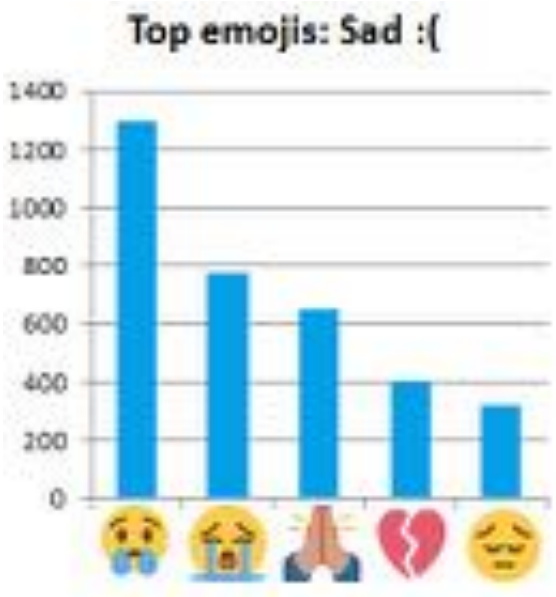
- Using the sentiment score complied for emoji by Novak et al. (2015), we calculated the average emoji-based sentiment score for each posts.
- In each comment containing emoji, the score is calculated as
- $\sum_1^n (Log(occurrences\ of\ emojii) + 1)\ * sentiment\ score\ of\ emojii$
- So that, for example, three hearts in one comments do not count to have three times the sentiment of comments containing only one heart.
- Then the average sentiment for a post is the mean of sentiment of comments (based on emoji) to this post.

Results: emoji sentiment vs. reactions









Sentiment Score:	**0.41**	**0.34**	**0.24**	**0.24**

Predicting Reactions Profile from Emojis

Featuers: emojis tf-idf, plus number of emojis and average emoji sentiment score.

Classifier: xgboost. Data: UK only (N_train = 3083, N_test = 771)

	precision	recall	f1-score	support
angry	0.61	0.32	0.42	60
funnyshock	0.37	0.16	0.22	180
like	0.68	0.90	0.77	480
sad	0.75	0.31	0.45	51
avg / total	0.62	0.64	0.61	771

Results: emoji sentiment vs. reactions

- We can see that the average emoji based sentiment score for cluster 3 (angry dominant) and cluster 4 (sad dominant) are lower than cluster 1 and 2.
- However, the difference is not pronounced, and the emoji based sentiment score for cluster 3 and 4 are still positive. Why?

Emoji and sentiment

Why is it that in posts with frequent angry reactions and sad reactions still have positive sentiment scores from comments emoji?

- Positive emoji still frequently used in comments relating to angry and sad reaction profiles.
- Positive emoji are sometimes used NOT to express positive emotion, but for politeness reasons. E.g. a smiley face can be used to soften a criticism/ disagreement
  - I don't think you are right ☺
- While negative emojis tend to indicate the global sentiment of the text, positive emoji can have a more local effect, e.g. recognizing something as ridiculous while overall feeling negative.
- Emojis may be used ironically.

Negative emojis in positive contexts

Emoji and sentiment

- Why is it that the average sentiment of profiles 1 is not much higher than profiles 3 and 4?
- Novack et al. (2015) scored sentiment using *entire texts containing emoji* rather than emoji by themselves.
- While this is a good approach to obtain the overall sentiment of texts containing emoji, it does not separate emotion expression versus politeness uses of positive emoji.
- Therefore, the Novack et al. (2015) sentiment score for, e.g. a smiley face, is likely lower than the perceived sentiment of a smiley face used purely to express emotions.

Conclusions:

- The current study studied Facebook reactions and emoji in comments to news pages in US, UK, France and Germany.
- Reactions: “like” most frequently (being default, plus the rest recently introduced). Slightly differences across countries and political stances
- However, people are more likely to share when the reaction is something other than “like” >> stronger emotional reaction leads to more sharing
- Four reaction profiles: “Just likes”, “Funny but angry?”, “Anger!”, “Sad ☹”. The first cluster is the most frequent.

Conclusions:

- Emoji: the most frequently used emoji in comments to news posts are DIFFERENT from general uses >> less personal conversations, more discussions.
- Emoji frequencies, unlike words in natural language, do not follow Zipf's law. >> the senses of emoji overlap more than that of words?
- Emoji distribution significantly different in different REACTION profiles. >> if we treat reaction as the overall sentiment, this suggest that emoji are good indicators of users' sentiment.
- However, sentiment score calculated based on Novak et al. (2015) showed less differentiation (low but still positive scores in "Anger!" and "Sad ☹" clusters, not much higher score in the other two).
  - We suggest this is to do with positive emoji sometimes used for politeness reasons or ironically, and the methods of Novak et al. do not address this issue.

References

- Naomi S Baron. 2009. The myth of impoverished signal: Dispelling the spoken-language fallacy for emoticons in online communication. In Jane Vincent and Leopoldina Fortunati, editors, Electronic emotions: The mediation of emotion via information and comunication technologies, pages 107–135. Peter Lang, Oxford.
- Marina Boia, Boi Faltings, Claudiu-Cristian Musat, and Pearl Pu. 2013. A:) is worth a thousand words: How people attach sentiment to emoticons and words in tweets. In Social Computing (SocialCom), 2013 International Conference on, pages 345–350. IEEE.
- Thomas Dimson. 2015. Emojineering: Machine learning for emoji trends by instagram. http://instagram-engineering.tumblr.com/post/117889701472/emojineering-part-1-machine-learning-for-emoji. Accessed: 2016-06-12.
- Eli Dresner and Susan C Herring. 2010. Functions of the nonverbal in cmc: Emoticons and illocutionary force. Communication theory, 20(3):249–268.
- Xia Hu, Jiliang Tang, Huiji Gao, and Huan Liu. 2013. Unsupervised sentiment analysis with emotional signals. In Proceedings of the 22nd international conference on World Wide Web, pages 607–618. International World Wide Web Conferences teering Committee.
- David A Huffaker and Sandra L Calvert. 2005. Gender, identity, and language use in teenage blogs. Journal of Computer-Mediated Communication, 10(2):00–00.
- Ryan Kelly and Leon Watts. 2015. Characterising the inventive appropriation of emoji as relationally meaningful in mediated close personal relationships. Experiences of Technology Appropriation: Unanticipated Users, Usage, Circumstances, and Design.

References

- Chiara Mazzocconi, Ye Tian, and Jonathan Ginzburg. 2016. Towards a multi-layered analysis of laughter. In Proceedings of JerSem, the 19th Workshop on the Semantics and Pragmatics of Dialogue.
- Petra Kralj Novak, Jasmina Smailovi´ c, Borut Sluban, and Igor Mozeti˘ c. 2015. Sentiment of emojis. PloS one, 10(12):e0144296.
- Satomi Sugiyama. 2015. Kawaii meiru and maroyaka neko: Mobile emoji for relationship maintenance and aesthetic expressions among japanese teens. First Monday, 20(10).
- Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. 2010. Sentiment strength detection in short informal text. Journal of the American Society for Information Science and Technology, 61(12):2544–2558.
- AleciaWolf. 2000. Emotionalexpressiononline: Gender differences in emoticon use. CyberPsychology & Behavior, (5):827–833.
- Jichang Zhao, Li Dong, Junjie Wu, and Ke Xu. 2012. Moodlens: an emoticon-based sentiment analysis system for chinese tweets. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, pages 1528–1531. ACM.