

Polarisation Analysis in case of Johnny Depp & Amber Heard



SICSS Aachen-Graz 2022 Project

Agenda

- Our Project Team
- Research Questions, Hypotheses & Motivation
- Data Collection
- Methods
- Results
- Limitations
- Learnings
- Next Steps

Our Project Team



Alina Kopkow



Dovlat Aliyeva



Katharina Roetzer



Muhsin Ciftci



**Nathalie Van
Raemdonck**



Yasaman Asgari



Zehui Yu

Research Questions & Motivation

How did the reactions - in terms of tweets, retweets, quotes - from “Team Depp” and “Team Heard” develop over the course of the 2022 Fairfax/Virginia trial?

- > How did support for Johnny Depp (JD) and Amber Heard (AH) develop over the course of the trial?
- > How did hostilities towards JD and AH develop over the course of the trial?

Hypotheses

- > Support for Johnny Depp **increased** over the course of the trial
- > Support for Amber Heard **decreased** over the course of the trial
- > Hostilities towards Johnny Depp **decreased** over the course of the trial
- > Hostilities towards Amber Heard **increased** over the course of the trial



Motivation



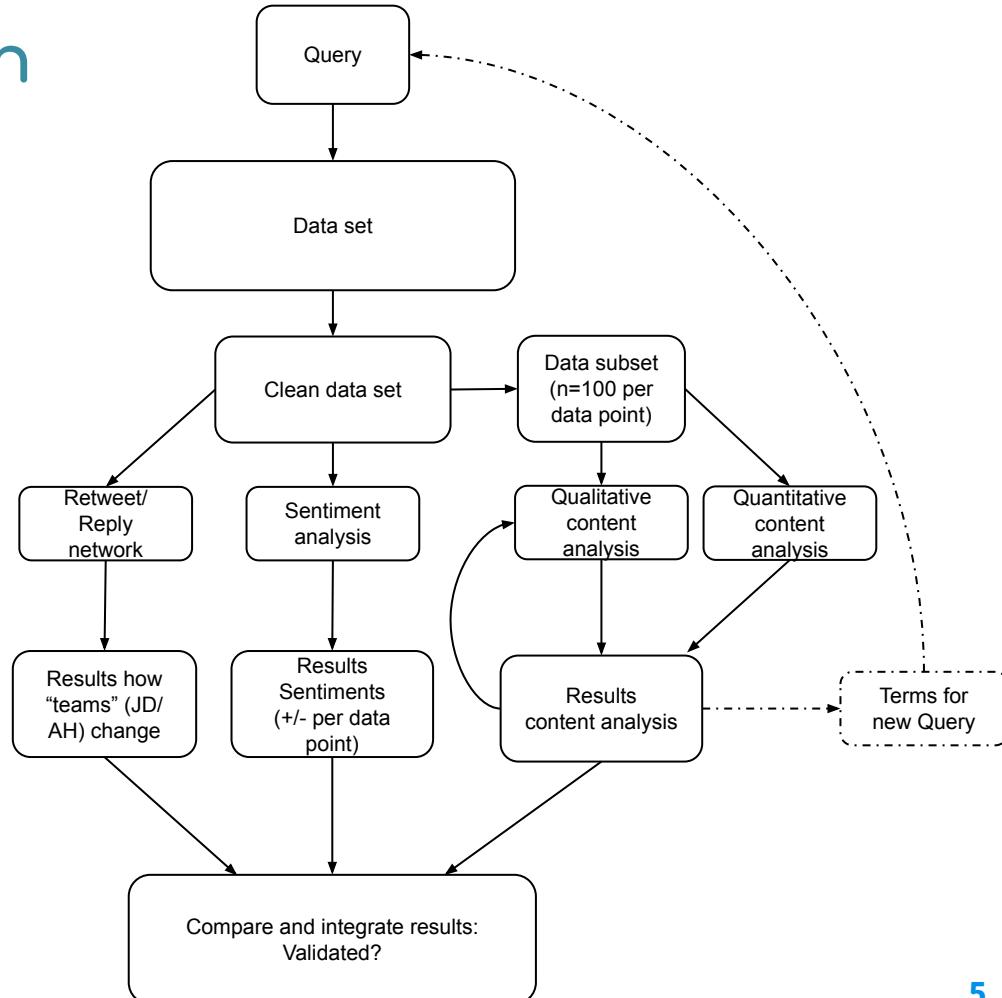
- > Find out more about the discourse on Twitter regarding the Depp v Heard case
- > Try out new methods, models, approaches, and integrate them (increased validity through triangulation?)
- > Utilise every team member’s skills, create and share ‘good’ models, codes, notebooks,... for future use
- > Consider everyone’s interests and learning goals
- > Learn from and with each other

Multi-Method Approach

Let's test our hypotheses ...

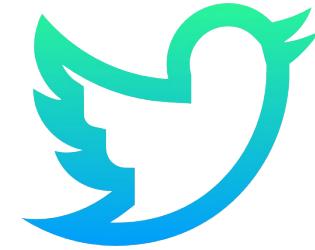
Based on our team's...

- Research interests
- Pre-existing skills
- Interest in learning new skills
- Background



Data Collection with Twitter API

- Data points t0 - t3
- Tweets containing “Amber Heard”, “Johnny Depp”
- Context ID
- Hashtags



“Team AH” - Supporting AH

#IstandWithAmberHeard
#AmberHeardIsASurvivor
#HumanizeHer
#IbelieveAmberHeard
#WEHEARDYOUAMBER
#AmberHeardDeservesJustice
#AmberHeardIsNotAlone

“Team AH”- Hostile towards JD

#DARVODepp
#JohnnyDeppIsAWifeBeater

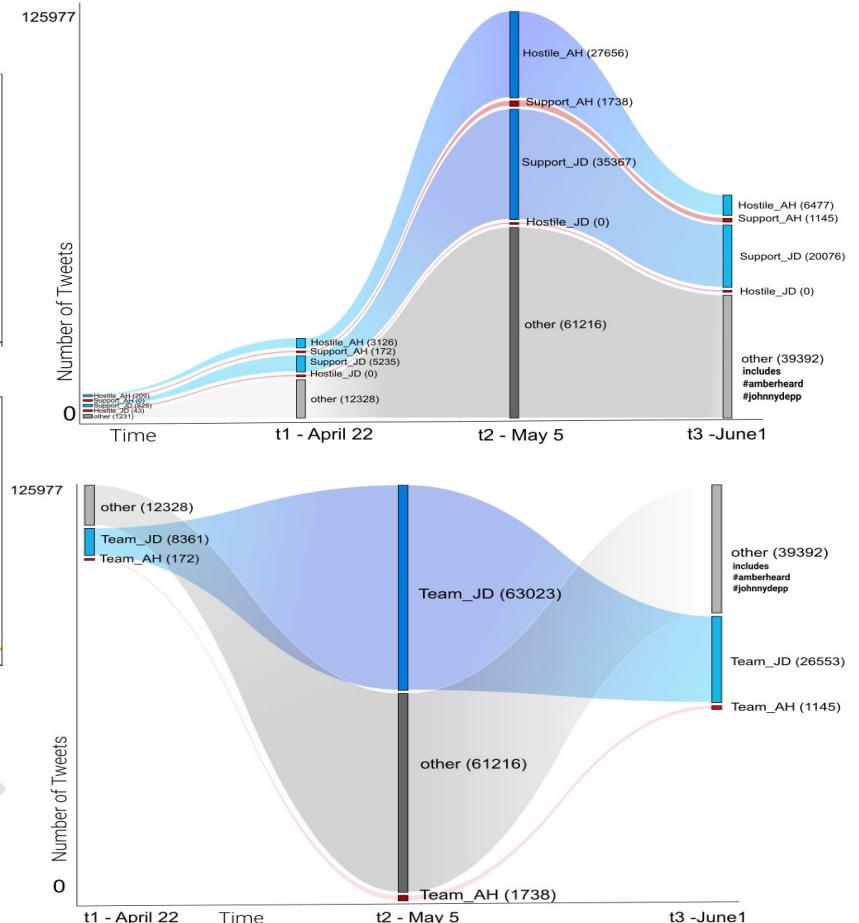
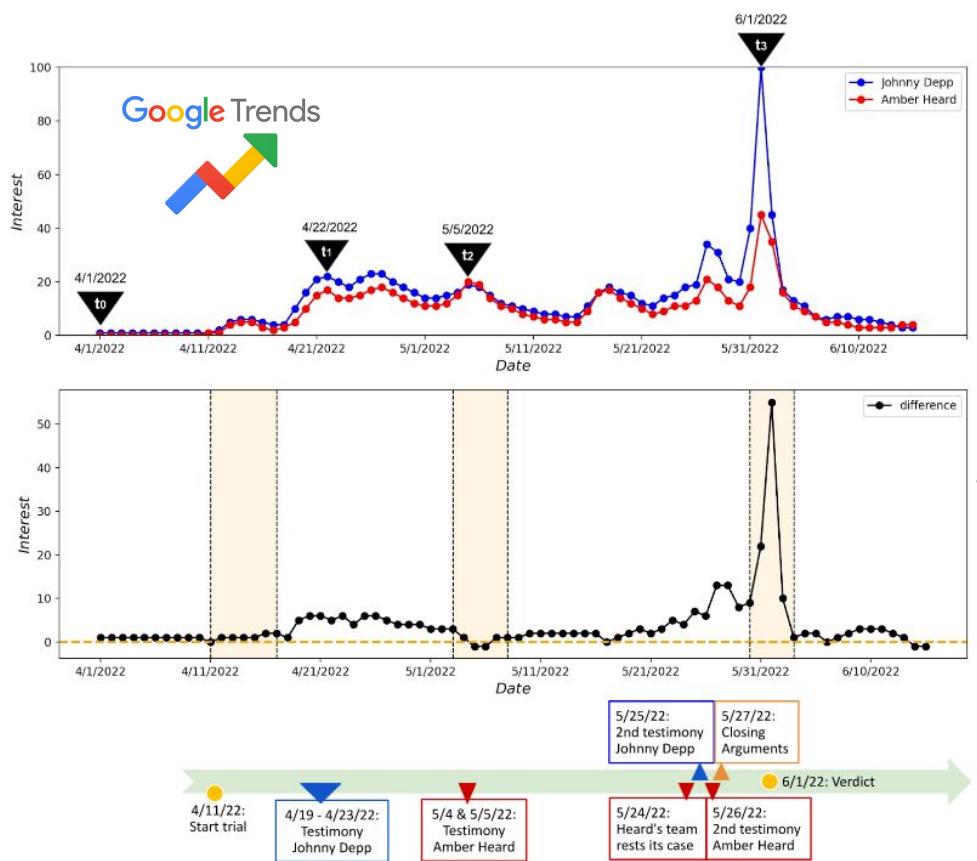
“Team JD” - Supporting JD

#JusticeForJohnnyDepp
#JusticeforJohnny
#JohnnyDeppWon
#TeamJohnnyDepp
#JohnnyDeppIsASurvivor

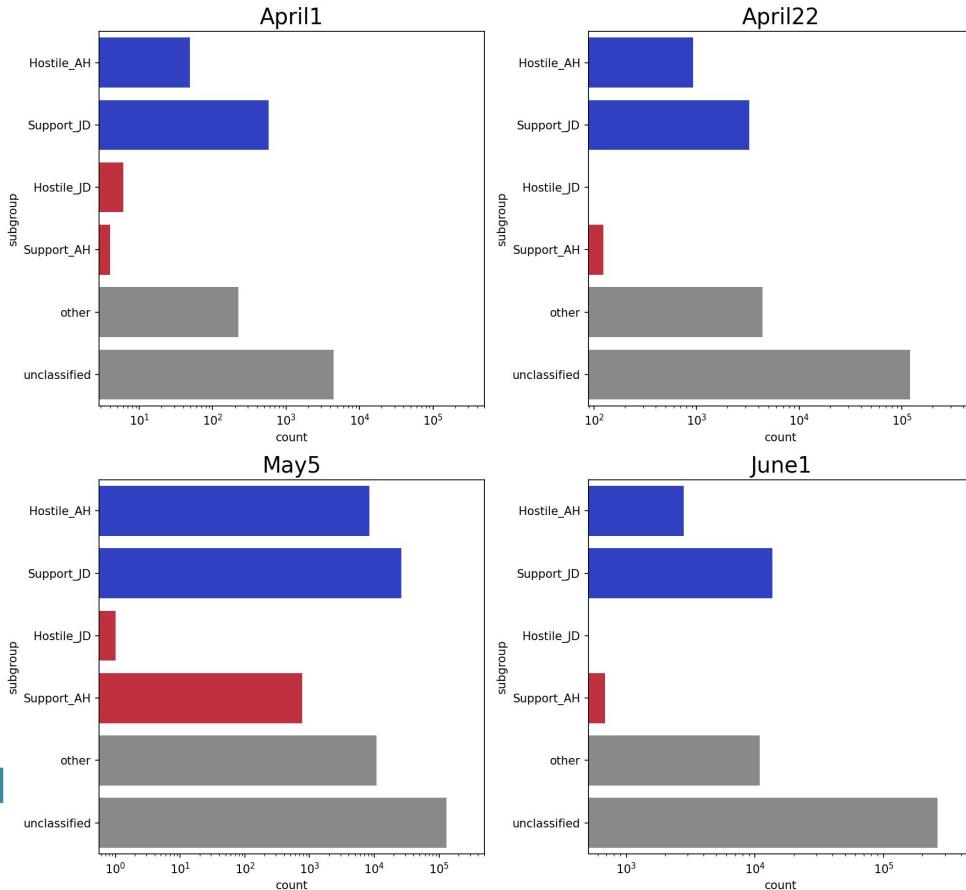
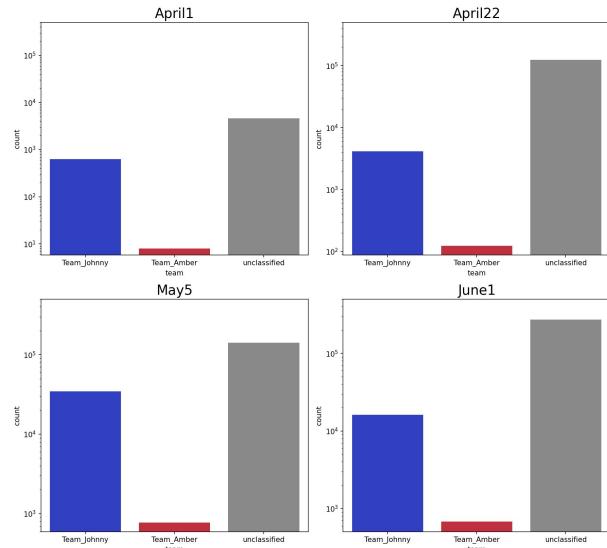
“Team JD” - Hostile towards AH

#AmberHeardIsALiar
#AmberHeardIsAnAbuser
#AmberHeardIsAHusbandBeater
#DARVOHeard
#amberturd

Data Collection with Twitter API



General Results: Subgroup Tweet Count



General Results: Most Commonly Used Words

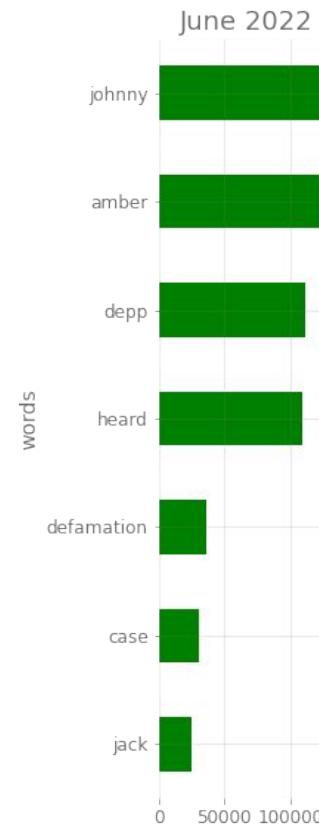
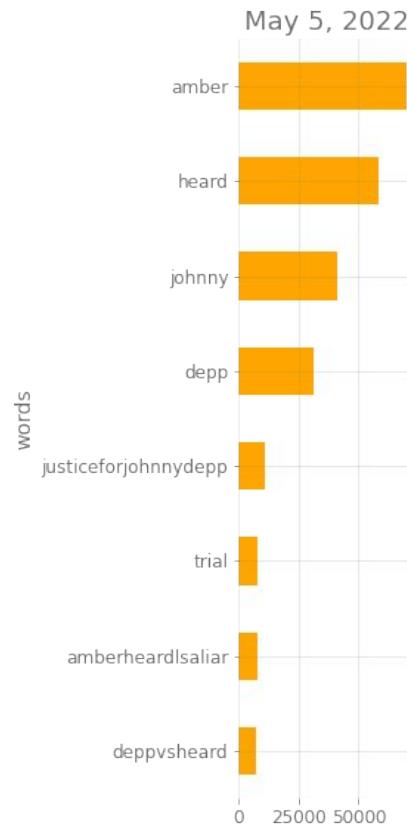
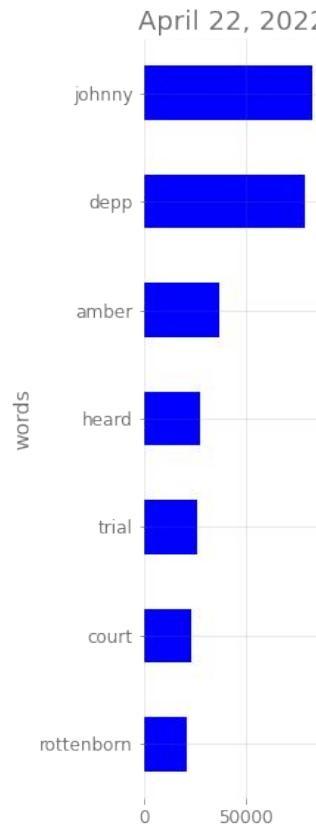
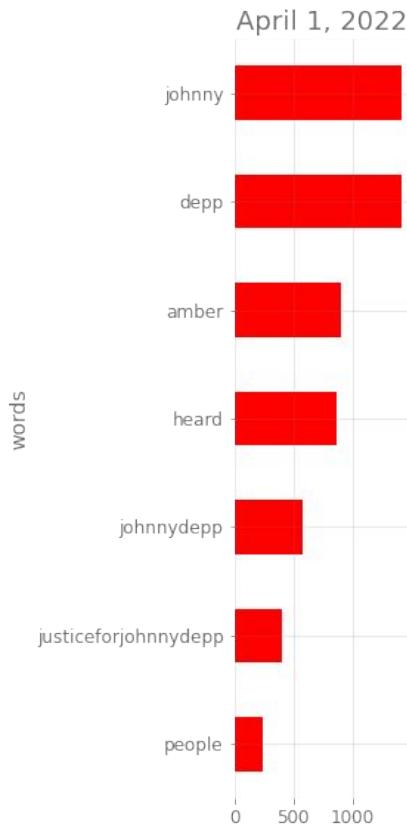


Illustration of Most Used Words

Team: Team_Johnny times getting
justiceforjohnnydepp nevertruthruth
censor any nevertruthruth truthwins
jd truth support jd
justiceforjohnnydepp johnnydepp
johnnydepp justiceforjohnnydepp

truth justiceforjohnnydepp
scam hoax

johnny depp
hoax perjury
ang closer

truthwins scam
johnnydepp justiceforjohnnydeppvirgin

days justiceforjohnnydeppvirginia2022
justiceforjohnnydeppvirginia2022 justiceforjohnnydepp

getting closer

depp justiceforjohnnydepp Jerry support

April 1
Team: Team_Amber

erratic end supports
dont johnnydeppisawifebeater
via boycottwarner justiceforamberheard
johnnydeppisarapist peo
listen misogyny
abusers boycott
unconditionally boycott
boycott disney
boycott fantasticebeasts

April22
Team: Team_Amber
amberheard birthday
sickening much
istandwithamberheard
amp justiceforamberheard
women istandwithamberheard
birthday girl
much johnnydepp
best amberheard
justiceforamberheard istandwithamberheard
hates women johnnydepp hates
thanks amp happy birthday
jack thanks
istandwithamberheard wearithyouramberheard
girl lily amberheard lily
lt3 happybirthdayamberheard best



Ghislaine Maxwell + Jeffrey Epstein Court

Team: unclassified

say much amp heard
love entertained rainey

depp abuser words
nobody

tell heart amber miss

heart
via believ mortbottom



Team: Team_Johnny
eally much sitting needs
carpet poor crying amberheardlsalair
ture tears amber dirty pity
hnydeppisimocent
sticeforjohnnydepp
media amber
berheard dirty
ch deppsheard
johnnydeppsvamberheard

Team: Team Johnny
team busch's sponsor
johnnydepp dream
deppin amber
heardiing believe
mepoo justiceforjohnnydepp
johnnydeppysamberheardtrial
clear istandwithjohnnydepp
justiceforjohnnydepp
amberhearddeservesprison
amberheardgeneration

Team: unclassified

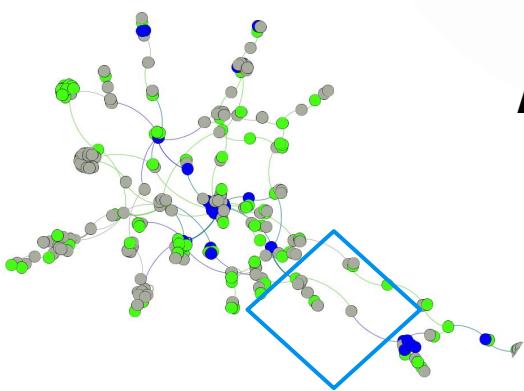
amber heard is a psychopath
heard johnny scratches
johnny scratches
honeymoon
amber pau back
mixx chin unseen nose appears
photo multiple stand

June1
am: Team_Amber
johnnydepp part
lot us depp ended
ex us believe
en
dollar **realizes lot**
rheard realizes **ex think**
too **teamamberheard**
ve johnnydepp
thamberheard justiceforanheard
believeinwopen

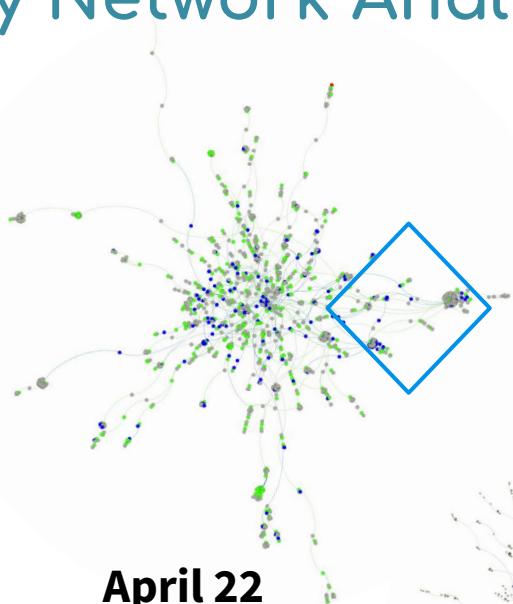
Team: unclassified

defamation
sued domestic case
johnny exwife actions
yes fact uk got
depp amber claims away
abuse

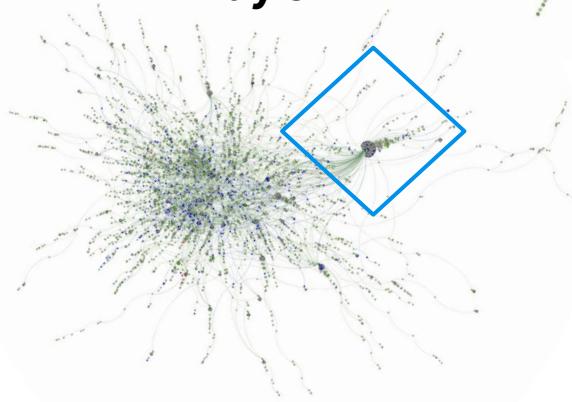
Results Reply Network Analysis



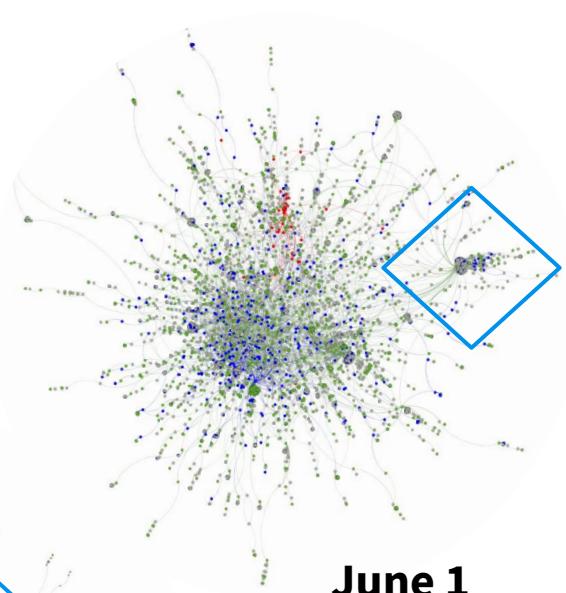
April 1



April 22



May 5

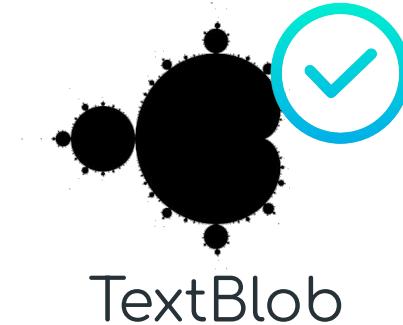
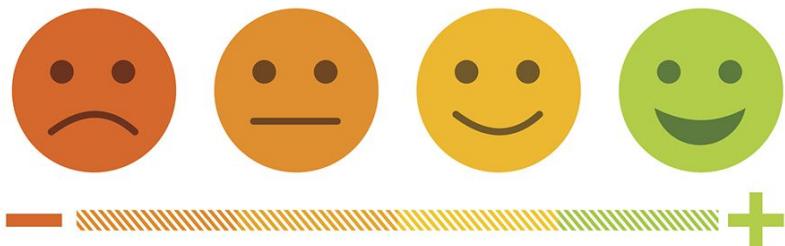


June 1

Results Retweet Network Analysis

Time	Edges	Nodes	Modularity
April 1	289	259	0.794
April 22	1662	1460	0.866
May 5	9200	7890	0.849
June 1	6485	5121	0.794

Results Sentiment Analysis



TextBlob

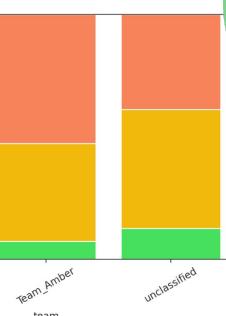
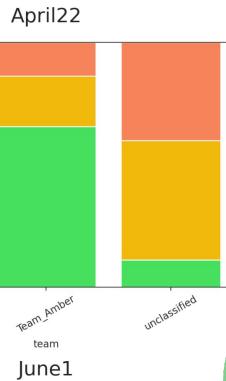
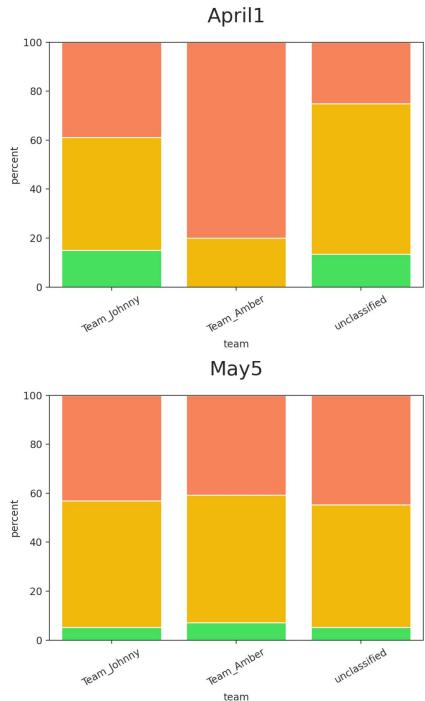
Time	Tweets	Non-RT	Coverage
April 1	5 261	1 700	32.3%
April 22	128 557	21 216	16.5%
May 5	177 114	58 277	32.9%
June 1	289 433	63 029	21.8%



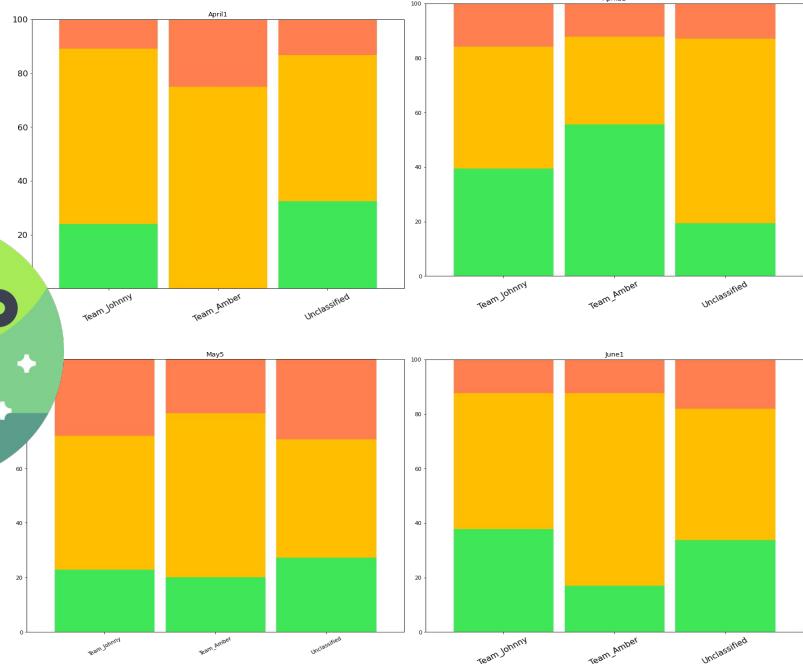
Roberta

Comparing Results Sentiment Analysis

Roberta (Excluding Retweets)



Text Blob (Including Retweets)

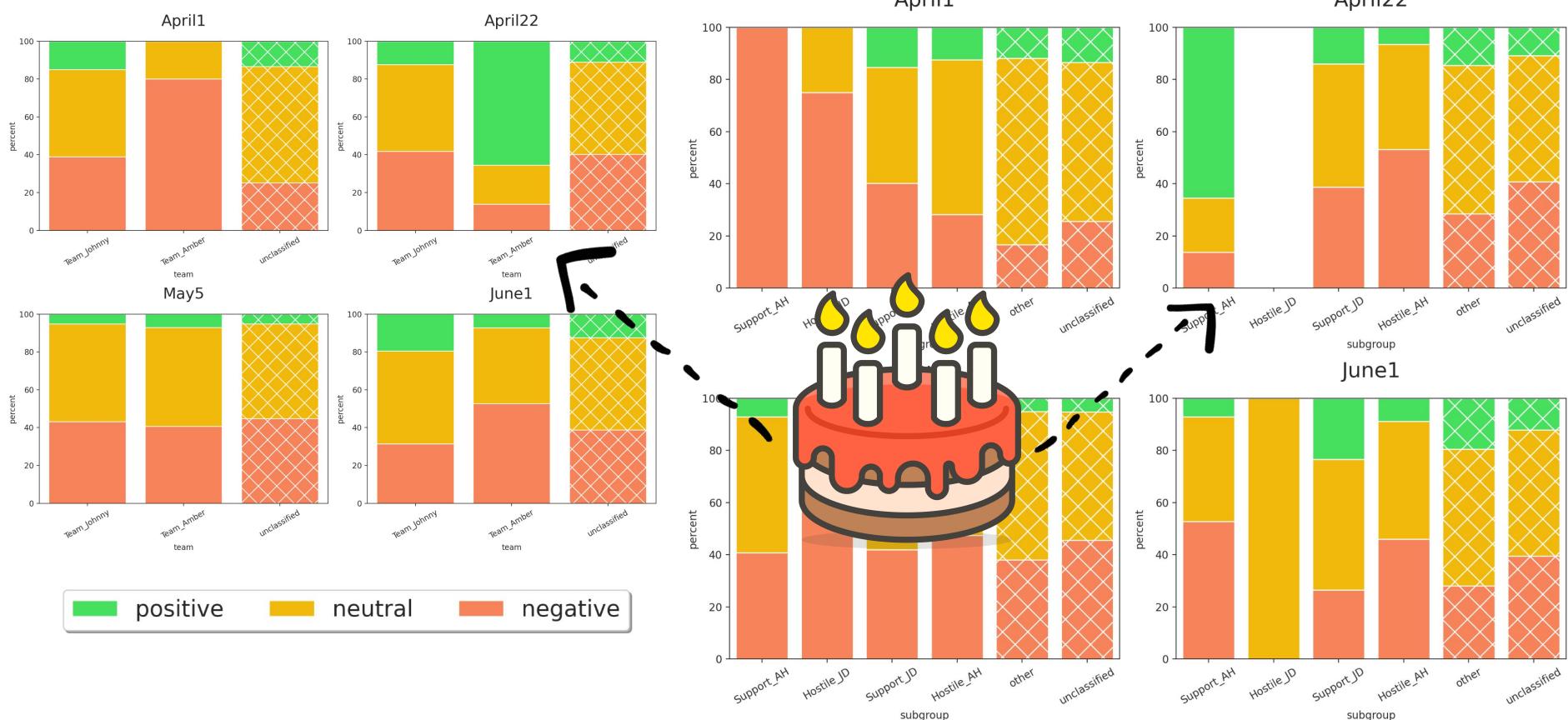


NOTICE!

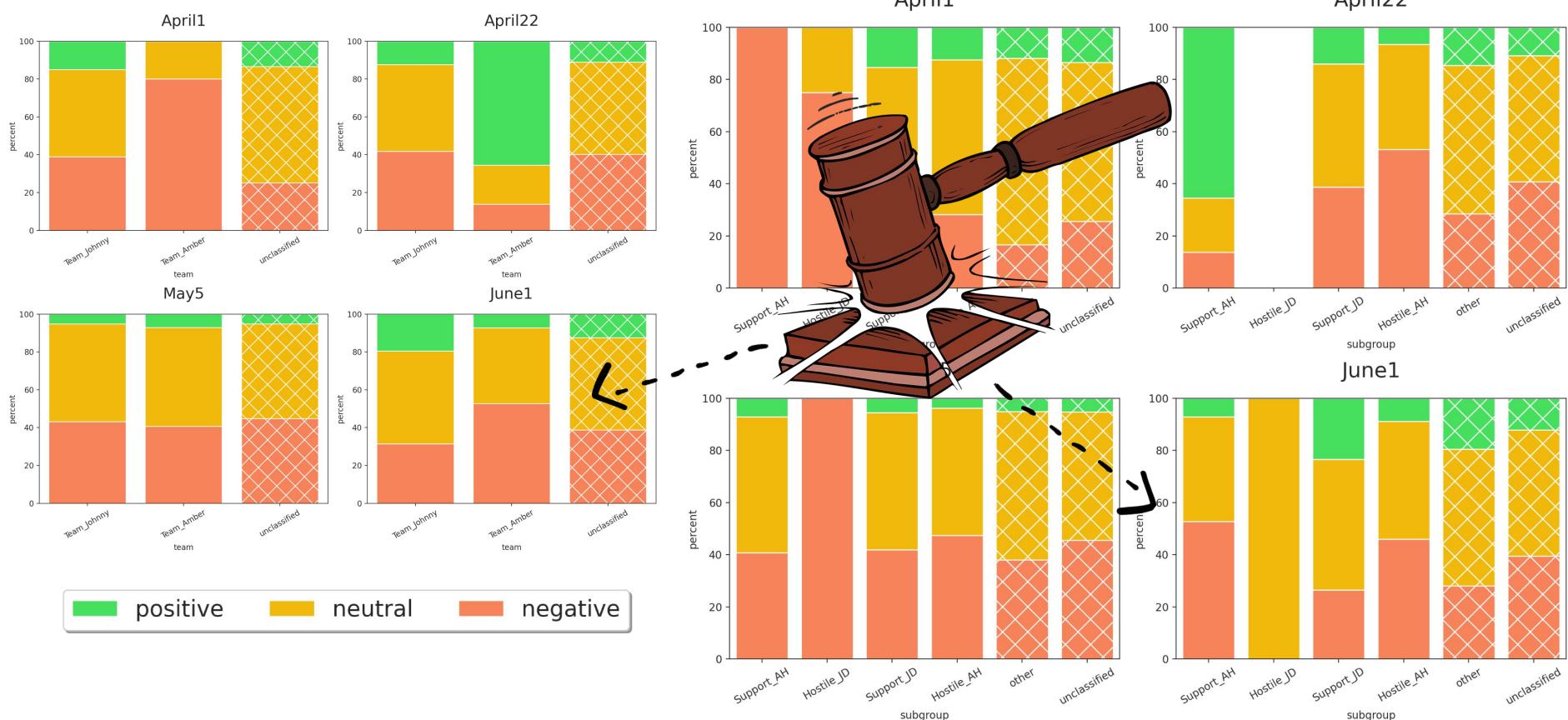


We did this to learn about different tools/models. To *properly* compare them, we would have had to use them on the same data set/corpus! We did *not* do that (we ran out of time...).

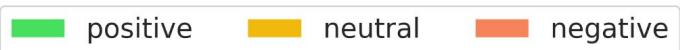
Results Sentiment Analysis, Roberta



Results Sentiment Analysis, Roberta



Results Sentiment Analysis



TextBlob (polarity ratio)

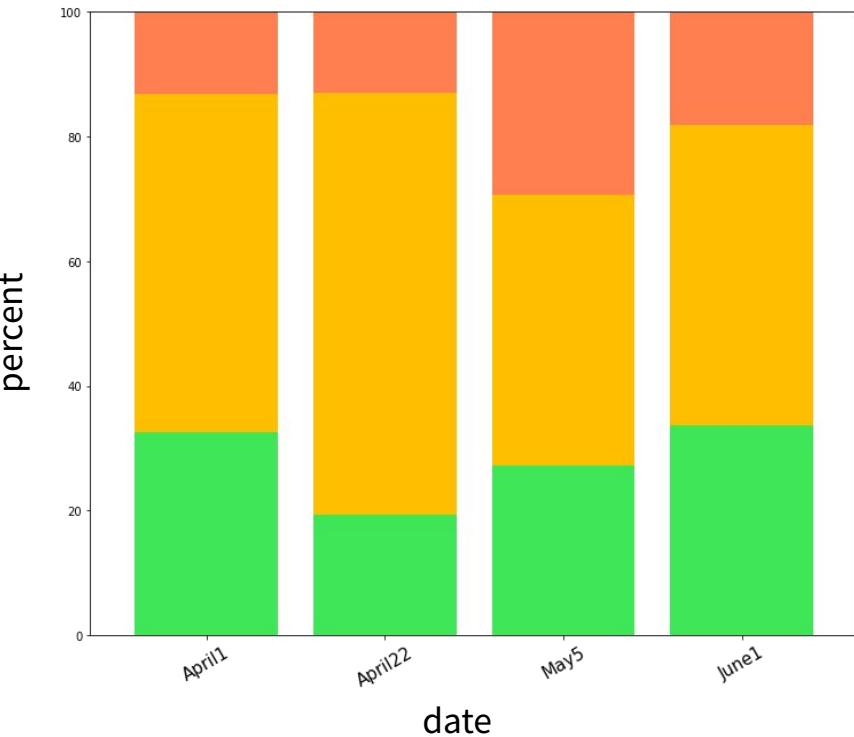
	Positive(%)	Neutral(%)	Negative(%)
--	-------------	------------	-------------

April1	32.48	54.29	13.23
--------	-------	-------	-------

April22	19.40	67.60	13.00
---------	-------	-------	-------

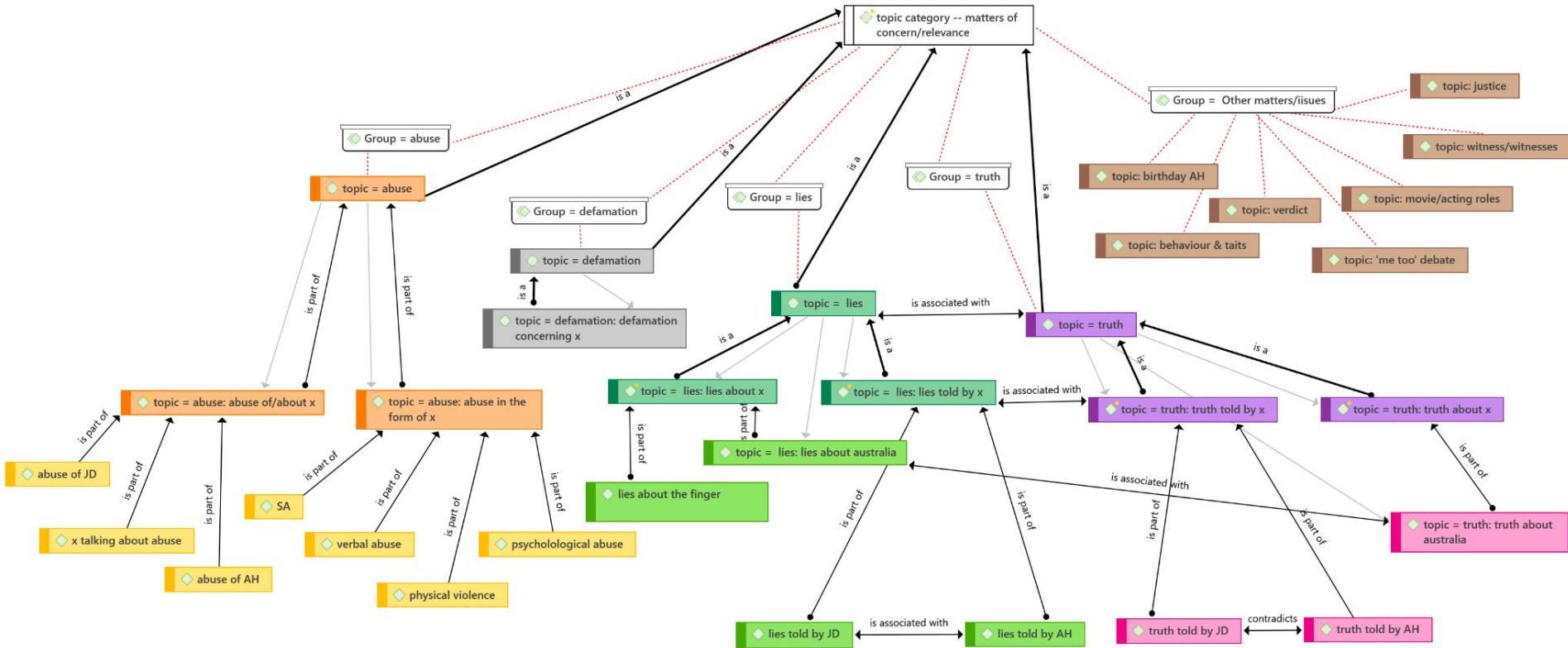
May5	27.20	43.46	29.34
------	-------	-------	-------

June1	33.64	48.29	18.06
-------	-------	-------	-------



Results Content Analysis

Structured qualitative analysis after Kuckartz
Using ATLAS.ti 22



Integrated Results

Main Result Network

- The reply network is modularized
- People had different tweets with different detected “team”
- Most of reply network nodes are “unclassified”
- Debate was mostly done on May 5

Main Results Sentiment Analysis

- Relation of sentiment values of Team Johnny stayed similar over time
- Relation of sentiment values of Team Amber changed a lot over time
- Sometimes sentiments contradict classification of hashtags

Main Result Content Analysis

- A lot of different topics in the tweets
- Barely anything on the topic of the trial (Defamation)
- Mainly discussions about lies, truth, and abuse

Please take every result with a grain of salt! They are ***preliminary***.

Limitations

Data Set

- We wanted to have all tweets from 1 week before to 1 week after the trial → was too much
- We picked 3 days (approx 12 hours for each, 24 hours would have taken too much time to download)
- Selectivity regarding hashtags
⇒ Affects representativity, generalisability, ...

- Content of the tweets in the category ‘other’ and ‘unclassified’ is (mostly) unknown

Sentiment Analysis

- TextBlob calculated the polarity rate based on the short sentence cleaned from tweets, which cannot represent the overall sentiment but just what can be extracted from the limited corpus used

Network Analysis

- cannot correctly depict user retweets (as nodes in network)
- we don’t know how many tweets we missed...

Content Analysis

- no hostile_JD?
- ‘other’ and ‘unclassified’ not investigated
- only started coding the data, we have structured topics, but don’t yet know what they ‘mean’

.. all results are preliminary .

Many Mistakes... Many Learnings!

Data collection with Twitter API

- Downloading data takes even longer (much longer) than expected ;)
→ *get more data and allocate more time for it*
- issue when downloading csv versus JSON (same query, but contains different info/parameters!) → *use twarc-csv*
- *don't collect data you cannot use* → *prioritise*

⇒ *Data cleaning takes up at least 50-60 percent of(.. our) analysis time*

Data Visualisation

- Hard or impossible to edit colours with pycatplot
→ *we used rankflow-tool*
<https://labs.polsys.net/tools/rankflow/> instead
- Use Gephi

General

- Using Colab with mounted Drive (all data and notebooks in Drive) works smoothly
→ *Use Colab GPU RAM !*
- Sarcasm is a problem

Sentiment Analysis

- The adoption of package should be suitable to the research question
- It is necessary to compare results validity from different modules, since the corpus used might vary a lot

Network Analysis

- ‘*reply_settings*’ is *not* needed for the reply network → *use ‘in_reply_to_user_id’ parameter*

Next Steps

General

- We did not get to the “second loop” in our research design
→ test new Twitter query and repeat :)
- Check if differences (teams, subgroups across time, etc.) are significant
- stronger integration of results from different methods/ triangulation

Sentiment Analysis

TextBlob: compare the polarity by groups and teams distinctively to indicate the polarity difference towards Amber and Depp and the change over time. It would also be possible to link the actual situation together with the public opinion to test our initial hypotheses.

Network analysis

- Include the ‘retweet.authorid’ in expansions parameter for further analysis and download the data as .json

Qualitative Analysis

- Focus on subgroups
- continue analysis, identify more topics of relevance to users AND check if meaningful!
- refine and validate code system

... stay in touch, keep learning and making productive mistakes.

Thank you for your attention!

Questions, feedback, critique, etc.

are very welcome 