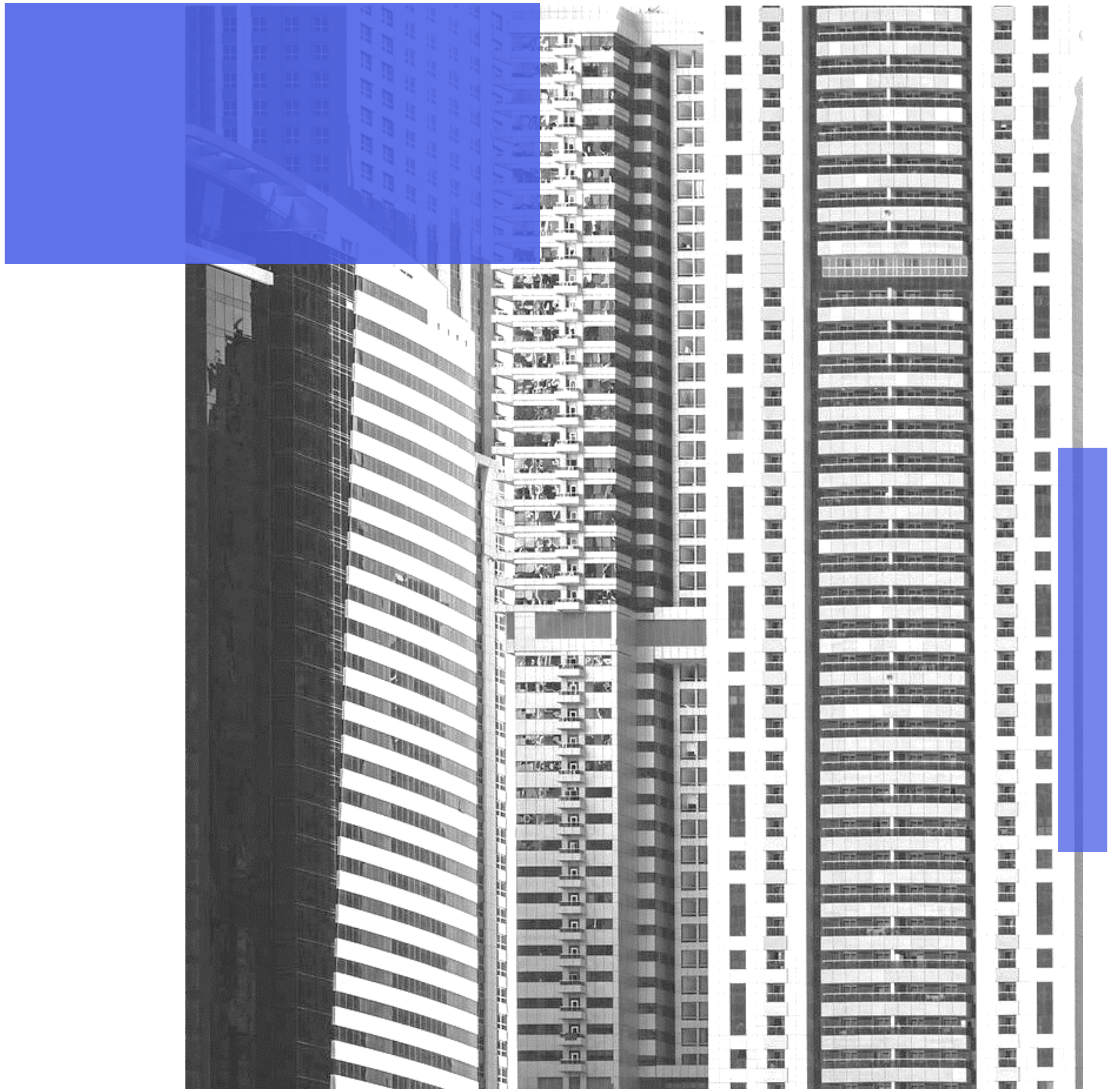
TEAM 13 – GABOR AND DANIEL

UNDERDOGS IN THE HOOD: COMPARING MASTODON SOCIAL & TRUTH SOCIAL



AGENDA

BACKGROUND AND RATIONALE

RESEARCH QUESTIONS

RESULTS

Q&A

BACKGROUND & RATIONALE

Social networks are omnipresent in modern lifestyle. Different social networks have distinguishing aspects and attract different audiences. Using graph theory and sentiment analysis will lead to insights into social implications and the effects these networks could have on societies at large (information bubbles, fake news, information dissemination).

A TALE OF THREE SOCIAL NETWORKS

2006: Twitter starts, popularising its microblogging structure for social media

2016: Mastodon is founded as an OS alternative to Twitter

2021: President Trump gets banned from Twitter following the US Capitol attack

2021: Truth Social is founded as an alternative to Twitter by Trump

2022: Twitter is bought by Elon Musk

2022: Significant userbase increase on Mastodon following the acquisition

WHAT DO WE CONTRIBUTE?

- The network we analyse share core design features (microblogging)
  - Tweets (= X posts) = Toots = Truths
- But they share a very different inherent user community!
- Can we detect these inherent differences using analysis methods?
- If we can, then what does it say about the networks?
- Is a network inherently more prone to adverse effects than others?

RESEARCH QUESTIONS

We have two Twitter alternatives. What questions could be answered using network and sentiment analysis ?

We've split our research into two parts based on these approaches.

NETWORKS & SENTIMENTS

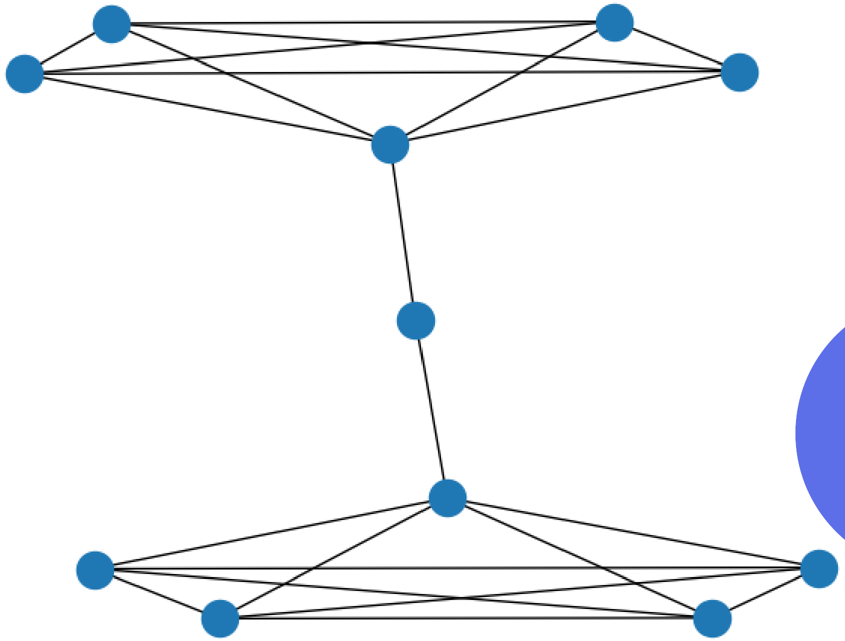
Network structure of users as follower-followee graph

1. How do the social networks differ in their network characteristics?
2. How do they compare to model graphs and their characteristics?
3. How do they differ from other protocols?

Sentiment analysis of social network content

1. How do the social networks differ in terms of their contents' positivity/negativity?

RESULTS: NETWORK ANALYSIS



The network analysis yielded that the two social networks behave exactly as expected from a human engineered network. Only slight differences can be found between the two networks.

THE DATA USED

Truth Social Dataset

(Gerard, Botzer & Weninger, 2023)

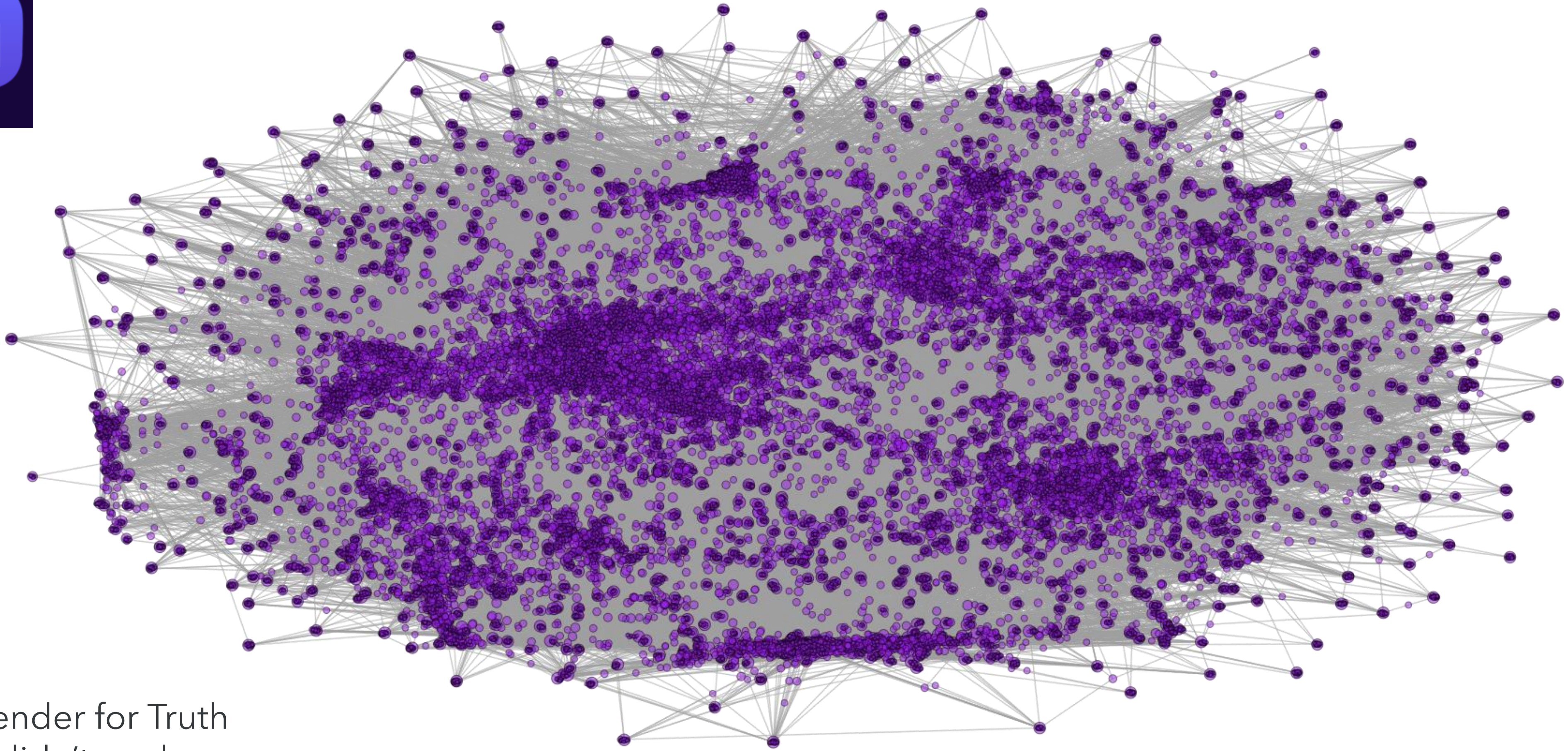
- Sample size: 454 458 users
- Population: ~ 4.2 million users
- 10.6% coverage in BFS manner
- Date: 2022 Sept.-Dec.

mastodon.social Dataset

Mined using R and rtoot

- Sample size: 54 285 users
- Population: ~ 950 000 users
- 5.7% coverage in BFS manner (limited BFS coverage per node)
- Date: 2024 June

A LOOK ON THE GRAPHS*



*The render for Truth Social didn't work

MASTODON VS. TRUTH SOCIAL

Metrics	Mastodon	Truth Social
Order	54 285	454 459
Size	140 030	4 002 115
Average degree	5.16	17.61
Components (W / S)*	1 / 50 209	9 589 / 452 043
Connection	Weak	Not
Connectivity (E / V)	0 / 0	0 / 0
Diameter	9	9
Mean path length	4.81	3.27
Transitivity	1.33%	1.05%

Due to sample size

Due to sampling method

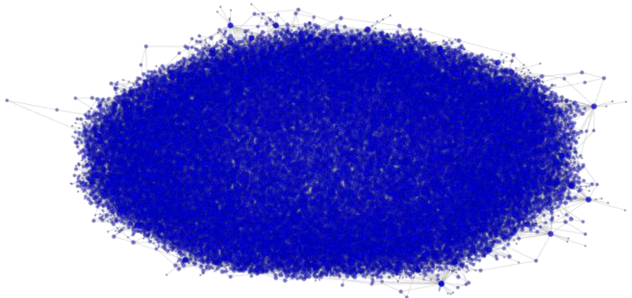
Similarity

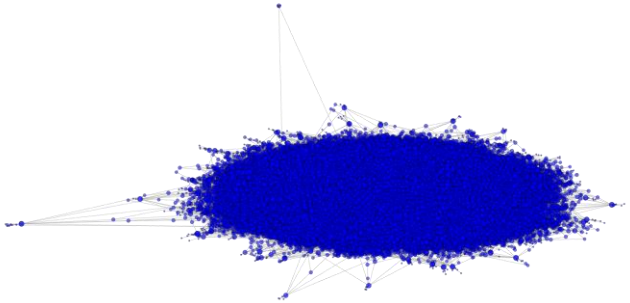
Difference

*Truth Social have an average path length and transitivity that suggests **clustering around hubs without neighbourhood connections**! (Note: the average degree too)*

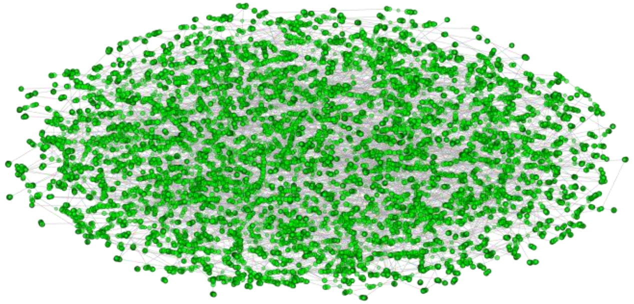
THESE ARE THE MODELS

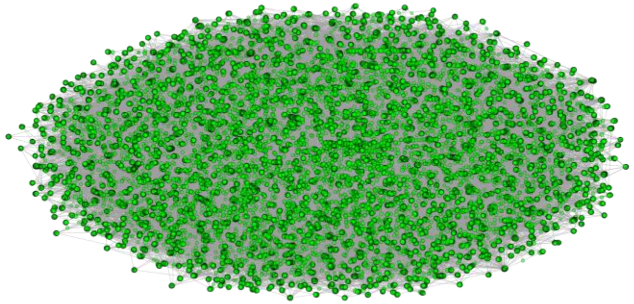
Erdős-Rényi
(GNP)



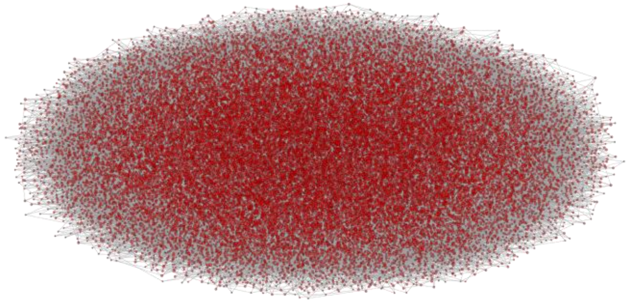


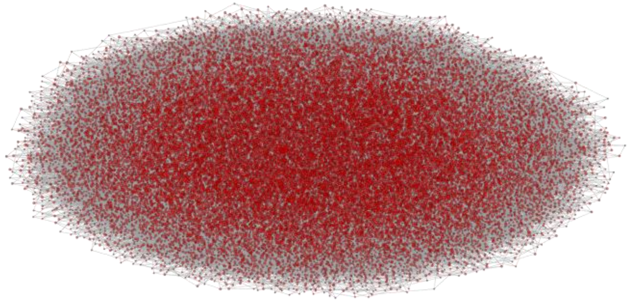
Watts-Strogatz
(SWG)





Barabási-Albert
(PA)





MASTODON VS. REAL-LIFE

Metrics	Mastodon	GNP (1:2)	SWG (1:2)	PA (1:2)
Order	54 285	27 143	27 143	27 143
Size	**140 030**	**65 055**	43 859	**81 423**
Average degree	**5.16**	**4.79**	3.23	**5.99**
Components	**1 / 50 209**	1 / 11432	1 / 1	**1 / 27 143**
Diameter	**9**	32	14	16
Mean path length	**4.81**	12.88	11.49	**4.14**
Transitivity	**1.33%**	0.0002%	29.4%	0.12%

Mastodon is most akin to preferential attachment graphs. This is ideal as they are the closest to social networks. Although human engineered networks never work like actual scale-free graphs due to real life constraints or design choices. This finding validates the structure of the Mastodon dataset in a preliminary sense.

TRUTH SOCIAL VS. REAL-LIFE

Metrics	**Truth Social**	**GNP (1:10)**	**SWG (1:10)**	**PA (1:20)**
Order	454 459	45 446	45 446	45 446
Size	4 002 115	94 701	162 468	408 968
Average degree	**17.61**	4.17	7.15	**18.00**
Components	**9 589 / 452 043**	1 / 23 887	1 / 30	**1 / 45446**
Diameter	**9**	38	**9**	18
Mean path length	**3.27**	14.85	7.22	**4.07**
Transitivity	**1.05%**	0.008%	42.55%	0.2%

Similar findings to the Mastodon network. Truth Social resembles in many aspects a preferential attachment graph even more and only differs in expected areas (diameter, transitivity), as the attachment effect is much larger in a social network.

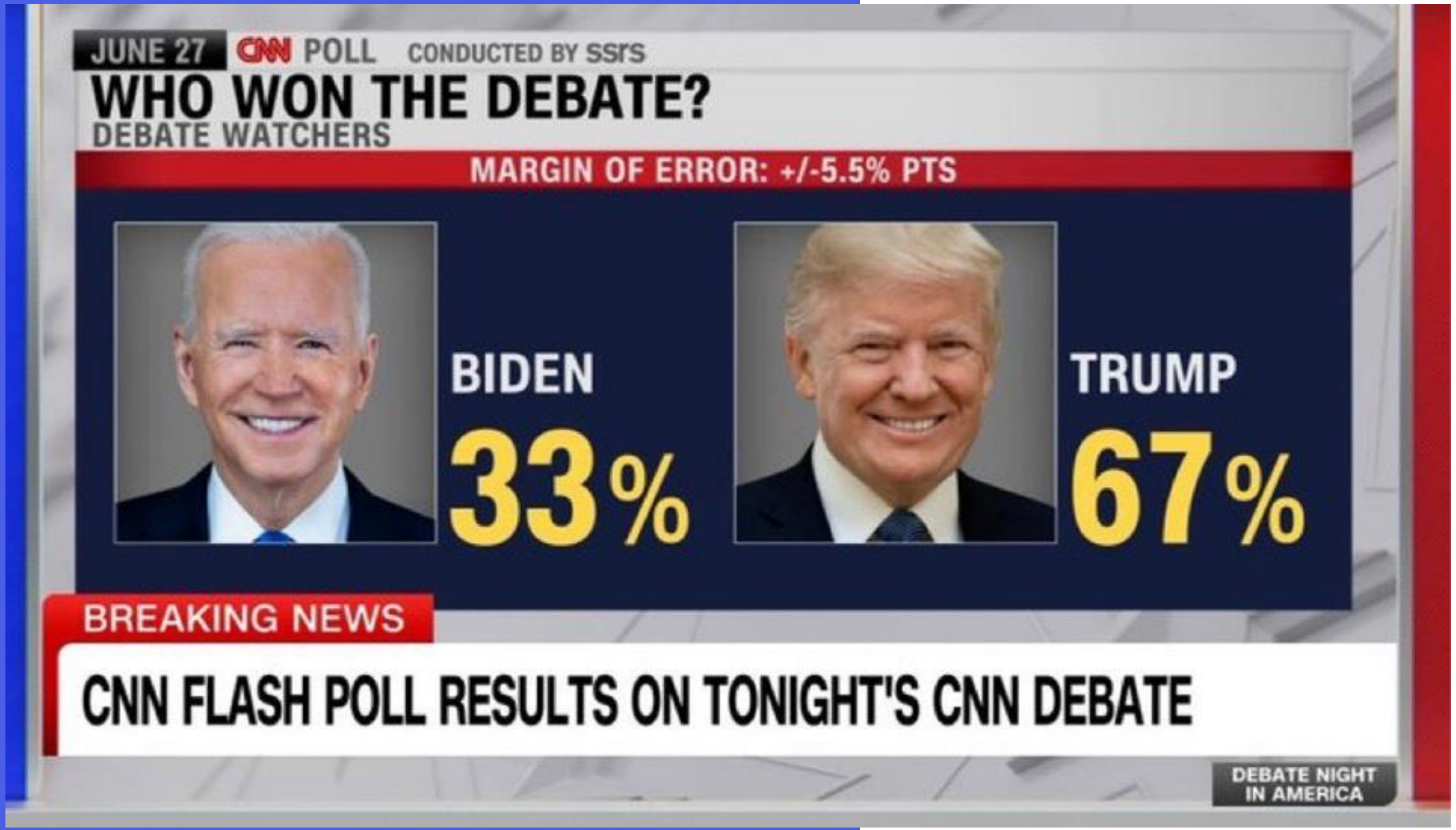
VS. OTHER SOCIAL NETWORKS

Metrics	Mastodon	Truth Social	Twitter (2012)*	Facebook (2011)**
Order	54 285	454 459	~175 000 000	63 400
Size	140 030	4 002 115	~20 000 000 000	12 580 000
Average degree	**5.16**	**17.61**	**2.83**	**396.8**
Components (W / S)*	1 / 50 209	9 589 / 452 043	One large, many small	One large, some small?
Connection	Weak	Not	Not	Not
Connectivity (E / V)	0 / 0	0 / 0	0 / 0	0 / 0
Diameter	**9**	**9**	N/A	**8.75**
Mean path length	**4.81**	**3.27**	**4.05**	N/A
Transitivity	**1.33%**	**1.05%**	N/A	**7.89%**

*Myers et al., 2014
**Catanese et al., 2011

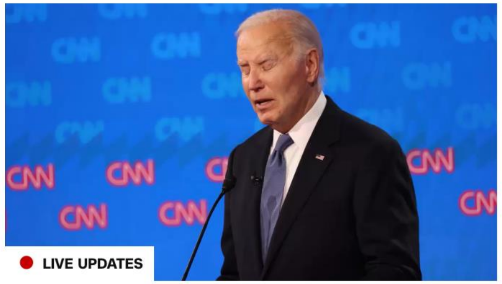
Not being a microblog immediately shows

RESULTS: THE DEBATE



CNN | PRESIDENTIAL DEBATE

Biden's poor showing and Trump's repeated falsehoods



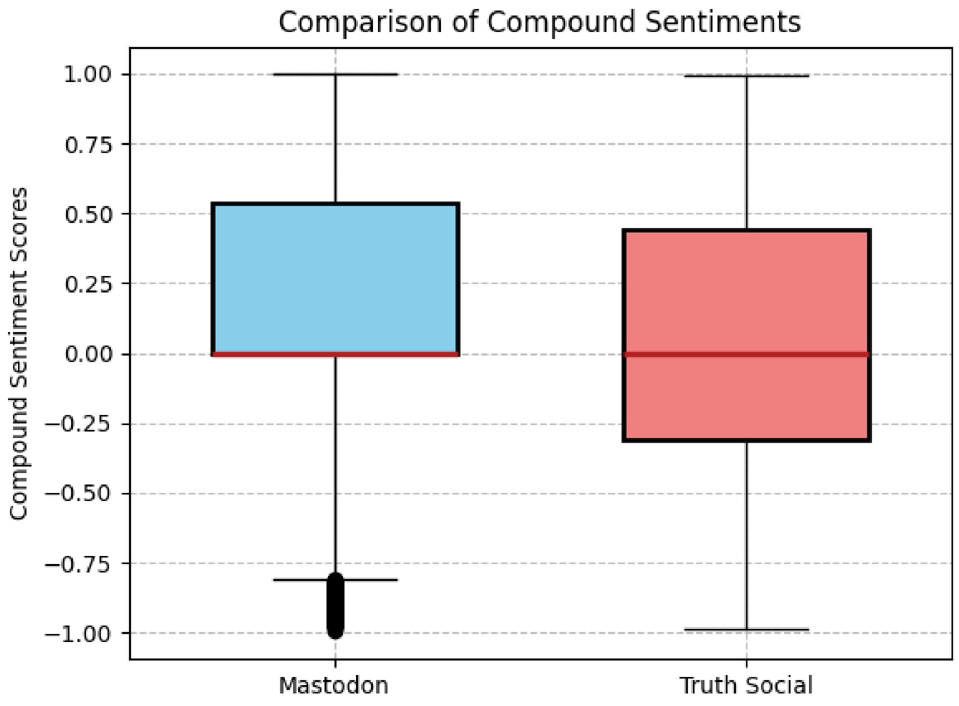
'That was very painful': Democrats lament Biden's performance as Trump dodges on multiple fronts

BLOODBATH: TRUMP DOMINANT IN DEBATE

BIDEN DISASTER: WORST PERFORMANCE — *OF ALL TIME?*

BREITBART

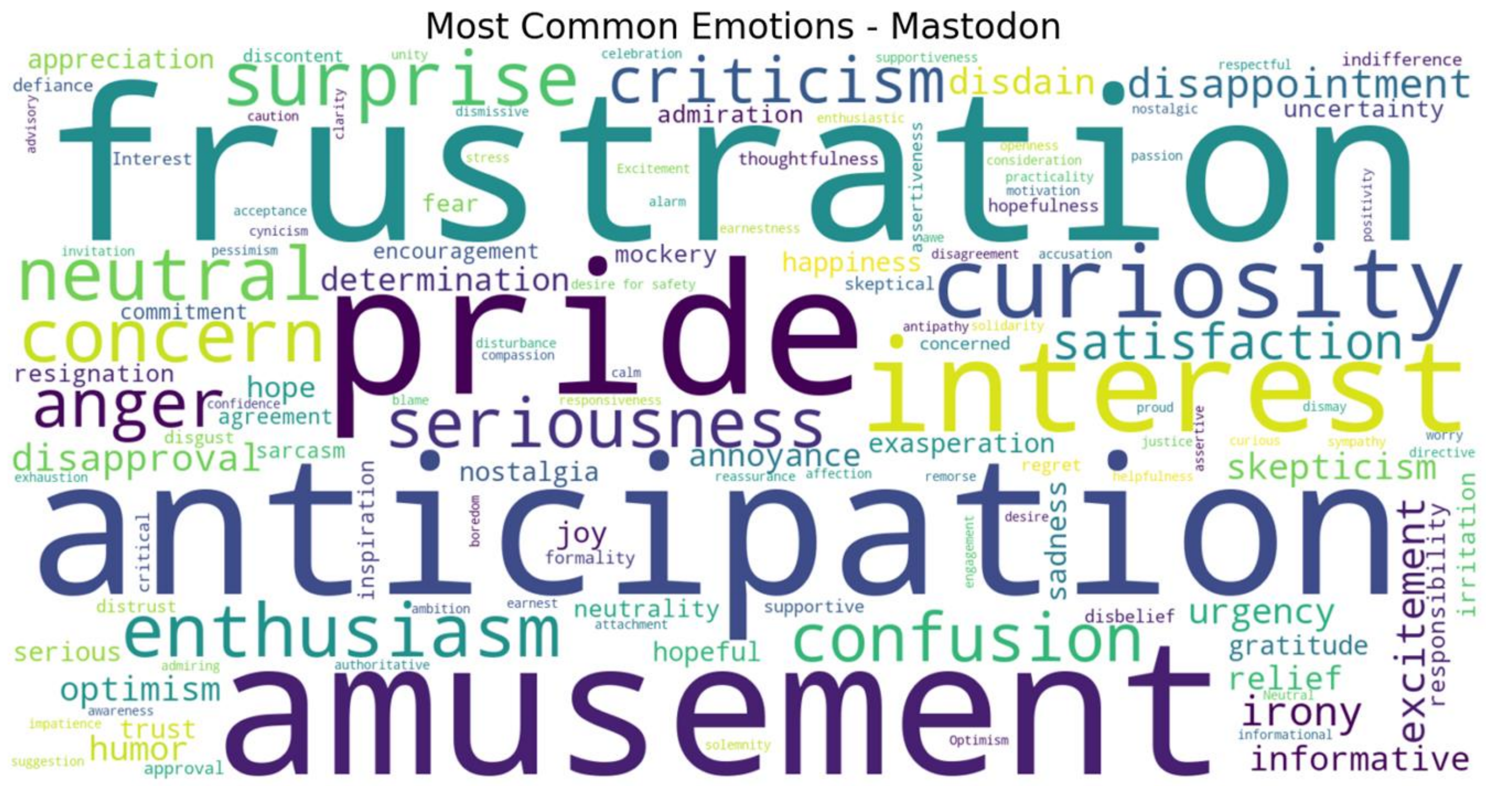
RESULTS: POLARITY



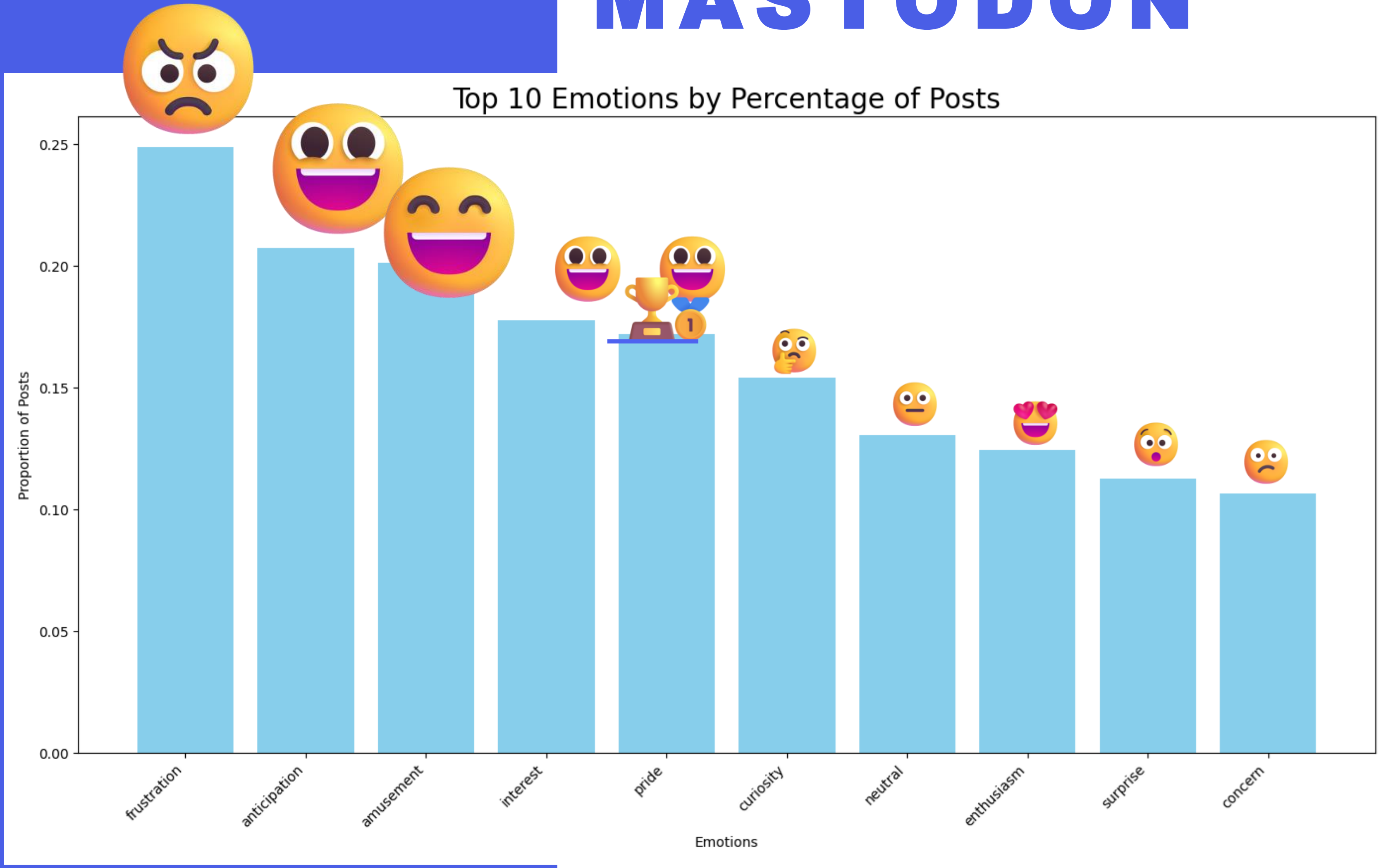
RESULTS: EMOTIONS ANALYSIS

Positive		Negative		Ambiguous
admiration 👏	joy 😃	anger 😡	grief 😢	confusion 😕
amusement 😂	love ❤️	annoyance 😒	nervousness 😬	curiosity 🤔
approval 👍	optimism 🤞	disappointment	remorse 😔	realization 💡
caring 🤗	pride 😌	disapproval 👎	sadness 😞	surprise 😲
desire 😍	relief 😅	disgust 🤮		
excitement 🤩		embarrassment 😳		
gratitude 🙏		fear 😨		

RESULTS: EMOTIONS ANALYSIS MASTODON

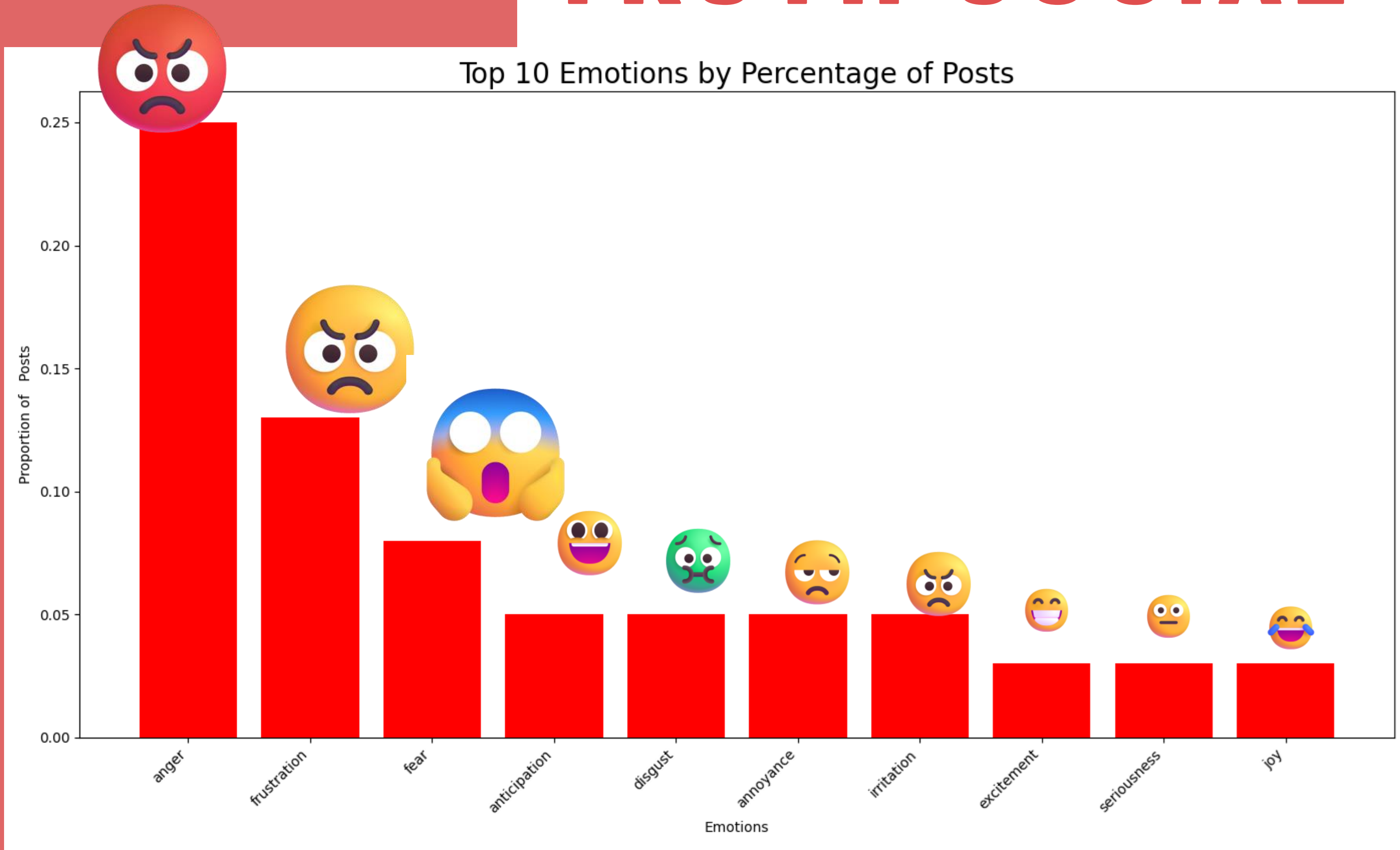


RESULTS: EMOTIONS ANALYSIS MASTODON



RESULTS: EMOTIONS ANALYSIS TRUTH SOCIAL

RESULTS: EMOTIONS ANALYSIS
TRUTH SOCIAL



RESULTS: SARCASM ANALYSIS

- Not only is he not in control, he was never in control
- *#FloridaManFlorida man sneezes his intestines out of his body at restaurant*
- This isn't America's Got Talent, goddamnit. This is about saving democracy.
- They hate women but want to imprison them in marriage
- CNN has decided, based on its pre-debate "coverage", that the two major-party candidates are very much equivalent.

- Vocabulary-based approaches (Maynard & Greenwood, 2014)
- We can do better?! GPT-4-T and $
    - Accuracy: 84,2%
    - Precision: 47,4% 👎
    - Recall: 81,8%
- No, we can't :) But we checked!

LIMITATIONS & IMPROVEMENT

Data gathering:
- Very time consuming
- BFS implementation not exact
- Dubious description of original mining method

Data analysis:
- Very computation intense (also due to R)
- Some metrics had to be left out (centralities)

How to improve:
- Parallelization in another environment
- Fetch both datasets using the same method
- Longer timeframe

QUESTIONS
ANSWERS

Please give us your insights on the topic and ask away. We would love to discuss!

We also have code:

[wu-mozol/DSAI2 (github.com)](https://github.com/wu-mozol/DSAI2)

THANK YOU

Daniil Dobriy & Gábor T. Mozol – Team 13

daniil.dobriy@wu.ac.at

gabor.tamas.mozol@wu.ac.at

REFERENCES 1

Barabási, A. L., & Eric Bonabeau. (2003, May 1). Scale-Free Networks. Scientific American Magazine.

Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. Science (New York, N.Y.), 286(5439), 509–512. doi:10.1126/science.286.5439.509

Bian, T., Xiao, X., Xu, T., Zhao, P., Huang, W., Rong, Y., & Huang, J. (2020). Rumor Detection on Social Media with Bi-Directional Graph Convolutional Networks. Proceedings of the . . . AAAI Conference on Artificial Intelligence, 34(01), 549–556. https://doi.org/10.1609/aaai.v34i01.5393

Catanese, S. A., De Meo, P., Ferrara, E., Fiumara, G., & Provetti, A. (2011, May 25). Crawling Facebook for social network analysis purposes. Proceedings of the International Conference on Web Intelligence, Mining and Semantics. Presented at the WIMS '11: International Conference on Web Intelligence, Mining, Sogndal Norway. doi:10.1145/1988688.1988749

Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., & Ravi, S. (2020). GoEmotions: A dataset of fine-grained emotions. arXiv preprint arXiv:2005.00547.

Doerr, B., Fouz, M., & Friedrich, T. (2012). Why rumors spread so quickly in social networks. Communications of the ACM, 55(6), 70-75.

Erdős, P., & Rényi, A. (1959). On random graphs I. Publ. Math. Debrecen, 6(290–297), 18.

Gerard, P., Botzer, N., & Weninger, T. (2023, June). Truth social dataset. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 17, pp. 1034-1040).

REFERENCES 2

Maynard, D. G., & Greenwood, M. A. (2014, March). Who cares about sarcastic tweets? investigating the impact of sarcasm on sentiment analysis. In Lrec 2014 proceedings. ELRA.

Myers, S. A., Sharma, A., Gupta, P., & Lin, J. (2014). Information network or social network? the structure of the twitter follow graph. Proceedings of the 23rd International Conference on World Wide Web, 493–498. Presented at the Seoul, Korea. doi:10.1145/2567948.2576939

Nettleton, D. F. (2013). Data mining of social networks represented as graphs. Computer Science Review, 7, 1–34. doi:10.1016/j.cosrev.2012.12.001

Raponi, S., Khalifa, Z., Oligeri, G., & Di Pietro, R. (2022). Fake News Propagation: A Review of Epidemic Models, Datasets, and Insights. ACM Trans. Web, 16(3). doi:10.1145/3522756

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. Social Networks, 29(2), 173–191. doi:10.1016/j.socnet.2006.08.002

Yang, X., Lyu, Y., Tian, T., Liu, Y., Liu, Y., & Zhang, X. (2021, January). Rumor detection on social media with graph structured adversarial learning. In Proceedings of the twenty-ninth international conference on international joint conferences on artificial intelligence (pp. 1417-1423).

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. Nature, 393(6684), 440–442. https://doi.org/10.1038/30918

We Are Social, & DataReportal, & Hootsuite. (February 22, 2024). Daily time spent on social networking by internet users worldwide from 2012 to 2024 (in minutes) [Graph]. In Statista. Retrieved May 16, 2024, from https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/