

Fantastic bots

And how to find them

Kaicheng Yang

Table of content

- Social bot detection
- Social bot activities
- Human perceptions of bots
- Bots as research instrument

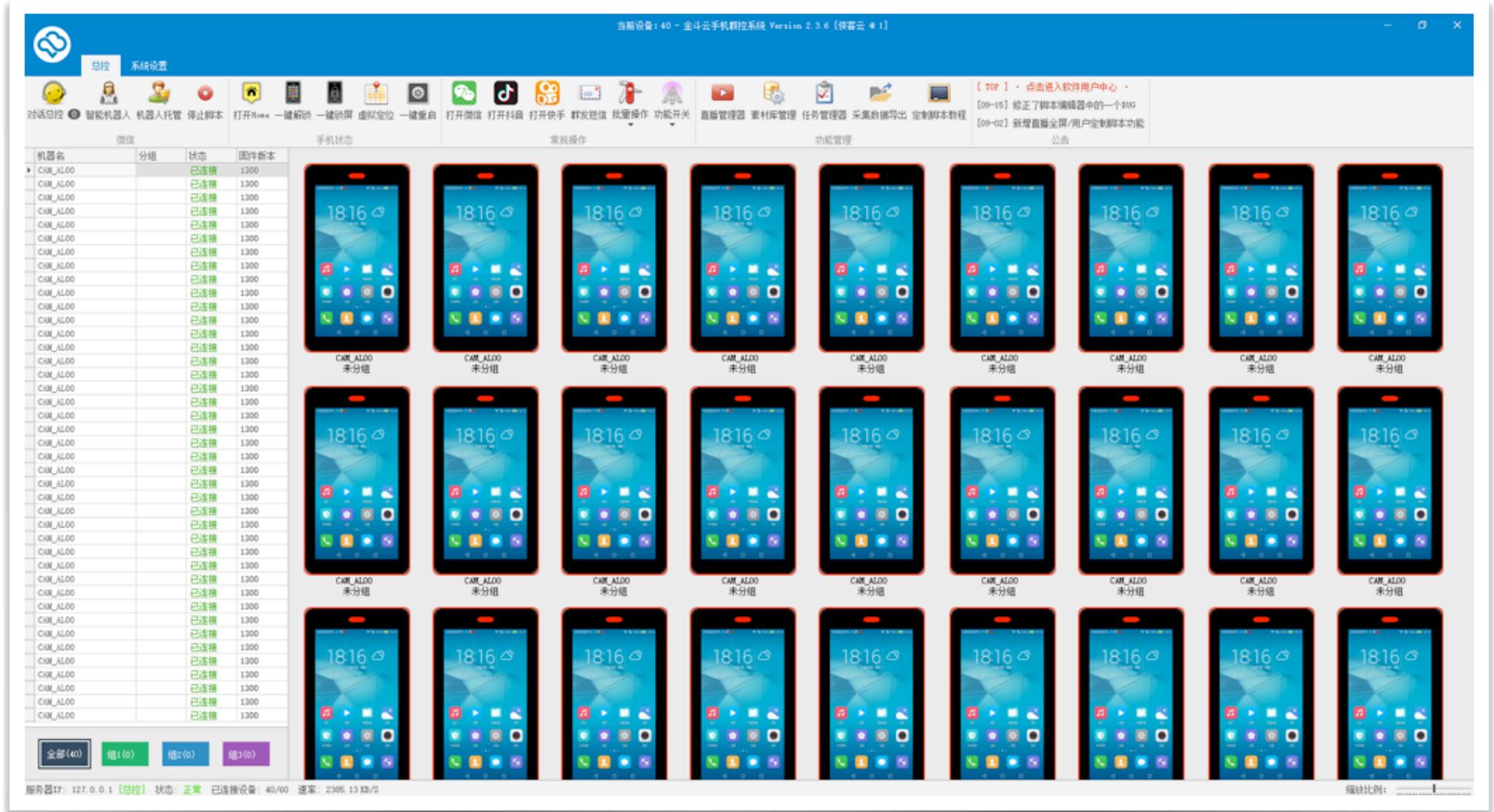
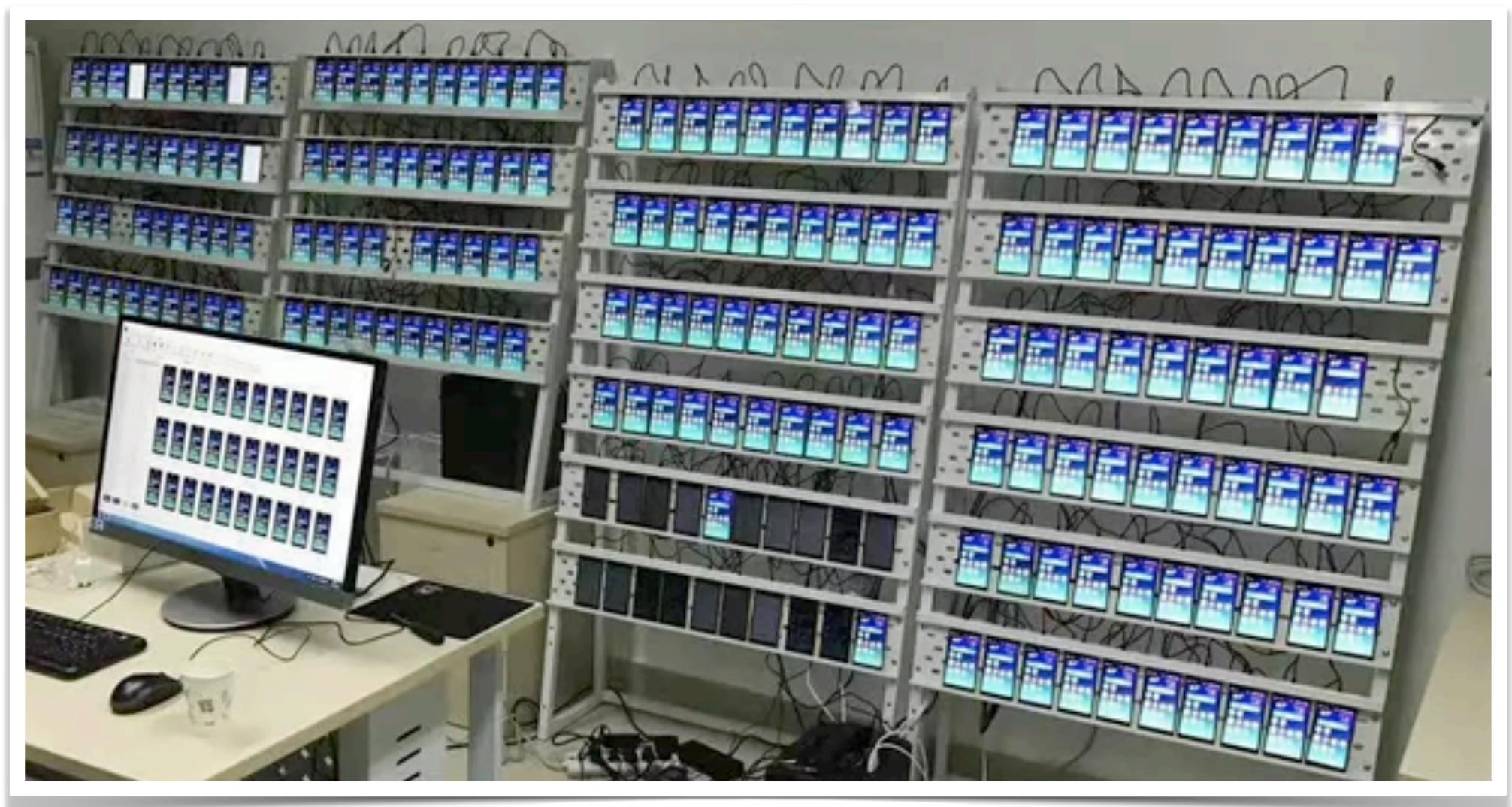
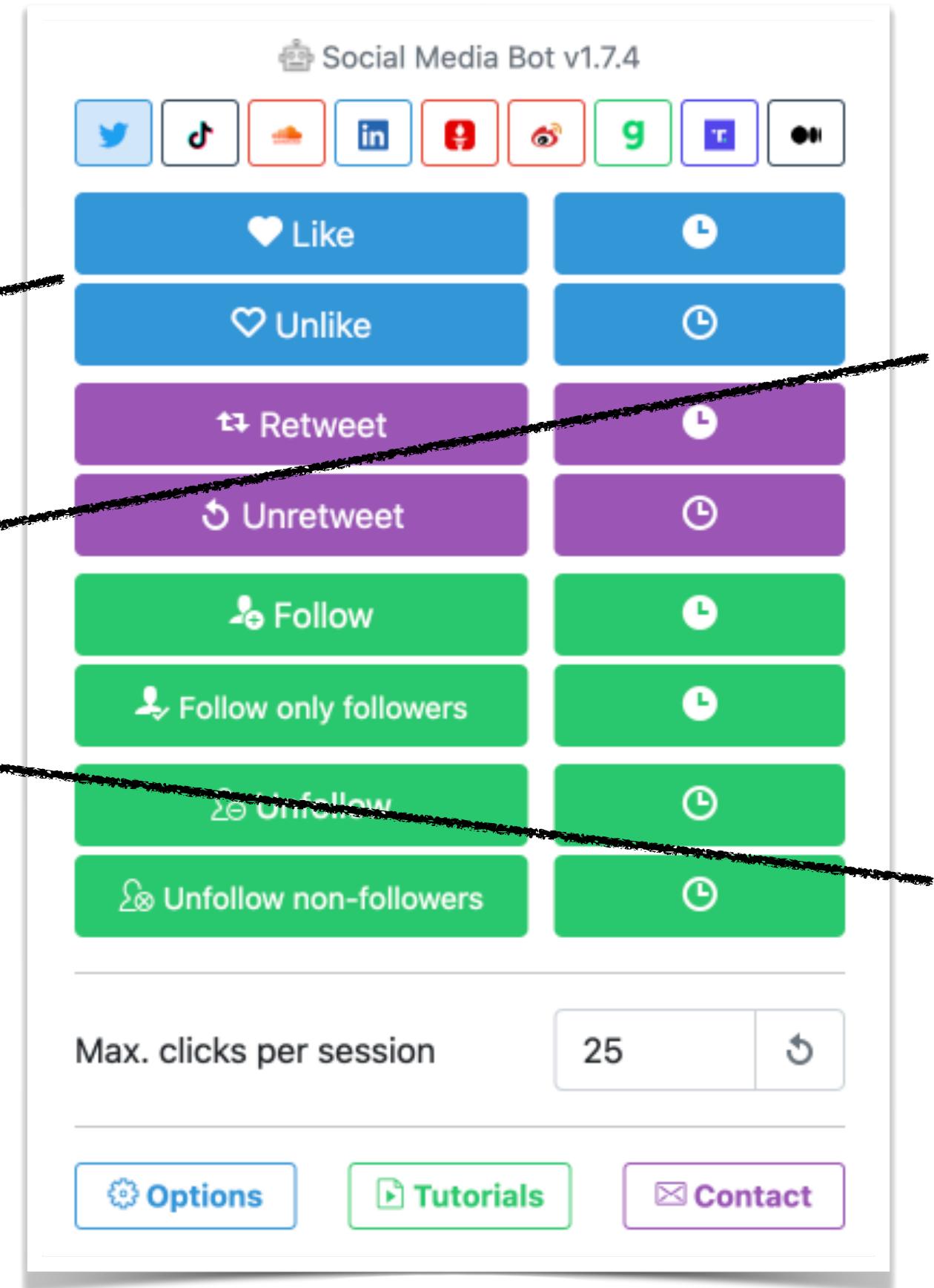
Definition of social bots

“ We broadly define social bots as social media accounts controlled at least in part by software. ”

botometer.osome.iu.edu/faq#definition

Automation technology

- API
- Browser extension
- Physical cell phones
- Virtual cell phones



Threats of malicious bots

f i s t m x n d @ ¥
PEER-REVIEWED JOURNAL ON THE INTERNET

**Social bots distort the 2016 U.S.
Presidential election online discussion
by Alessandro Bessi and Emilio Ferrara**

AJPH OPEN-THEMED RESEARCH

**Weaponized Health Communication: Twitter Bots
and Russian Trolls Amplify the Vaccine Debate**

David A. Broniatowski, PhD, Amelia M. Jamison, MAA, MPH, SiHua Qi, SM, Luhwah AlKulaib, SM, Tao Chen, PhD, Adrian Benton, MS, Sandra C. Quinn, PhD, and Mark Dredze, PhD

ARTICLE

DOI: 10.1038/s41467-018-06930-7

OPEN

The spread of low-credibility content by social bots

Chengcheng Shao  ^{1,2}, Giovanni Luca Ciampaglia  ³, Onur Varol  ¹, Kai-Cheng Yang¹, Alessandro Flammini^{1,3} & Filippo Menczer  ^{1,3}

JOURNAL OF MEDICAL INTERNET RESEARCH

Himelein-Wachowiak et al

Viewpoint

Bots and Misinformation Spread on Social Media: Implications for COVID-19

PEW RESEARCH CENTER | OCTOBER 10, 2018

SOCIAL MEDIA BOTS DRAW PUBLIC'S ATTENTION AND CONCERN



**About two-thirds of Americans have heard about
social media bots, most of whom believe they are
used maliciously**

\$3 = 1,000 retweets!

ID	Service	Rate per 1000	Min order	Max order	Description
Twitter Retweets					
5615	Twitter Arab Gulf Retweets + Bonus LIKES [Start time= 0-24H]	\$3.60	10	500	View
4606	Twitter Arab Gulf Retweets	\$6.39	100	5000	View
3974	Twitter Retweets [Arab MIX] [Start time= 0-1H - FAST]	\$12.00	100	3000	View
6183	Twitter Retweets Max 5K SuperInstant Speed: 5K/Day	\$2.63	10	5000	View
6184	Twitter Retweets Max 10K SuperInstant Speed: 10K/Day	\$2.88	10	10000	View
6185	Twitter Retweets Max 20K SuperInstant Speed: 20K/Day	\$3.13	10	20000	View

markenetspanel.com

Bots supercharged with AI?

AI-synthesized faces are indistinguishable from real faces and more trustworthy

Sophie J. Nightingale  and Hany Farid [Authors Info & Affiliations](#)

Edited by William Press, Computer Sciences and Integrative Biology, University of Texas at Austin, Austin, TX; received November 11, 2021; accepted December 20, 2021

February 14, 2022 | 119 (8) e2120481119 | <https://doi.org/10.1073/pnas.2120481119>

The Next Great Misinformation Superspreaders: How ChatGPT Could Spread Toxic Misinformation At Unprecedented Scale

We tempted the AI chatbot with 100 false narratives from our catalog of Misinformation Fingerprints™. 80% of the time, the AI chatbot delivered eloquent, false and misleading claims about significant topics in the news, including COVID-19, Ukraine and school shootings.

By [Jack Brewster](#), [Lorenzo Arvanitis](#), and [McKenzie Sadeghi](#)

All Internet Observatory News / Blogs / January 11, 2023

Forecasting potential misuses of language models for disinformation campaigns—and how to reduce risk

Authors: Josh A. Goldstein, Girish Sastry, Micah Musser, Renée DiResta, Matthew Gentzel, Katerina Sedova

Stanford Internet Observatory

How to identify bots?

The screenshot shows the Botometer website interface. At the top, there is a navigation bar with links to FAQ, API, Publications, Bot Repo, BotAmp, BEV, Lite, and a user profile (@OSoMe_IU). The main header features the "Botometer®" logo and the text "An OSoMe project (bot•o•meter)". To the right of the header is a cartoon illustration of a blue bird and a blue robot bird interacting.

Below the header, a text box contains information about the service: "Botometer (formerly BotOrNot) checks the activity of a Twitter account and gives it a score. Higher scores mean more bot-like activity." It also mentions requirements for Twitter authentication and permissions, and provides links to the FAQ and Why? sections. It concludes by stating that Botometer is a joint project of the Observatory on Social Media (OSoMe) and the Network Science Institute (IUNI) at Indiana University.

Below this text box are four buttons: "trutherbotprop", "Check user", "Check followers", and "Check friends".

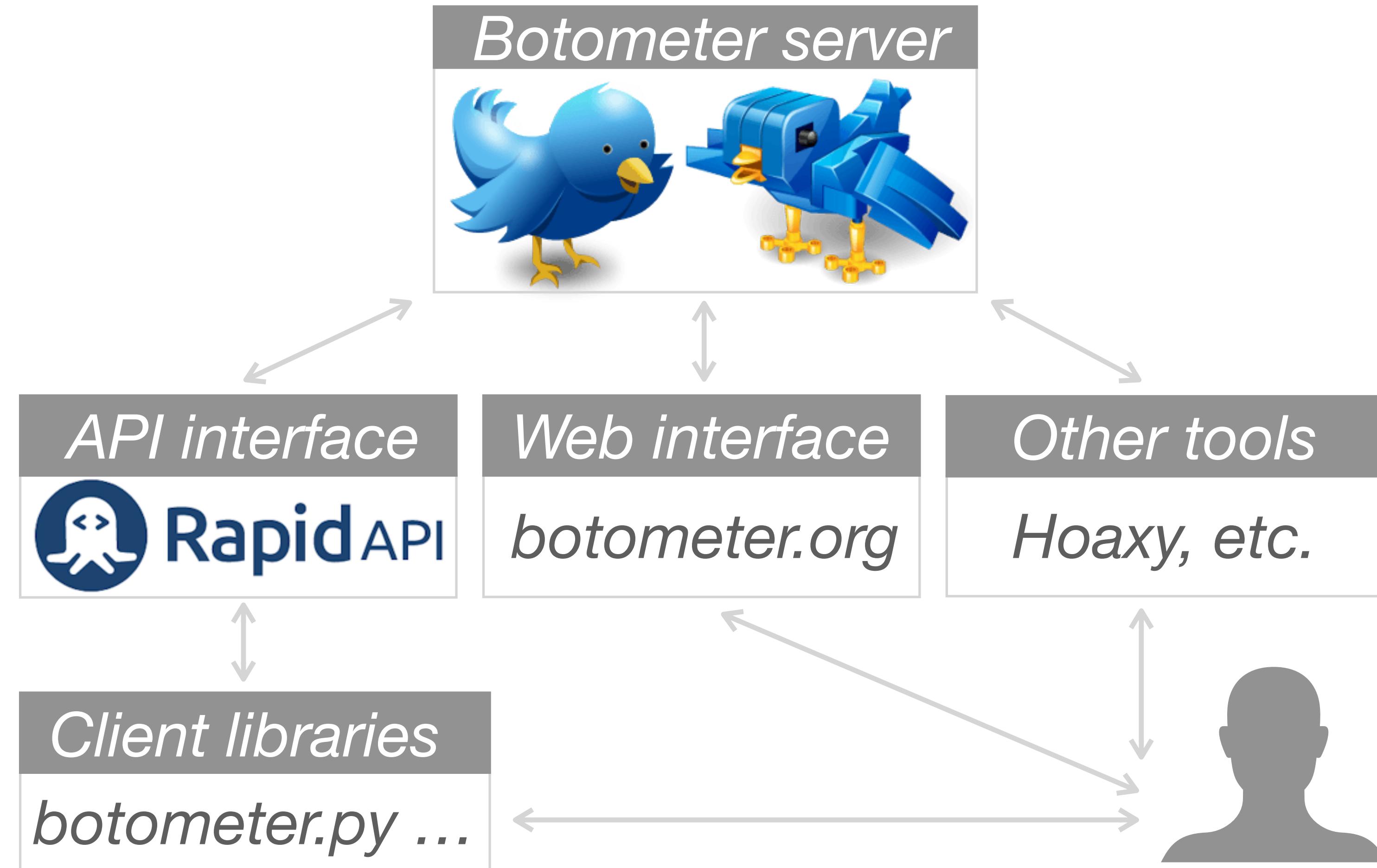
The main content area displays three Twitter accounts with their bot scores:

- @trutherbotprop: Score 4.2 / 5 (Orange background)
- @elonmusk: Score 0.9 / 5 (Green background)
- @yang3kc: Score 0.4 / 5 (Blue background)

Each account entry includes a small profile picture, the Twitter handle, and a checkmark icon if applicable. To the right of each score is a semi-circular progress bar with a black needle pointing to the value. An "x" icon is located at the end of each progress bar.

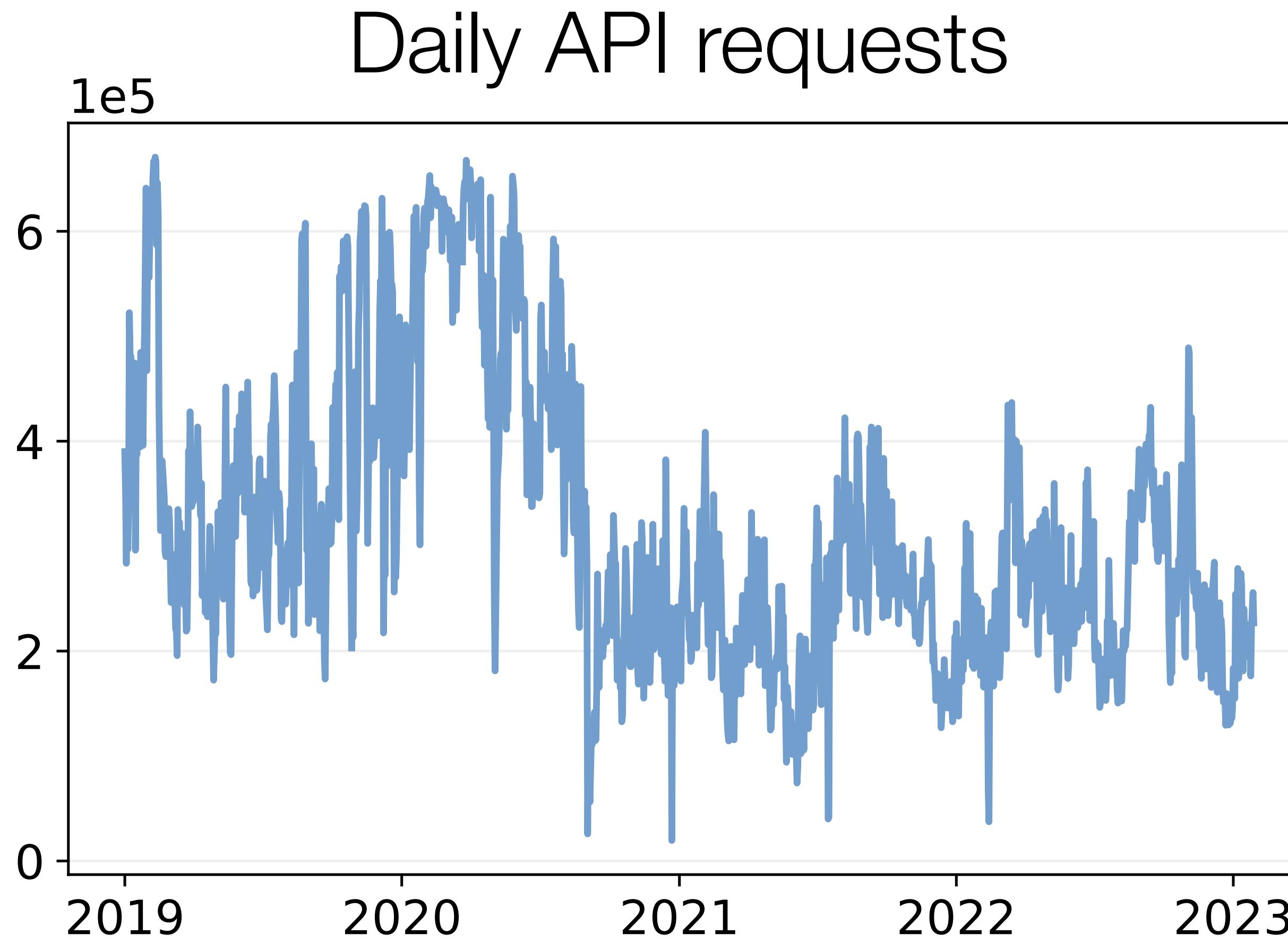
At the bottom right of the main content area is a button labeled "Export data".

Botometer ecosystem



Hoaxy: hoaxy.osome.iu.edu

Botometer has many users



Botometer supported hundreds of research projects

Science

Current Issue First release papers Archive About Submit manuscript

HOME > SCIENCE > VOL. 363, NO. 6425 > FAKE NEWS ON TWITTER DURING THE 2016 U.S. PRESIDENTIAL ELECTION

Vol. 363, No. 6425

RESEARCH ARTICLE



Fake news on Twitter during the 2016 U.S. presidential election

NIR GRINBERG , KENNETH JOSEPH , LISA FRIEDLAND , BRIONY SWIRE-THOMPSON , AND , DAVID LAZER [Authors Info & Affiliations](#)

SCIENCE • 25 Jan 2019 • Vol 363, Issue 6425 • pp. 374-378 • DOI:10.1126/science.aau2706

nature communications

Explore content About the journal Publish with us

[nature](#) > [nature communications](#) > [articles](#) > [article](#)

Article | Open Access | Published: 02 January 2019

Influence of fake news in Twitter during the 2016 US presidential election

[Alexandre Bovet](#) & [Hernán A. Makse](#)

[Nature Communications](#) 10, Article number: 7 (2019) | [Cite this article](#)

85k Accesses | 167 Citations | 326 Altmetric | [Metrics](#)

nature communications

Explore content About the journal Publish with us

[nature](#) > [nature communications](#) > [articles](#) > [article](#)

Article | Open Access | Published: 22 September 2021

Neutral bots probe political bias on social media

[Wen Chen](#), [Diogo Pacheco](#), [Kai-Cheng Yang](#) & [Filippo Menczer](#)

[Nature Communications](#) 12, Article number: 5580 (2021) | [Cite this article](#)

7784 Accesses | 332 Altmetric | [Metrics](#)

Botometer supported journalistic reporting

FACT CHECK

Newsweek

Fact Check: Did Zelensky's Wife Go on \$40K Shopping Spree in Paris?

BY BRENDAN COLE ON 12/21/22 AT 1:47 PM EST

BUSINESS INSIDER

Pro-Russian Twitter bots heavily targeted South Africa after it didn't vote Moscow's way

Dave Chambers, Business Insider SA

11 Nov 2022



PODER 360

Extremistas se comunicavam pelo código “Festa da Selma”

Termo usado por bolsonaristas radicais alude a cumprimento militar; chamadas para invasão do Planalto, Congresso e STF correram em redes sociais

Elezioni, la battaglia social dietro le linee. Così i partiti guadagnano visibilità

Fratelli d'Italia la formazione che mostra un'attività maggiore di bot. Il caso dell'hashtag #Credo di Salvini e le mosse del Pd

QUOTIDIANONAZIONALE

Elon Musk cited **this tool** in his bot dispute with Twitter. Its **creator** has thoughts **Botometer** **Me**



By [Clare Duffy](#), CNN Business

Updated 4:59 PM ET, Tue August 9, 2022



MALVINA MITCHELL/GETTY IMAGES/FILE

New York (CNN Business) – Kaicheng Yang, a researcher at Indiana University's Observatory on Social Media, received quite the surprise last week. [Botometer](#), a tool he helped build to examine automated activity on Twitter, had been mentioned in court documents in the legal battle between Twitter and Elon Musk over their \$44 billion acquisition deal.

Botometer 101

Dataset/Software | [Published: 20 August 2022](#)

Botometer 101: social bot practicum for computational social scientists

[Kai-Cheng Yang](#) , [Emilio Ferrara](#) & [Filippo Menczer](#)

[*Journal of Computational Social Science*](#) **5**, 1511–1528 (2022) | [Cite this article](#)

How Botometer works

Default background?

Suspicious username?

Description length

Number of friends

Number of followers



And

Content

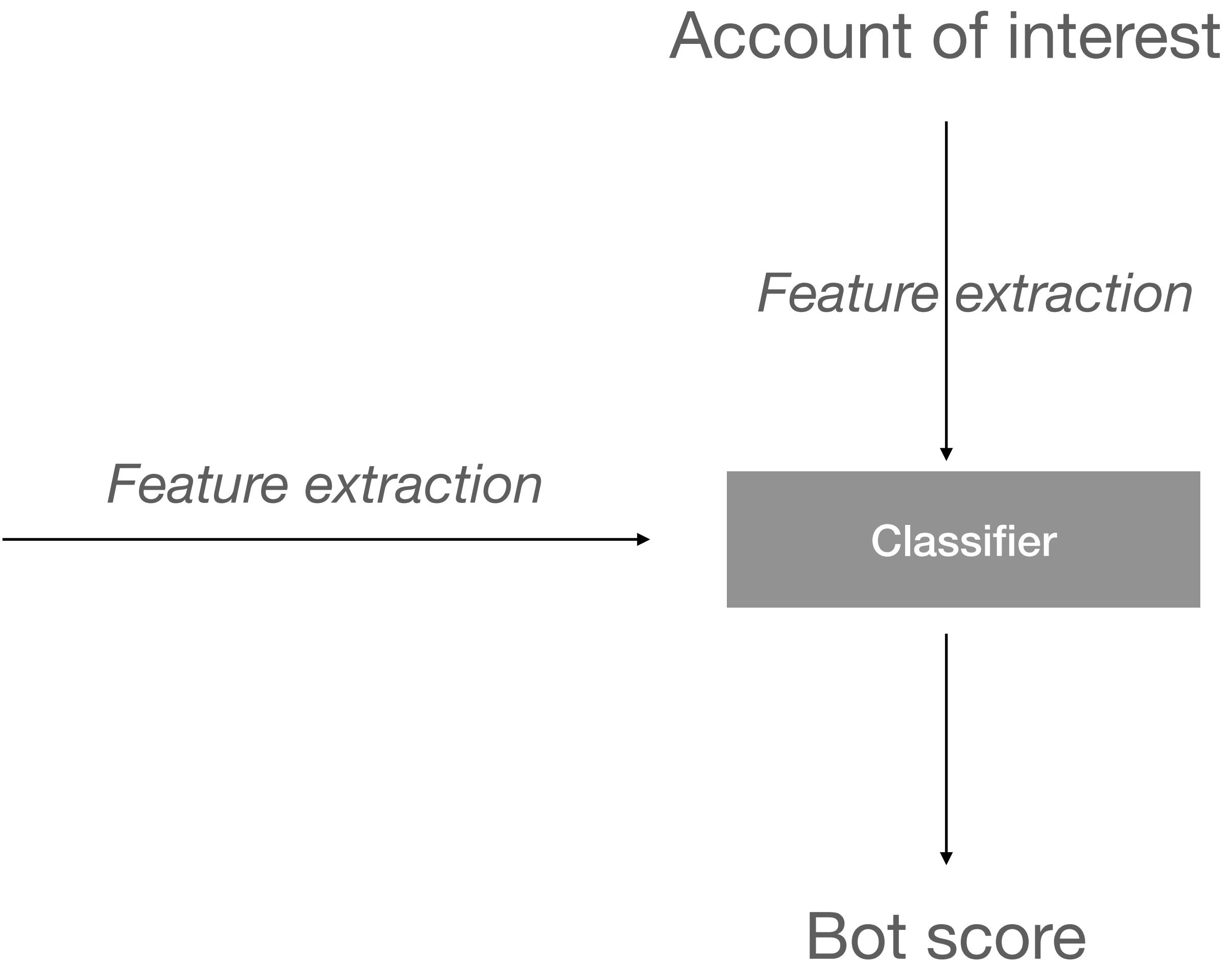
Temporal patterns

Social network

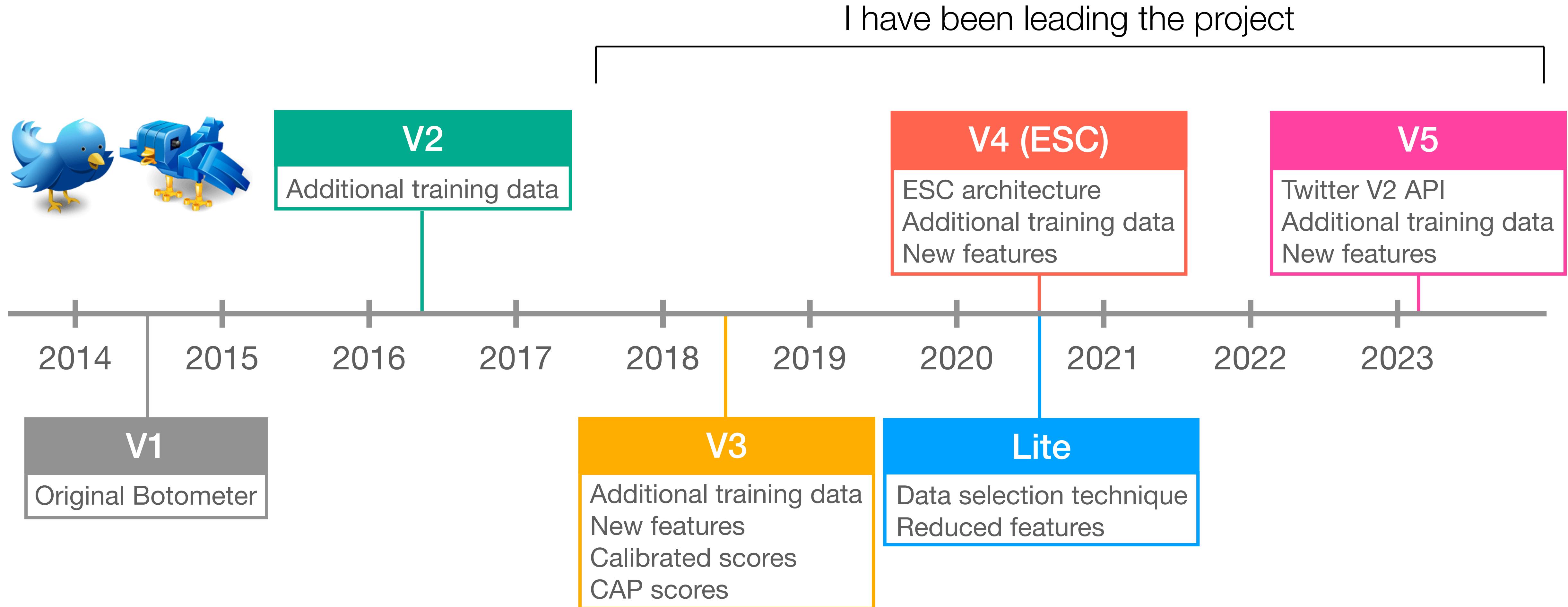
How Botometer works

Dataset	#bots	#human
caverlee	15,483	14,833
varol-icwsm	733	1,495
cresci-17	7,049	2,764
pronbots	17,882	0
celebrity	0	5,918
vendor-purchased	1,087	0
botometer-feedback	139	380
political-bots	62	0
gilani-17	1,090	1,413
cresci-rtbust	353	340
cresci-stock	7,102	6,174
midterm-18	42,446	8,092
botwiki	698	0
verified	0	1,987
Total	94,124	43,396

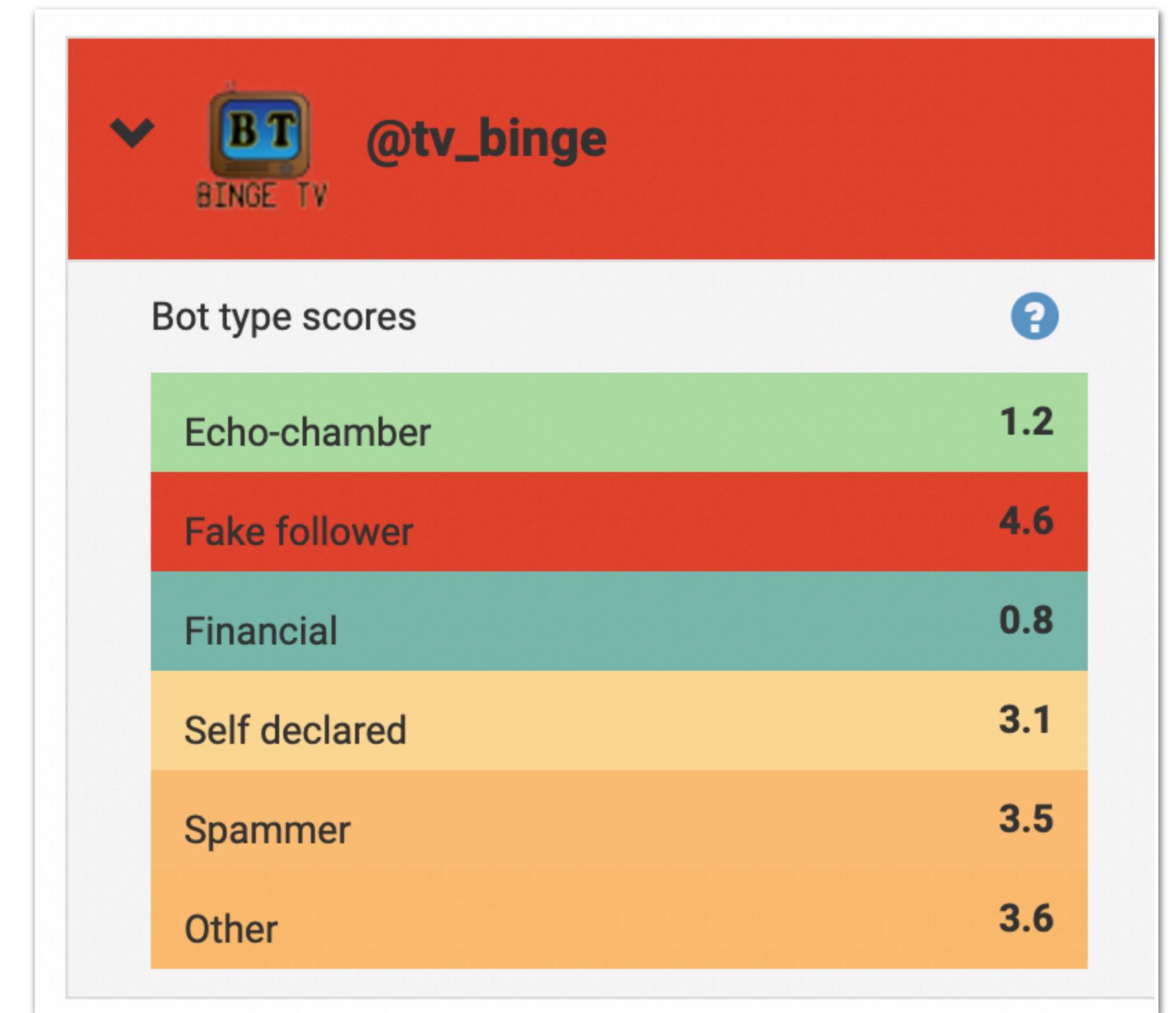
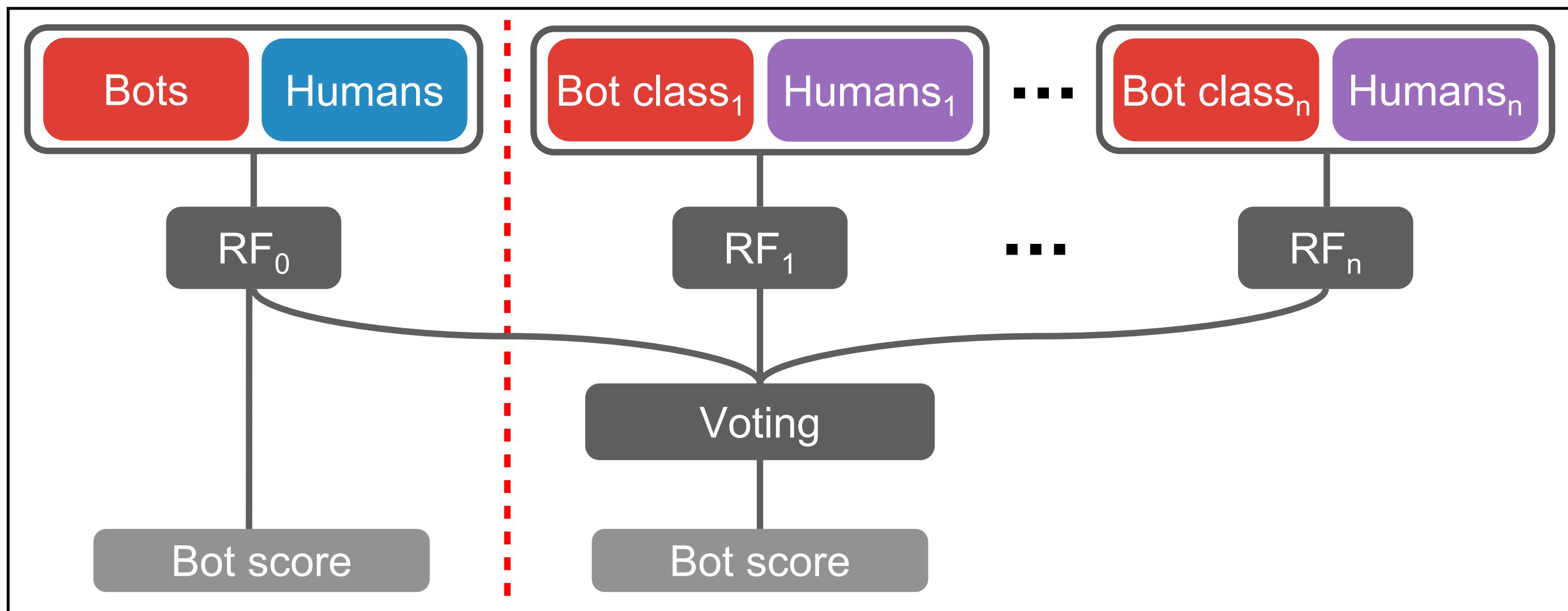
AUC=0.99



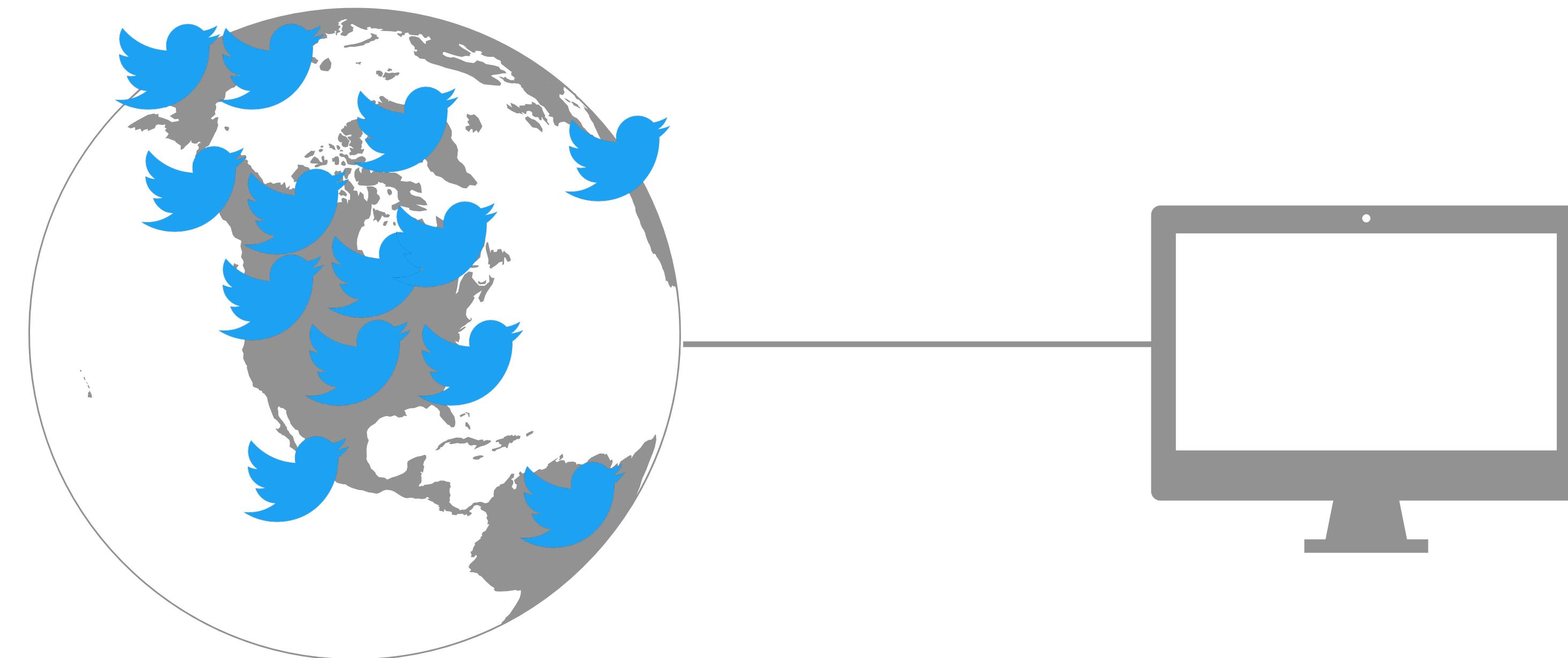
History of the project



V4: Adding explanations

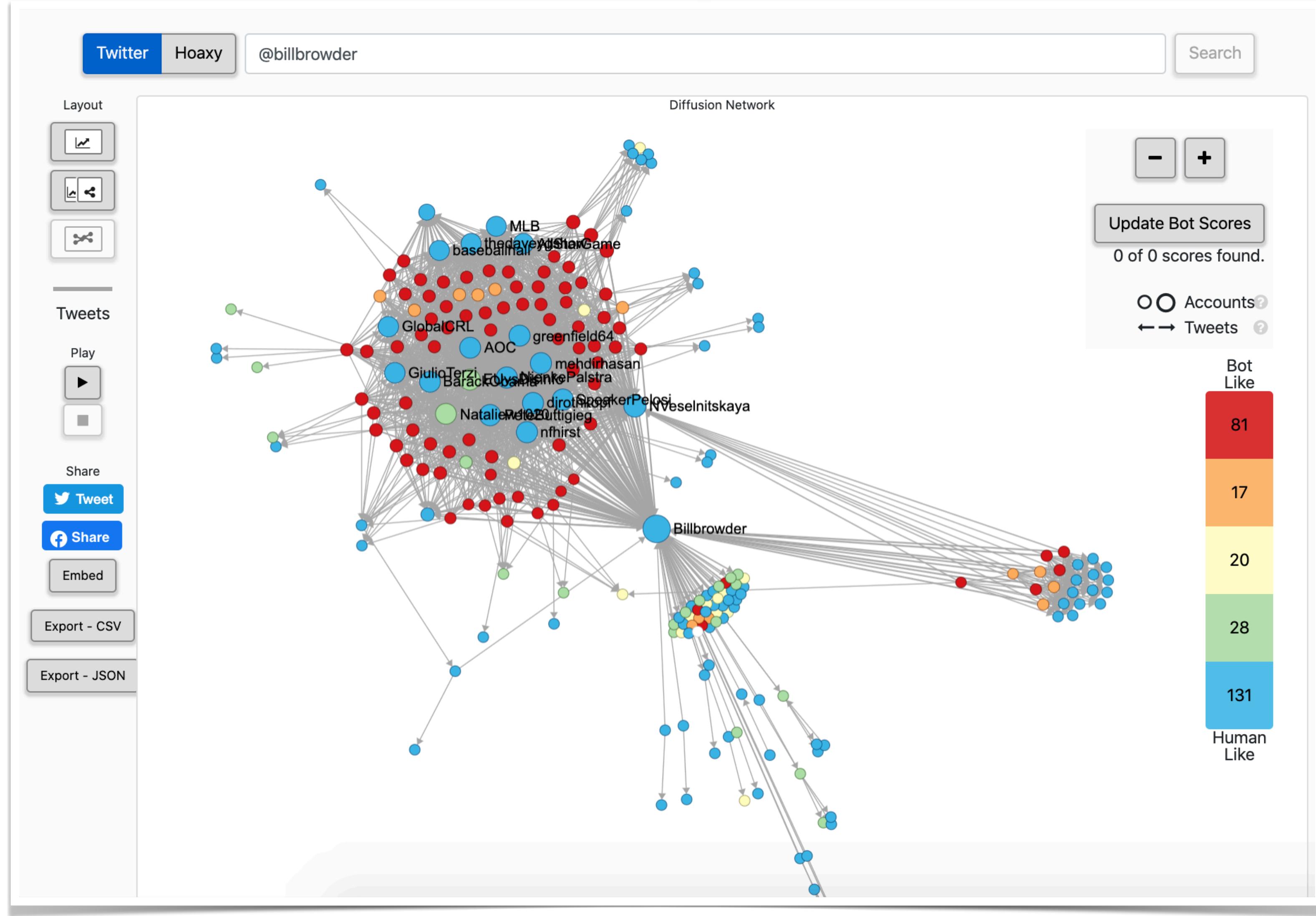


Lite: Evaluate millions of accounts within minutes



A desktop is sufficient for processing 900M tweets/day

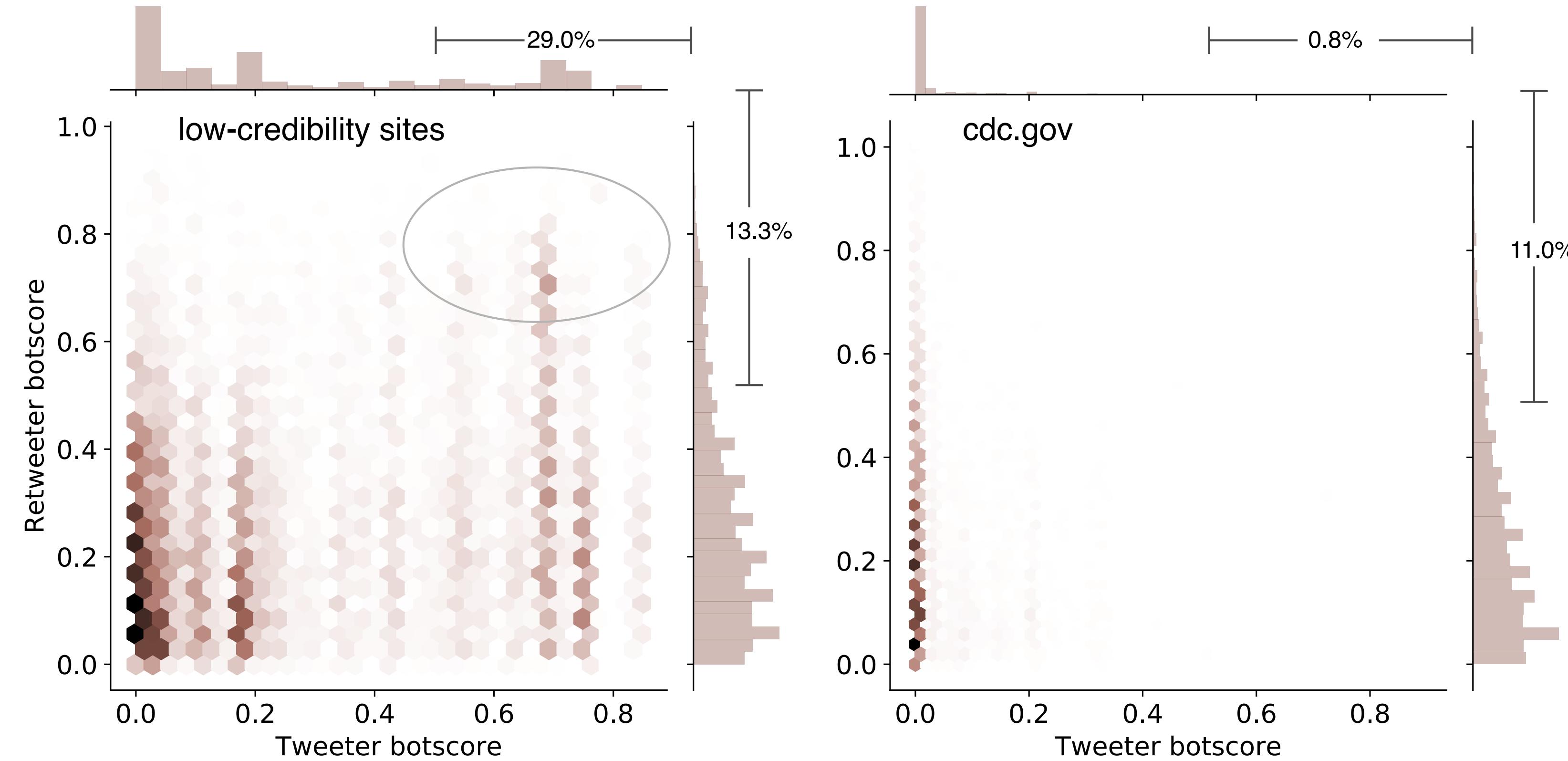
Capturing bots in action



A botnet pushing a Youtube video attacking American activist Bill Browder

hoaxy.osome.iu.edu

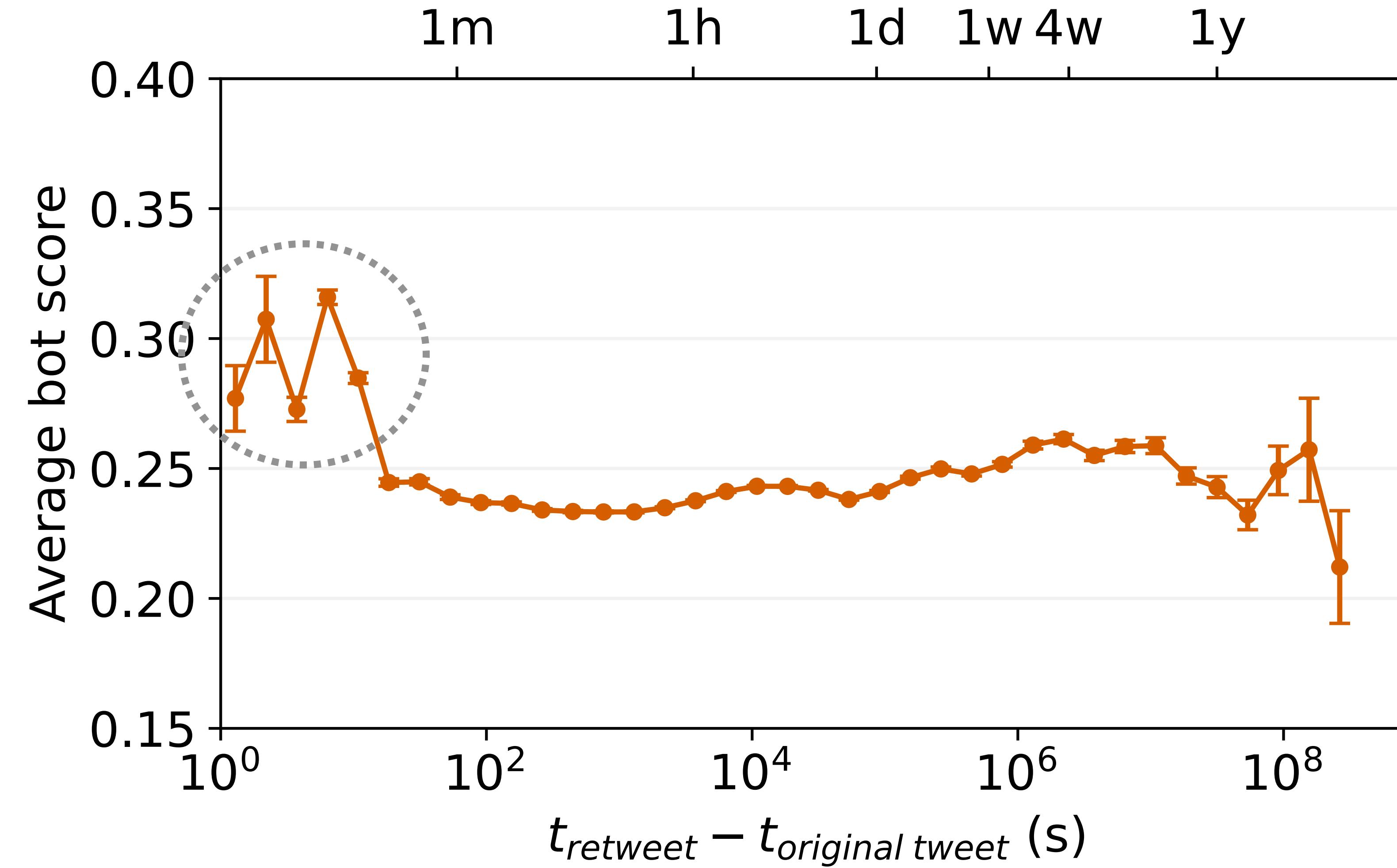
Bots amplify COVID-19 misinformation



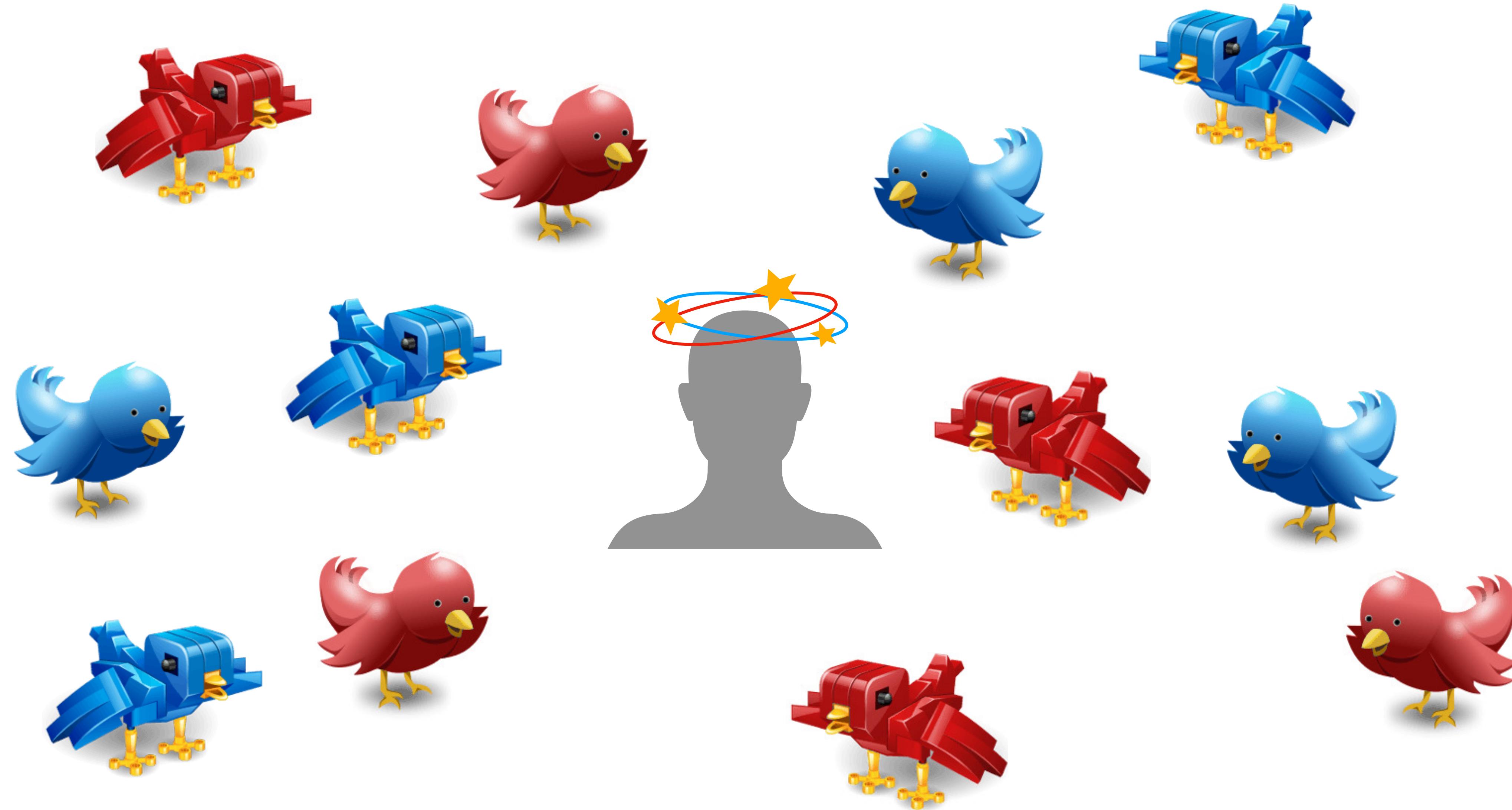
thegatewaypundit.com

childrenshealthdefense.org

Bots amplify COVID-19 misinformation



Human perception of bots



Experiment design



← **Carla** ❤️ #BLM 🌈
225.8K Tweets

WE THE PEOPLE
#wtpBLUE

@bright8694

...

Follow

Carla ❤️ #BLM 🌈
@bright8694

Vote like your life depends on it in local, state, and federal elections. It does! Resist! Stephen Colbert, Seth Meyers, & Trevor Noah keep me sane. wtpBLUE

📍 Atlanta, GA Joined December 2018

14.5K Following 17.5K Followers

Not followed by anyone you're following

Tweets Tweets & replies Media Likes

Pinned Tweet

Carla ❤️ #BLM 🌈 @bright8694 · Mar 3

I have not been an activist since the Nixon years. The fight against tyranny continues. It is Democracy vs. Fascism. Truth vs. lies.

We must fight back with love, truth, purpose, and in huge numbers. This is a fight we must win!

We are stronger together!

15 35 107

← **Janet Frank**
131.8K Tweets

...

Follow

Janet Frank
@JanetFr11616397

Joined December 2016

18.2K Following 17.9K Followers

Not followed by anyone you're following

Tweets Tweets & replies Media Likes

Janet Frank Retweeted

RetMilVet17 🐾 @mil_vet17 · 6h

Replying to @CoVet_81 @Scott4MAGA and 18 others

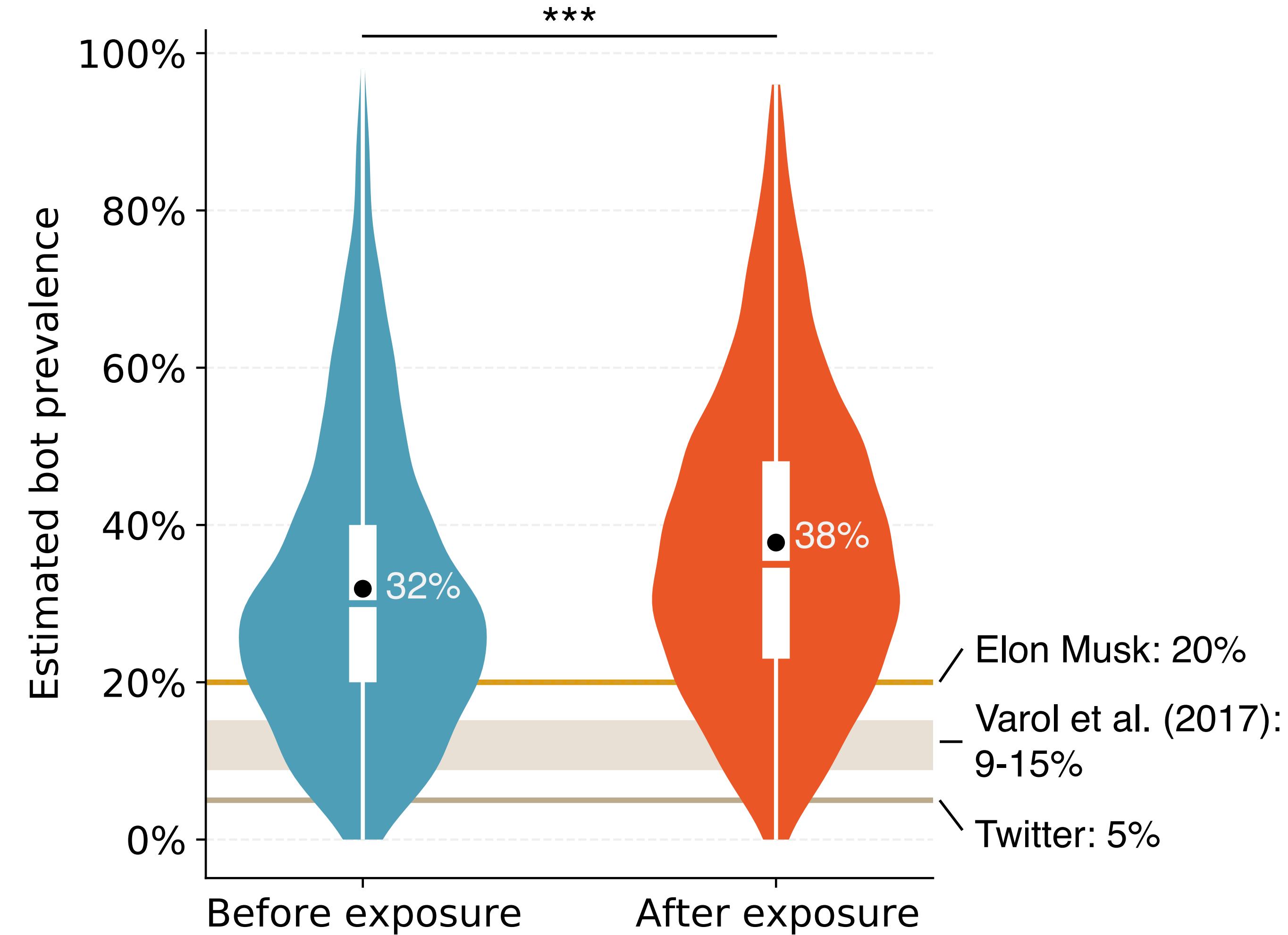
Thank you, Cheryl 😊

Always an honor my dear friend 🙏

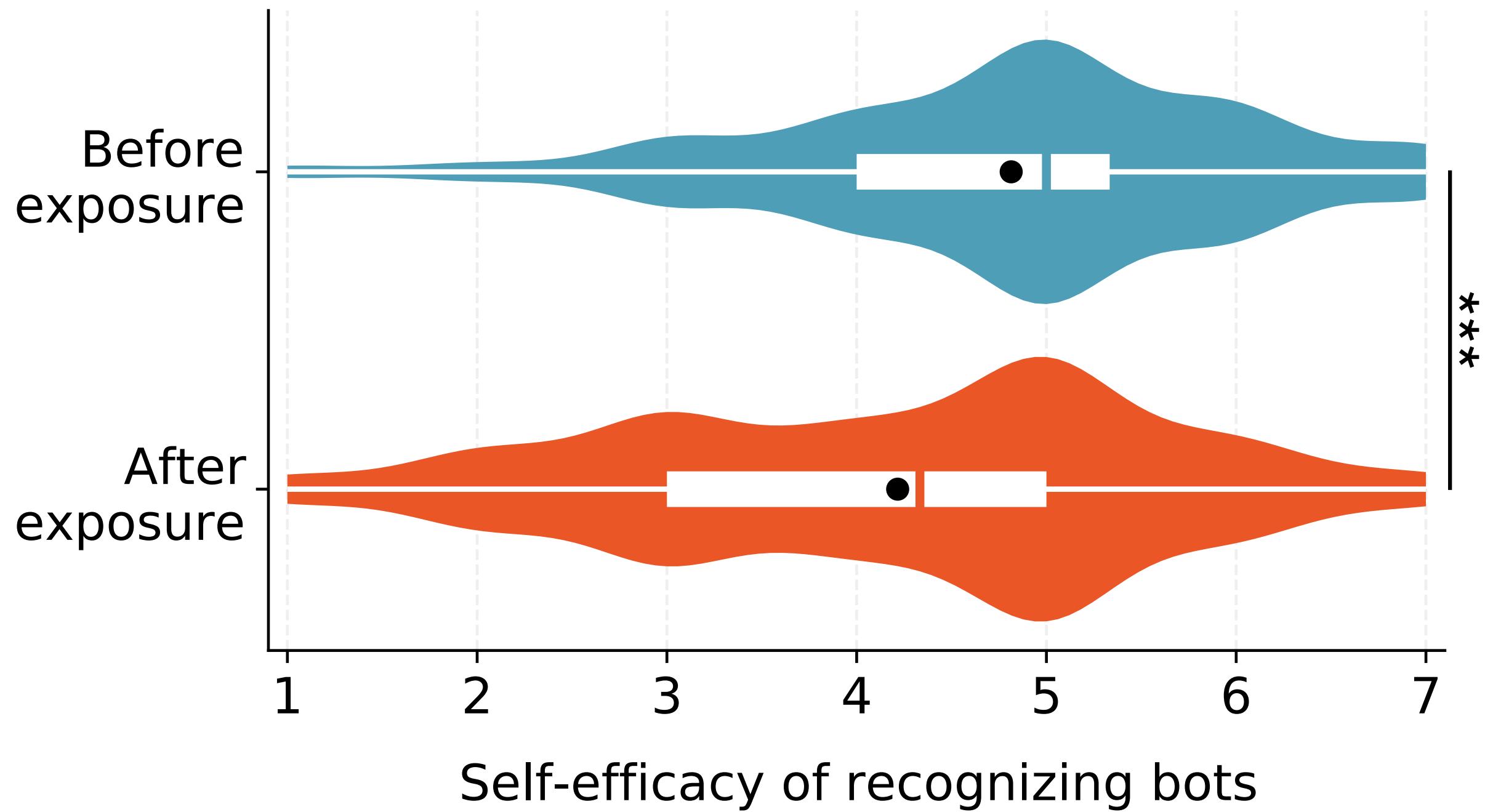
They can no longer hide the truth, so they think throwing Joe under the bus will make all go away. Too late, We the People are awake now 🤝🇺🇸

🇺🇸❤️💪 Please Follow👉 @CoVet_81 🤝❤️💪

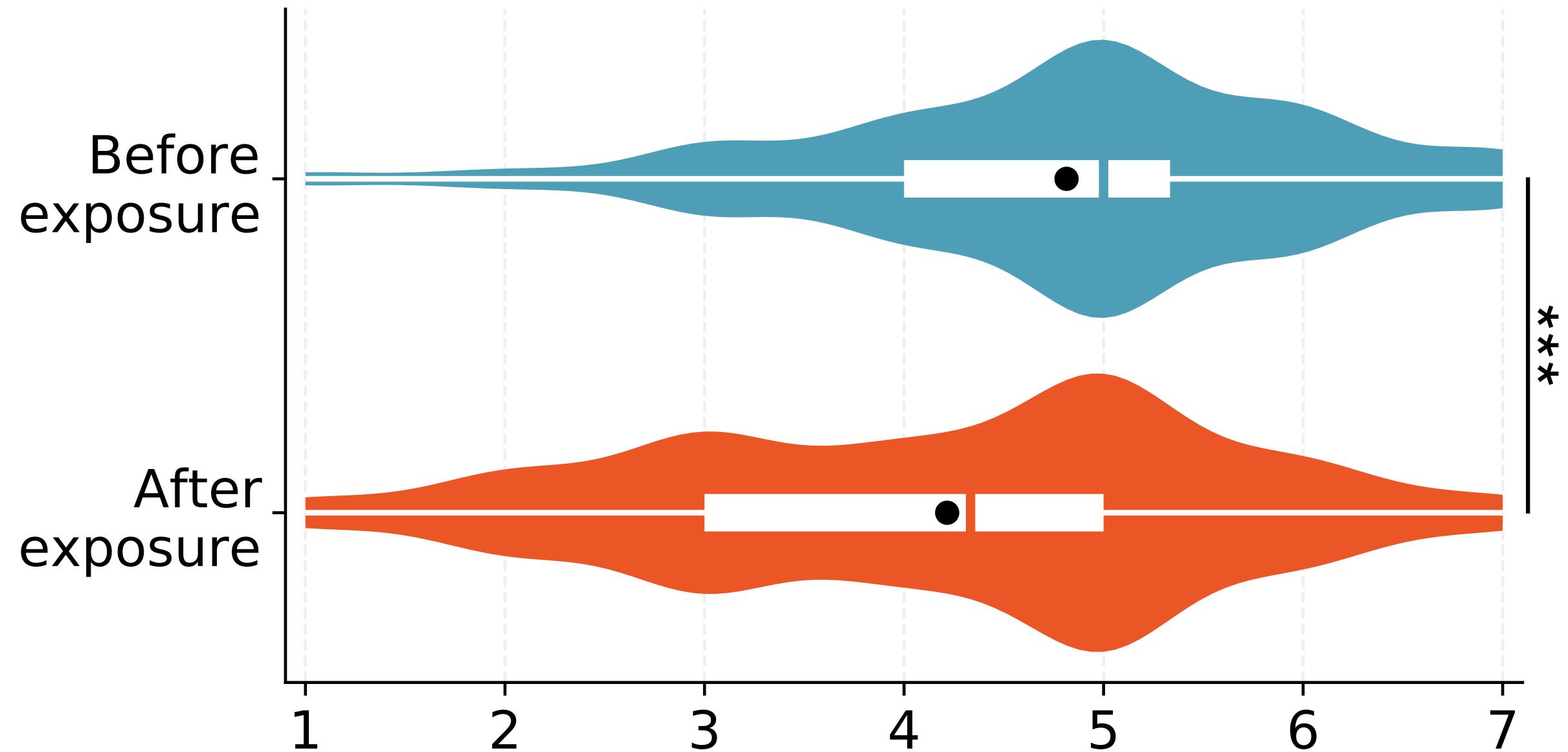
How many bots are there?



Do people think they can identify bots?

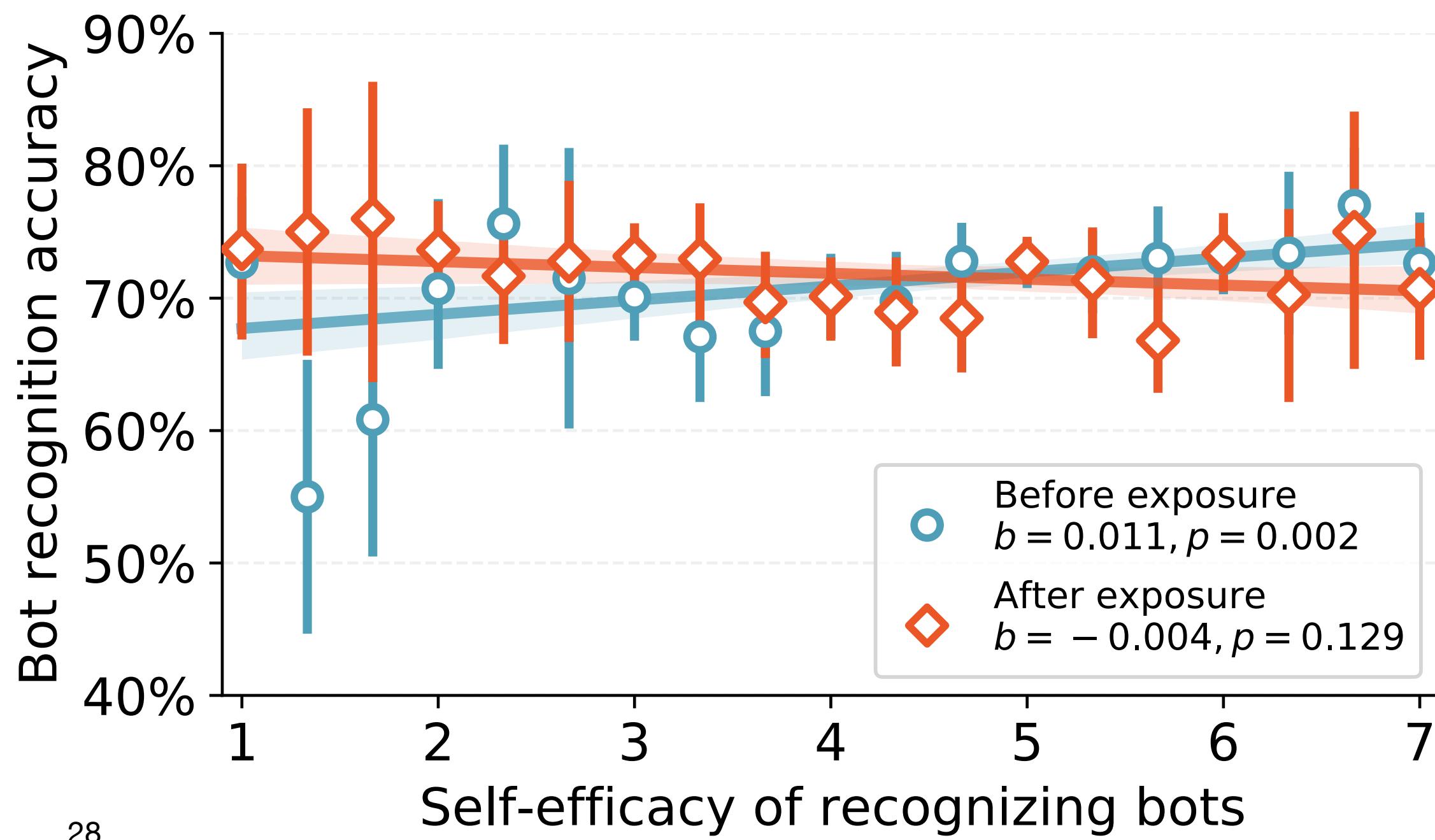


Do people think they can identify bots?

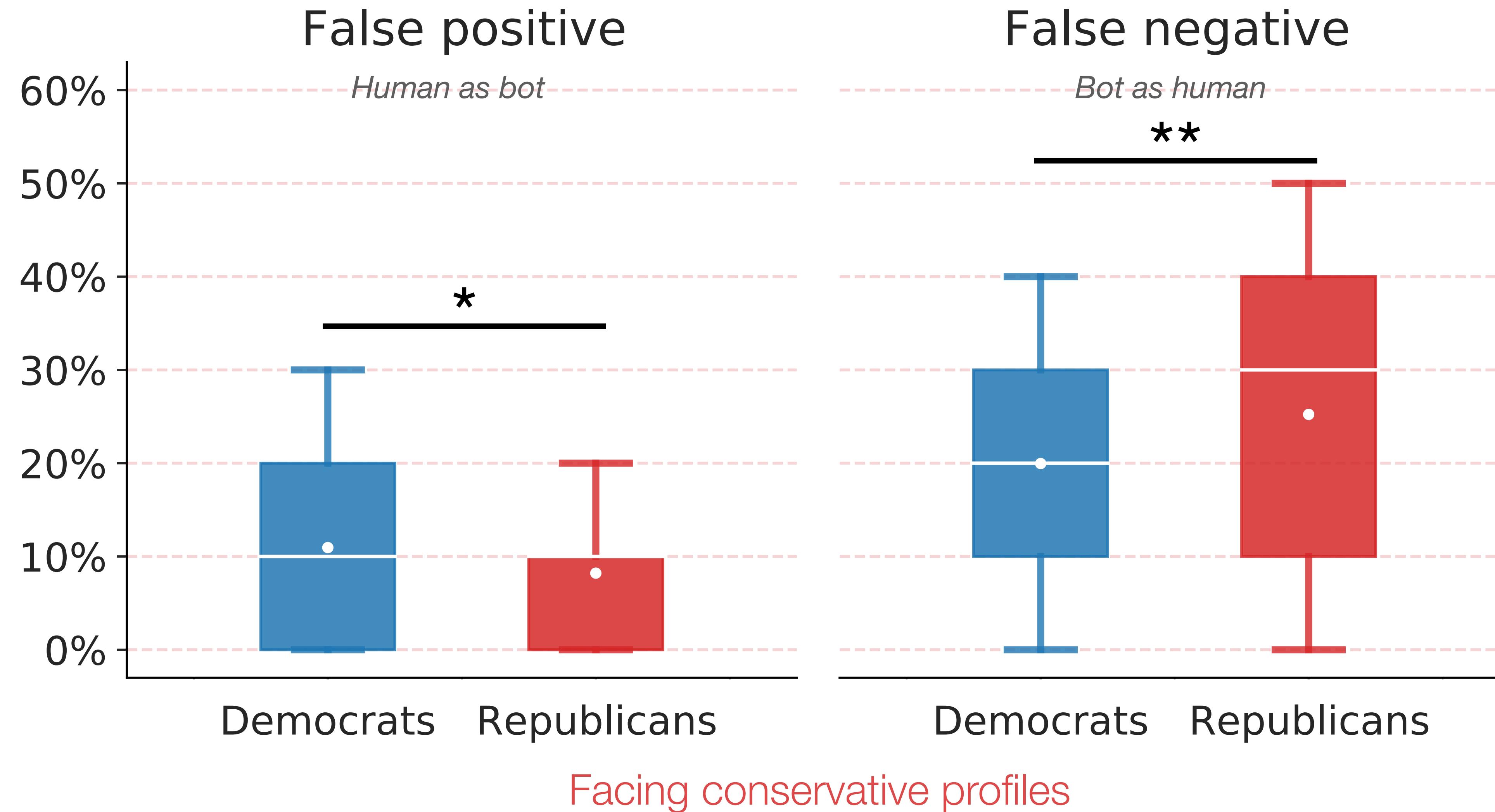


Can they?

In submission



Why did they make mistakes?



Bots as research instruments

Left/liberal

U.S. political spectrum

Right/conservative



Left

 **The Nation** 
@thenation

The place for debate on the left. Subscribe: bit.ly/2TSODfo
 Washington, DC  washingtonpost.com
 Joined March 2007

9,043 Following 1.2M Followers

Center Left

 **The Washington Post** 
@washingtonpost

Democracy Dies in Darkness.
 Washington, DC  washingtonpost.com
 Joined March 2007

1,716 Following 18.3M Followers

Center

 **USA TODAY** 
@USATODAY

The pulse of the nation in the palm of your hand.
 USA TODAY HQ, McLean, Va.  usatoday.com
 Joined August 2008

901 Following 4.4M Followers

Center Right

 **The Wall Street Journal** 
@WSJ

Sign up for our newsletters and email alerts:
wsj.com/newsletters
 New York, NY  wsj.com
 Joined April 2007

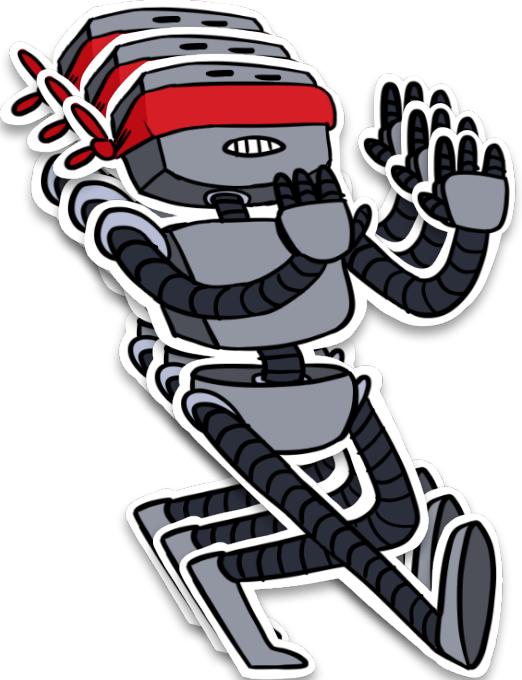
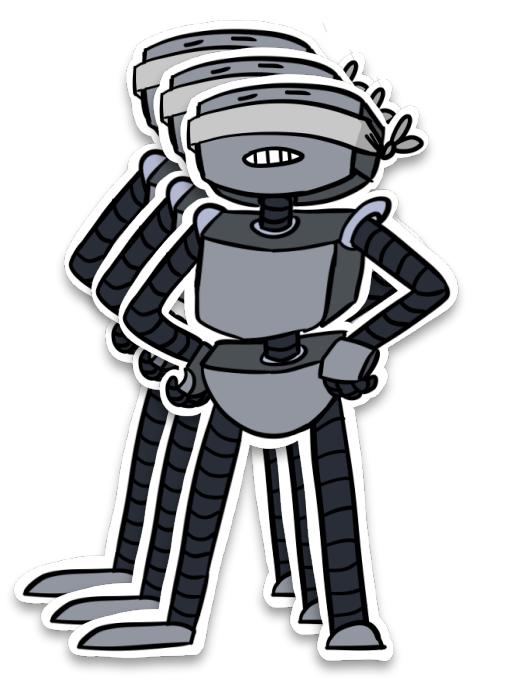
1,045 Following 19.1M Followers

Right

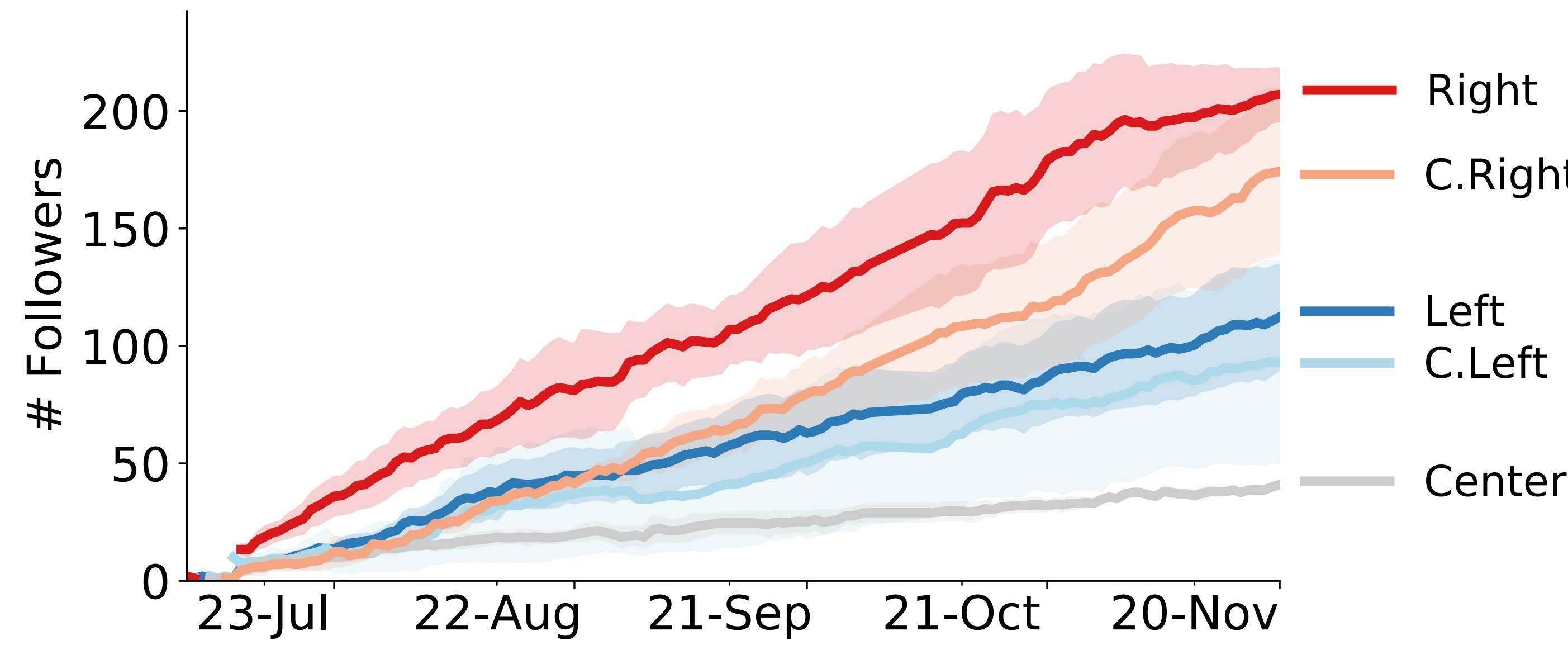
 **Breitbart News** 
@BreitbartNews

News, commentary, and destruction of the political/media establishment.
 Breitbart.com
 Joined January 2012

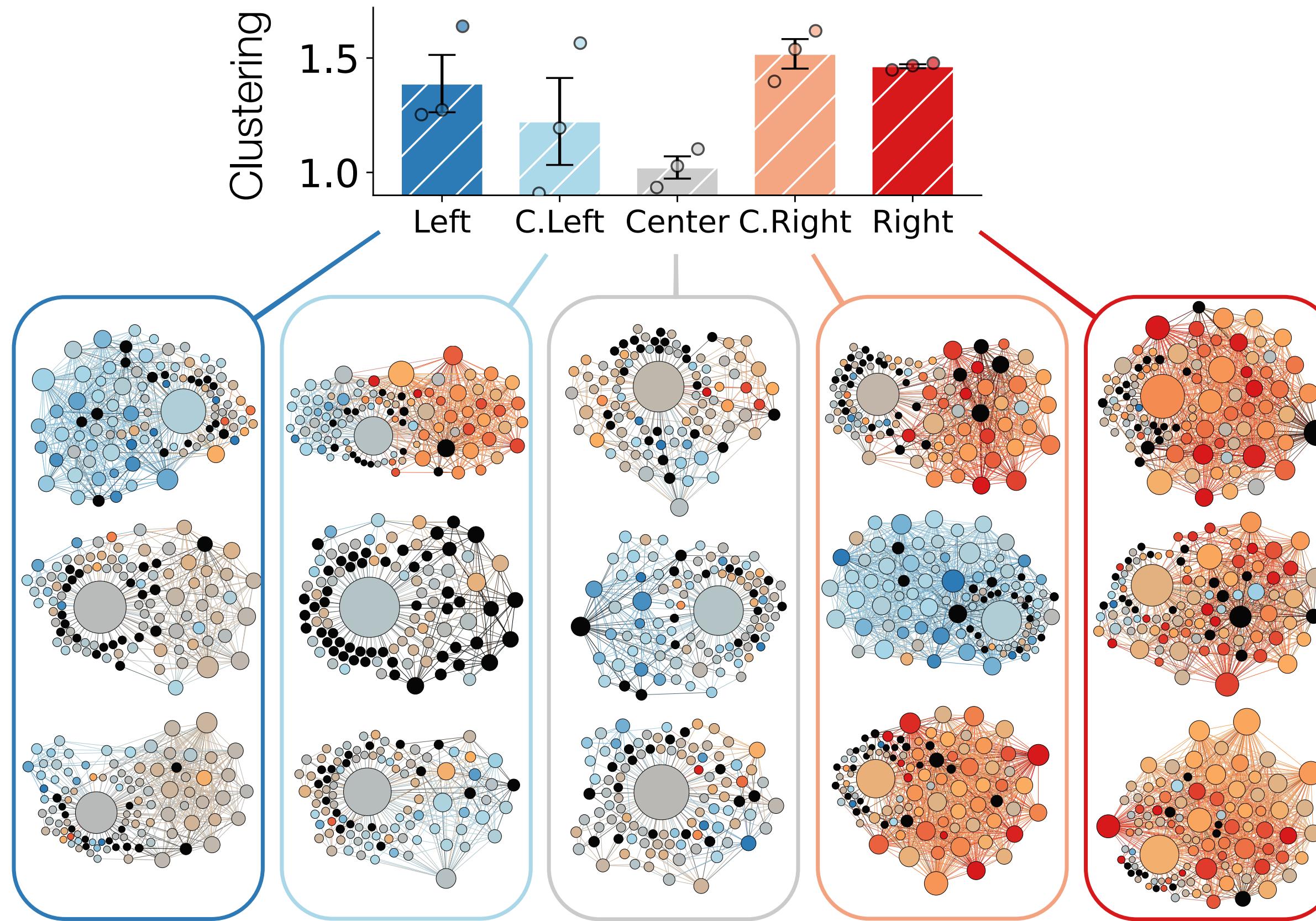
111 Following 1.4M Followers



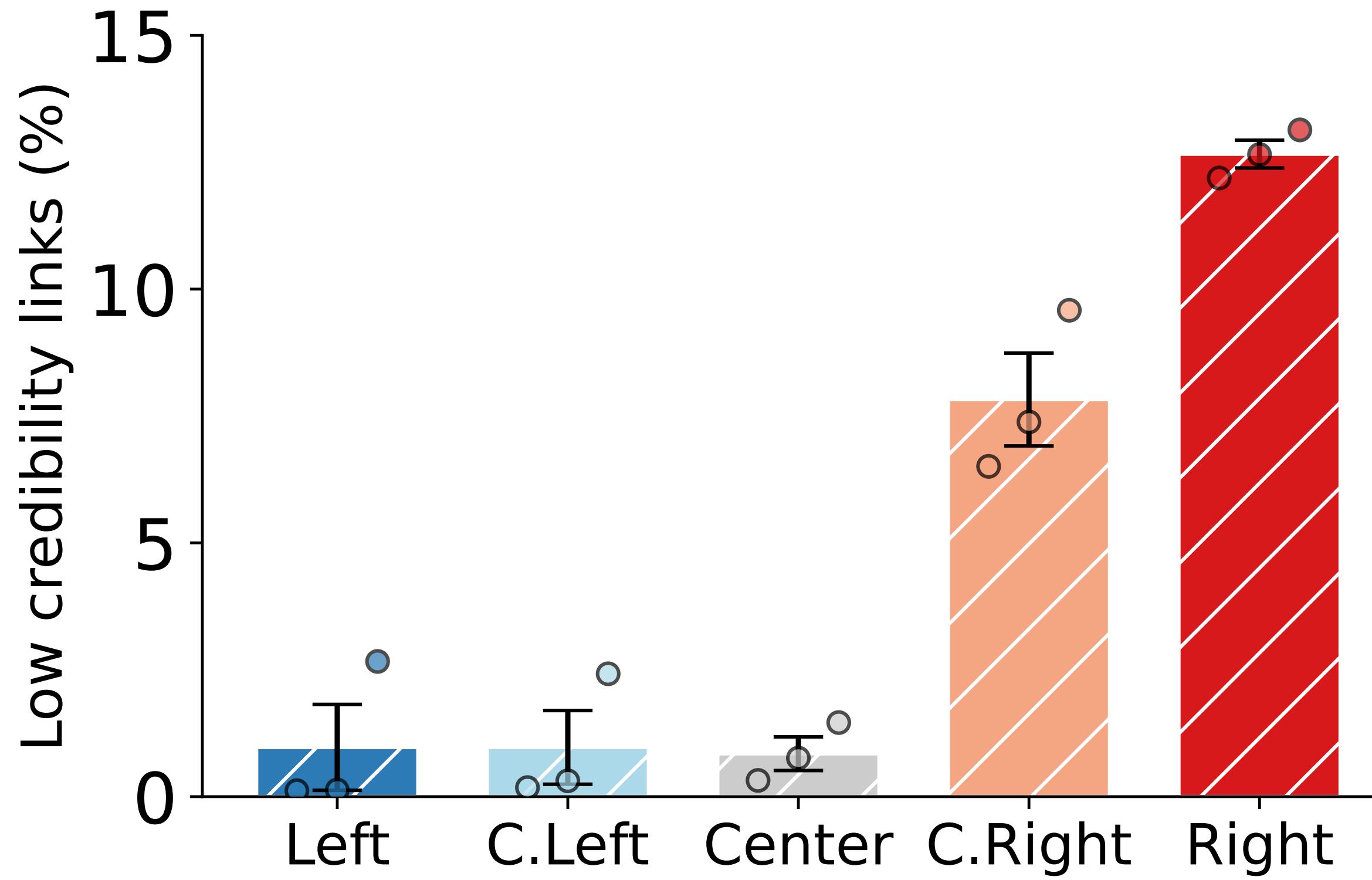
Follower growth



Political echo chamber



Exposure to misinformation



What's next?

 <p>Ernesto Gilmore @ErnestoGilmore1 One day, I'm gonna make the onions cry. Joined April 2022 7 Following 0 Followers</p>	 <p>Liam Ho @LiamHo18 One day, I'm gonna make the onions cry. Joined April 2022 12 Following 0 Followers</p>	 <p>Kyle Coffey @KyleCof88349162 One day, I'm gonna make the onions cry. Joined April 2022 21 Following 0 Followers</p>
 <p>Helen Lang @HelenLa71170902 One day, I'm gonna make the onions cry. Joined April 2022 21 Following 0 Followers</p>	 <p>Anahi Jefferson @JeffersonAnahi One day, I'm gonna make the onions cry. Joined April 2022 21 Following 0 Followers</p>	 <p>Noelle Good @NoelleG75485714 One day, I'm gonna make the onions cry. Joined April 2022 16 Following 0 Followers</p>

Profile pictures generated by GAN (Generative Adversarial Network)

Source: @conspirator0



Write a Twitter profile of a house wife in Nebraska



Here is a potential Twitter profile for a housewife in Nebraska:



Name: Sarah Johnson

Location: Omaha, Nebraska

Bio: Midwestern housewife and mom of two. Passionate about cooking, gardening, and spending time with my family.

Tweets:

- Just finished making a delicious apple pie from scratch. Yum!
- Took the kids to the park today and had so much fun playing on the swings.
- Going to try out a new recipe for dinner tonight - excited to see how it turns out!
- Can't believe how fast the summer is flying by. Enjoying every minute of it though.
- Just signed up for a new gardening class - can't wait to learn some new techniques!
- So proud of my kids for their hard work in school. They are growing up so fast!

Feel free to modify and personalize this Twitter profile to better fit your needs.

Source: ChatGPT

Takeaways

- Threats of bots are real, but human perception is biased
- It requires innovative approaches and human factor is crucial
- Think about constructive usage of bots
- Fight against malicious actors is a never-ending arms race

Research interests

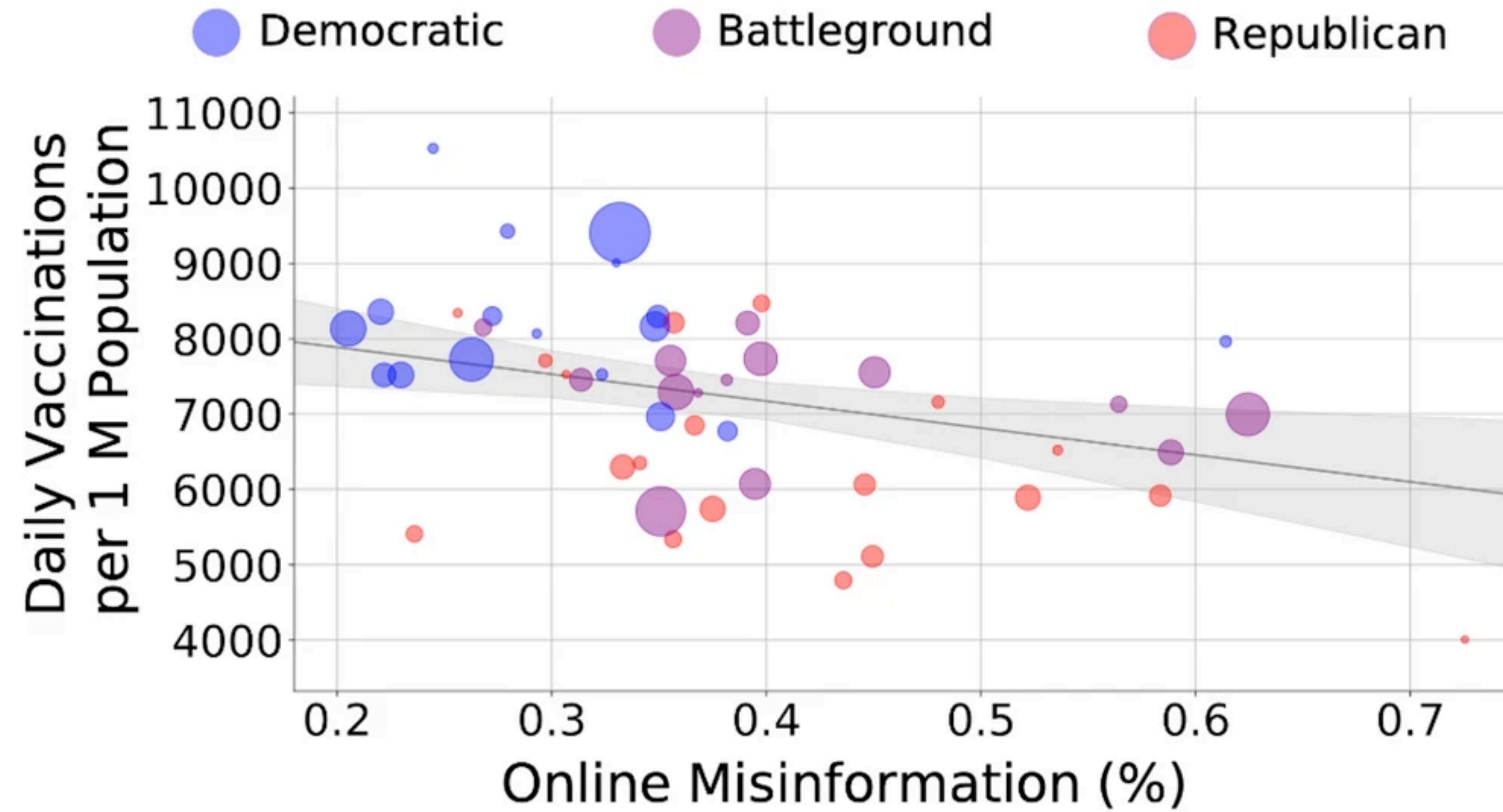
Social media integrity

- Bad actors such as social bots (*AAAI, CIKM, New Media & Soc., Hum. Commun. Res.*)
- Spread of misinformation (*Big Data & Soc., Nat. Comms., ICWSM*)
- Platform bias (*Nat. Comms.*)

Public health

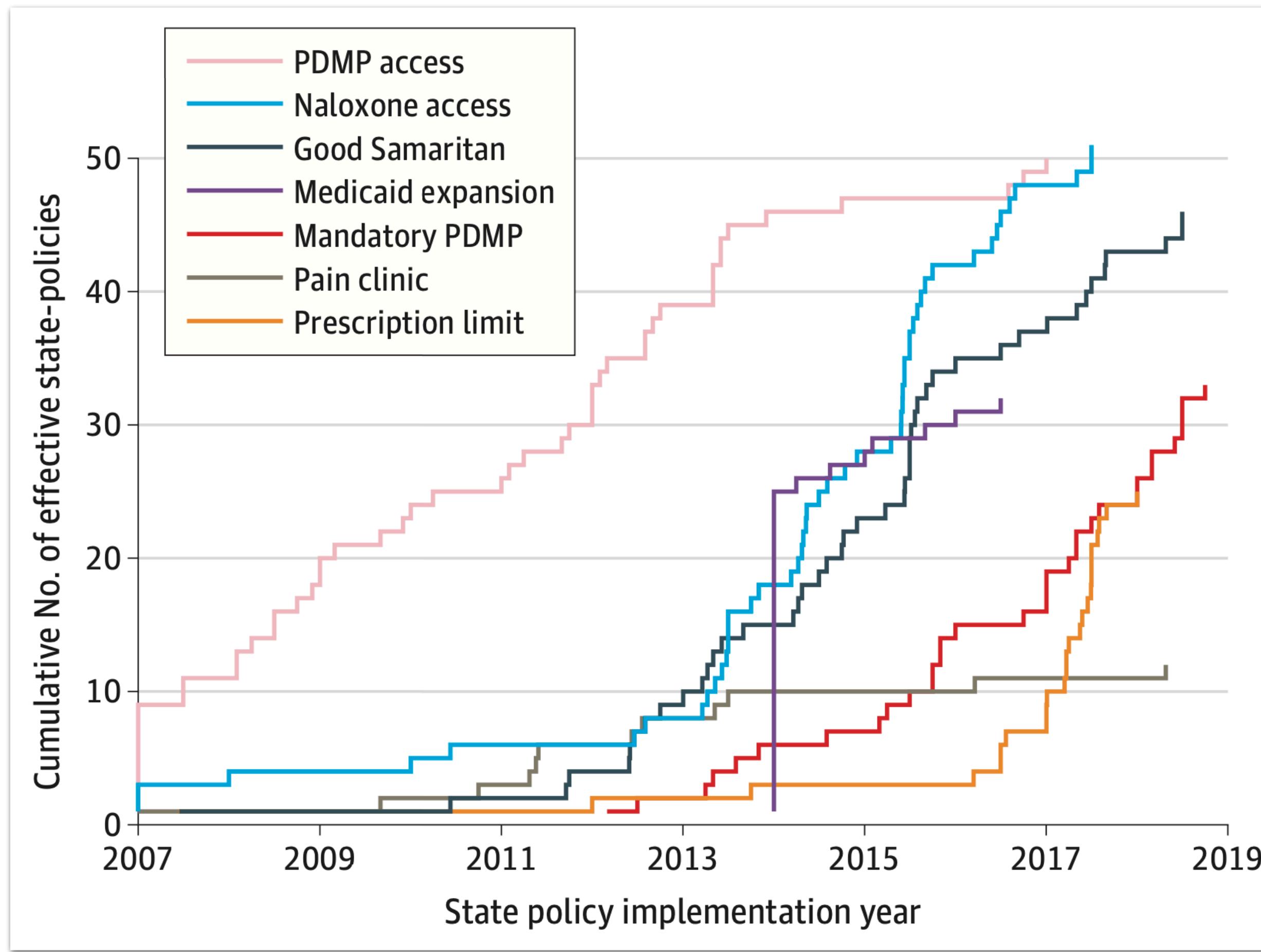
- COVID-19 misinformation and health outcomes (*Big Data & Soc., ICWSM, Sci. Rep.*)
- U.S. opioid crisis
 - Evaluation of intervention policies (*JAMA Net. Open*)
 - High-risk patients and prescribers (*Addiction, PLoS ONE*)

Online vaccine misinformation is linked to COVID-19 vaccination refusal

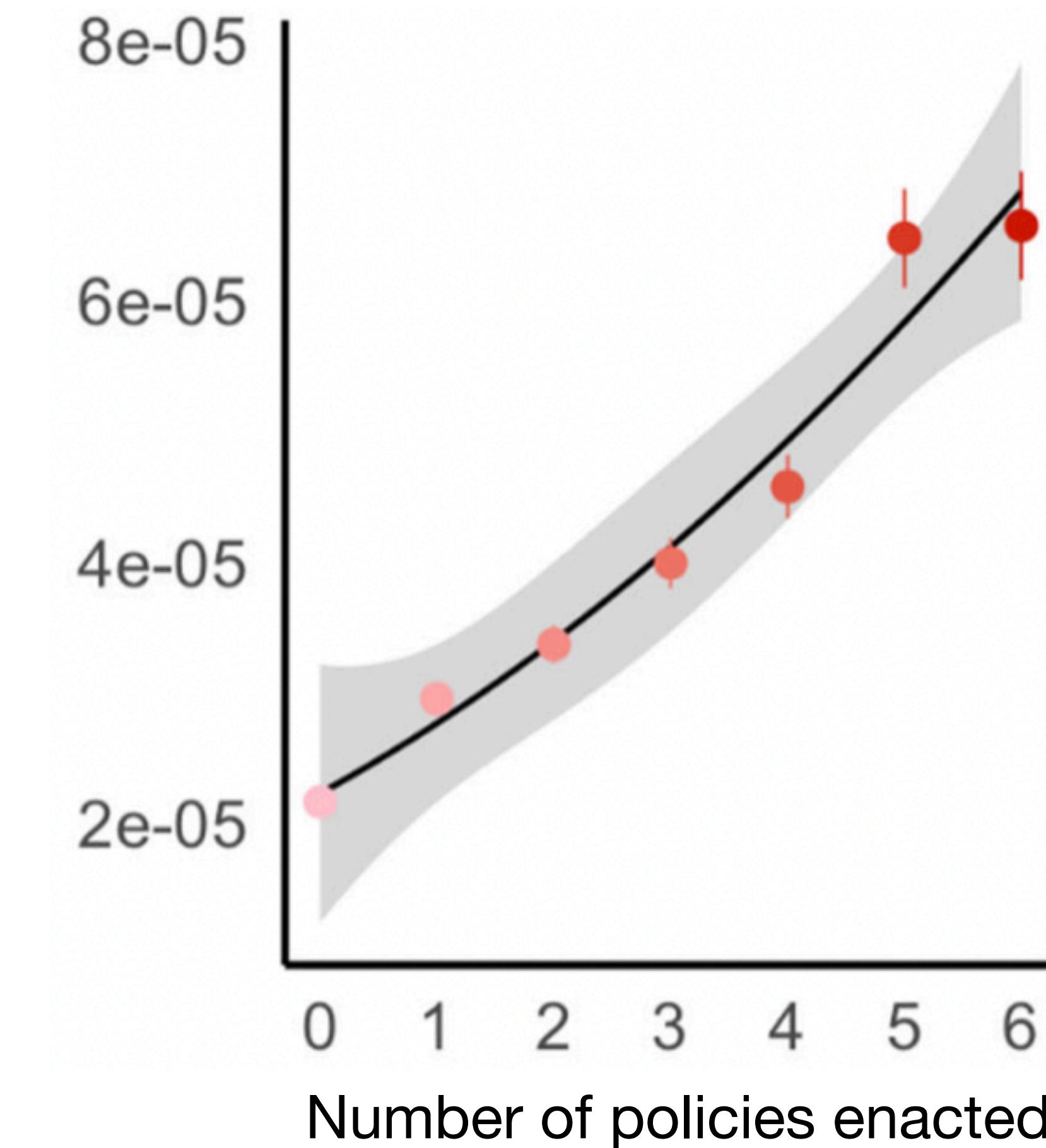


State policy interventions targeting the US opioid epidemic have unintended consequences

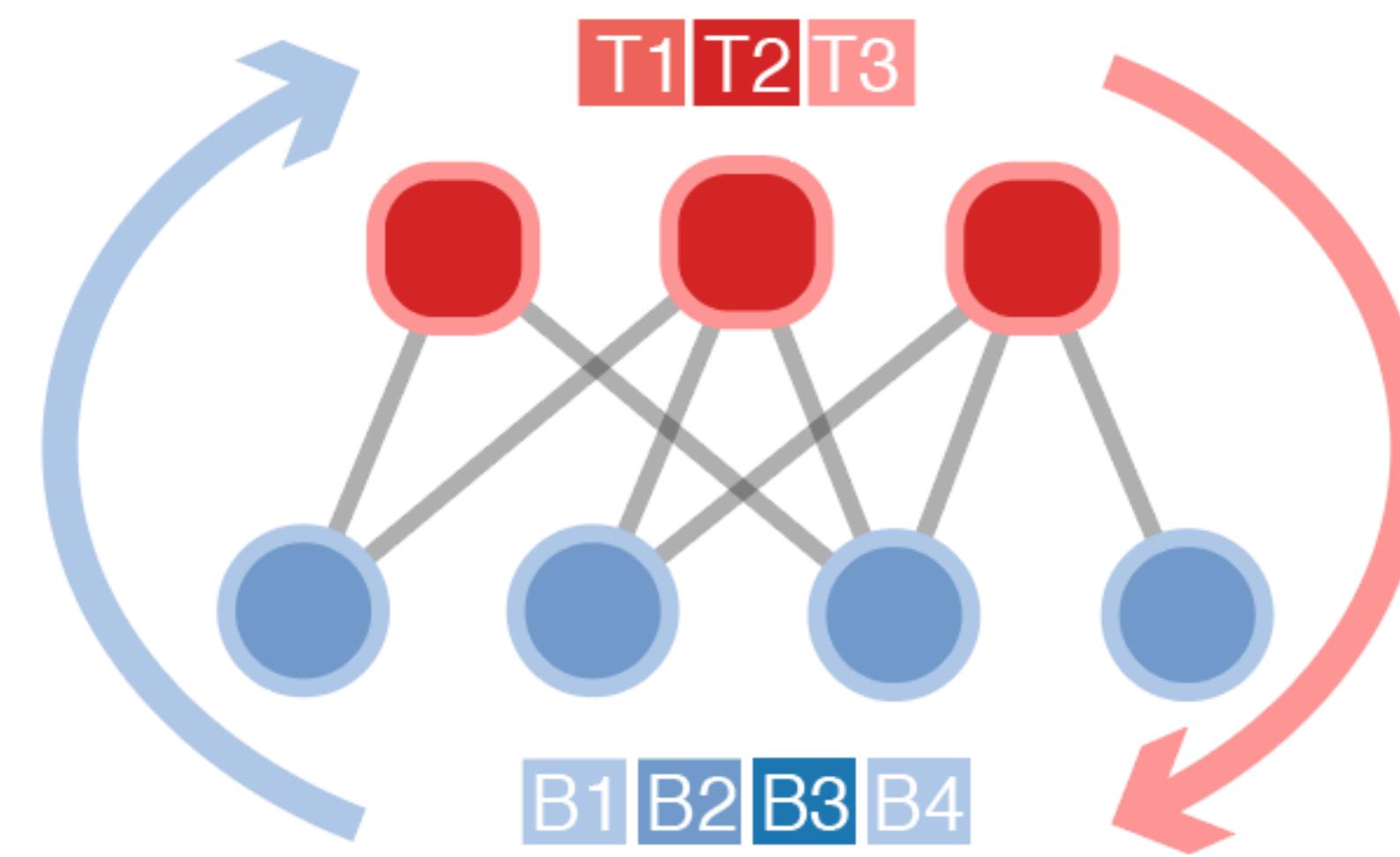
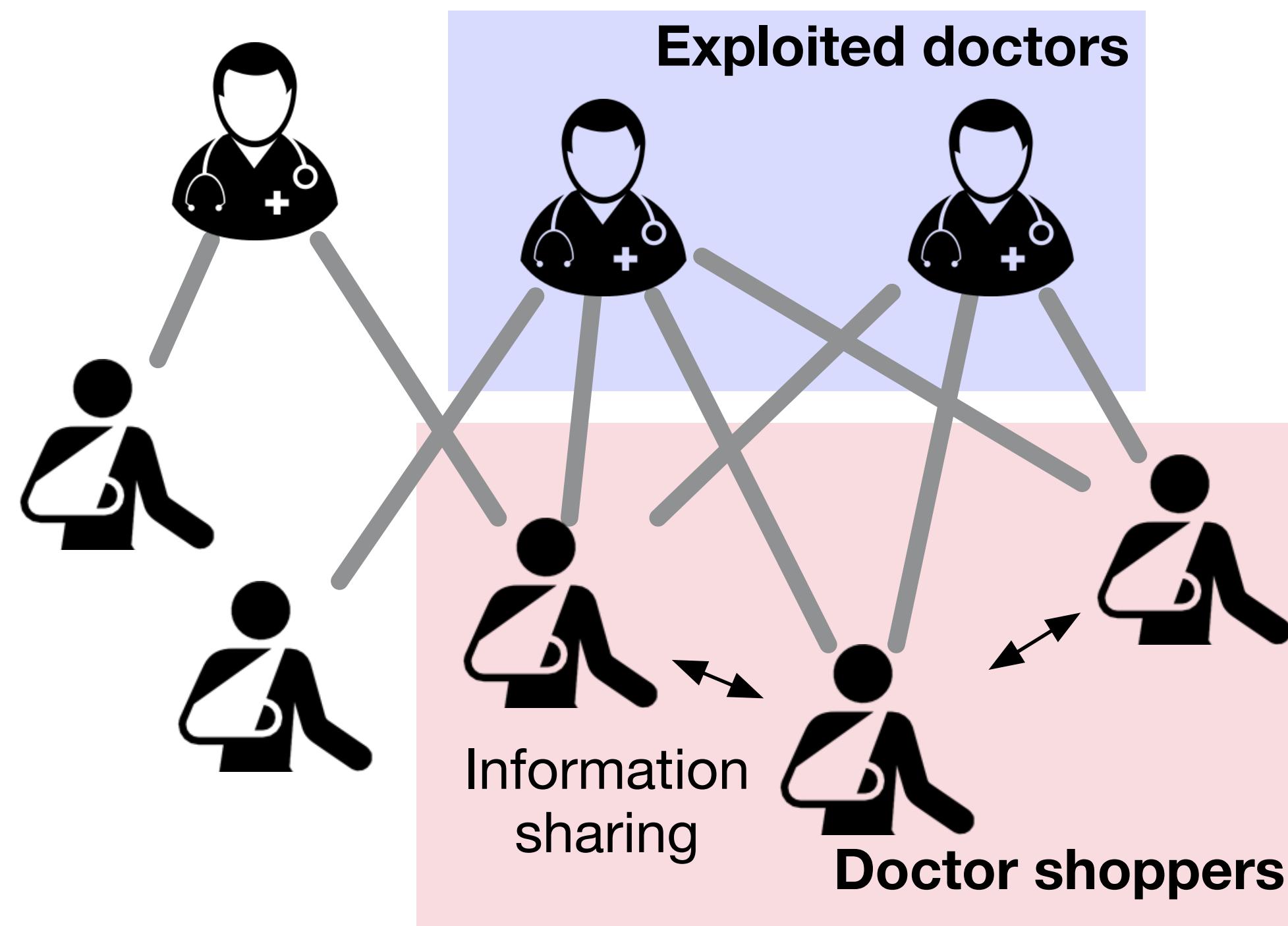
State policies



Overdose deaths per million



Using social network analysis to identify high-risk patients



Thank you!

Advisors



Filippo Menczer



Yong-Yeol Ahn



Brea Perry



Alessandro Flammini

Contact

yangkc@iu.edu

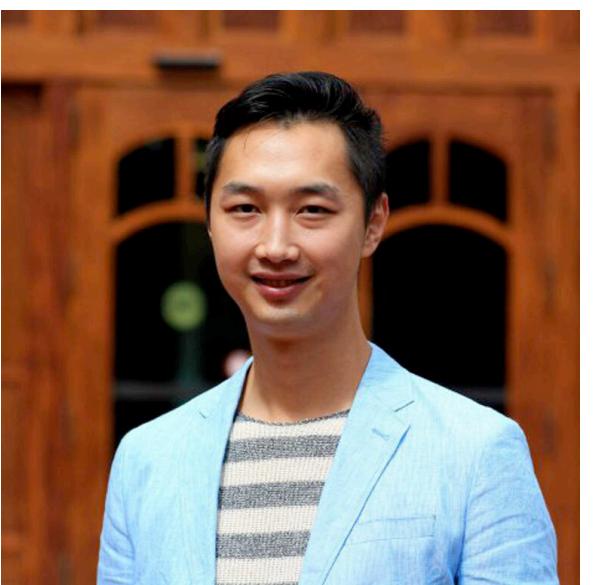
www.kaichengyang.me

@yang3kc

Collaborators



Onur Varol



Harry Yan



Christopher
Torres-Lugo



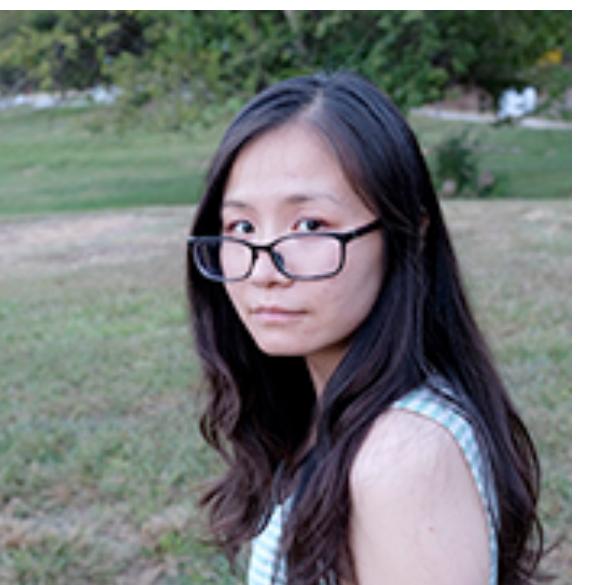
Mohsen Sayyadi



Francesco Pierri

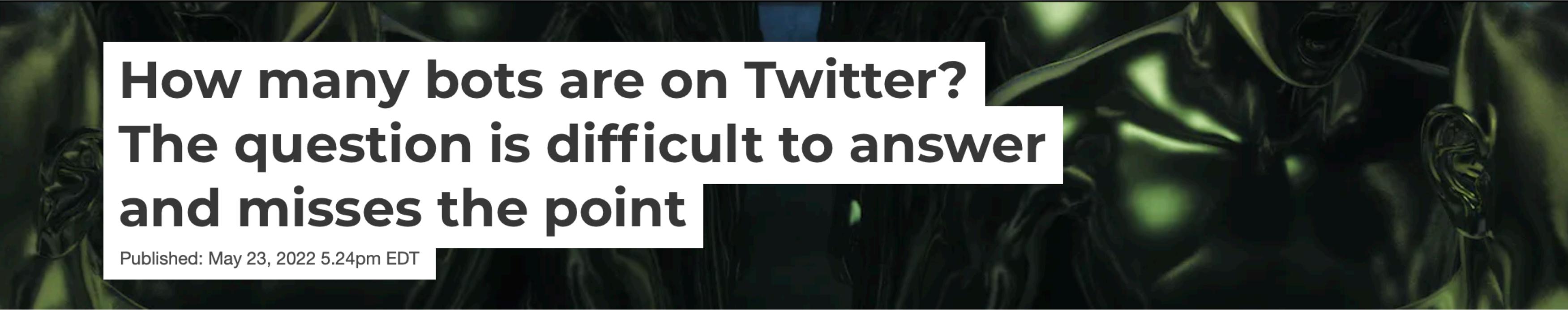


Diogo Pacheco



Wen Chen

Appendix



How many bots are on Twitter? The question is difficult to answer and misses the point

Published: May 23, 2022 5.24pm EDT

Yes, worry about Twitter, but don't worry whether there are hordes of spambots running rampant there. gremlin/E+ via Getty Images

[Email](#) [Twitter](#) 73 [Facebook](#) 119 [LinkedIn](#) [Print](#)

Twitter reports that fewer than 5% of accounts are fakes or spammers, commonly referred to as “bots.” Since his offer to buy Twitter was accepted, Elon Musk has repeatedly questioned these estimates, even dismissing Chief Executive Officer Parag Agrawal’s public response.

Later, Musk put the deal on hold and demanded more proof.

So why are people arguing about the percentage of bot accounts on Twitter?

As the creators of [Botometer](#), a widely used bot detection tool, our group at the Indiana University [Observatory on Social Media](#) has been studying inauthentic accounts and manipulation on social media for over a decade. We brought the concept of the “social bot” to the foreground and [first estimated their prevalence on Twitter in 2017](#).

Authors

 **Kai-Cheng Yang**
Doctoral Student in Informatics,
Indiana University

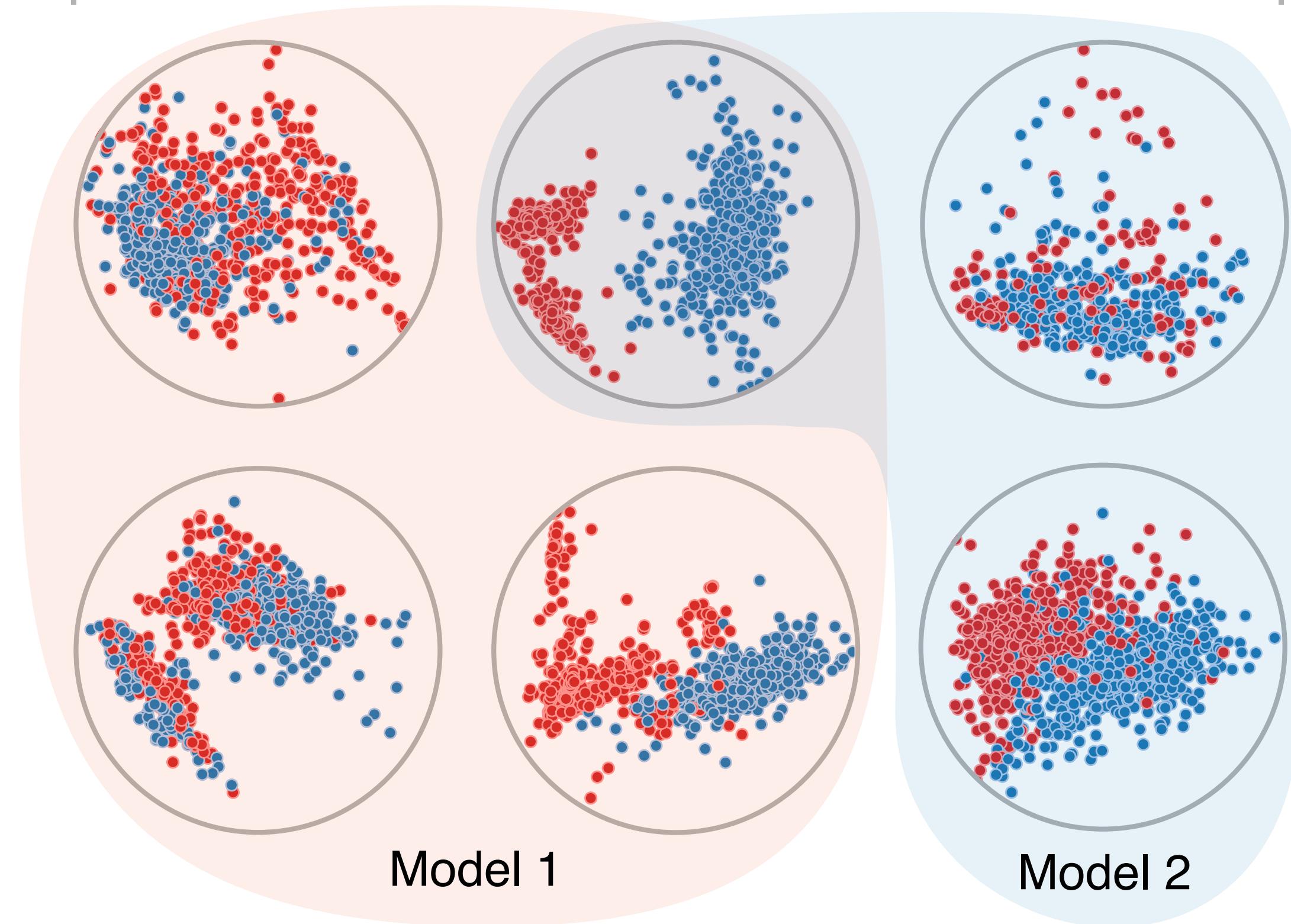
 **Filippo Menczer**
Professor of Informatics and
Computer Science, Indiana
University

Disclosure statement

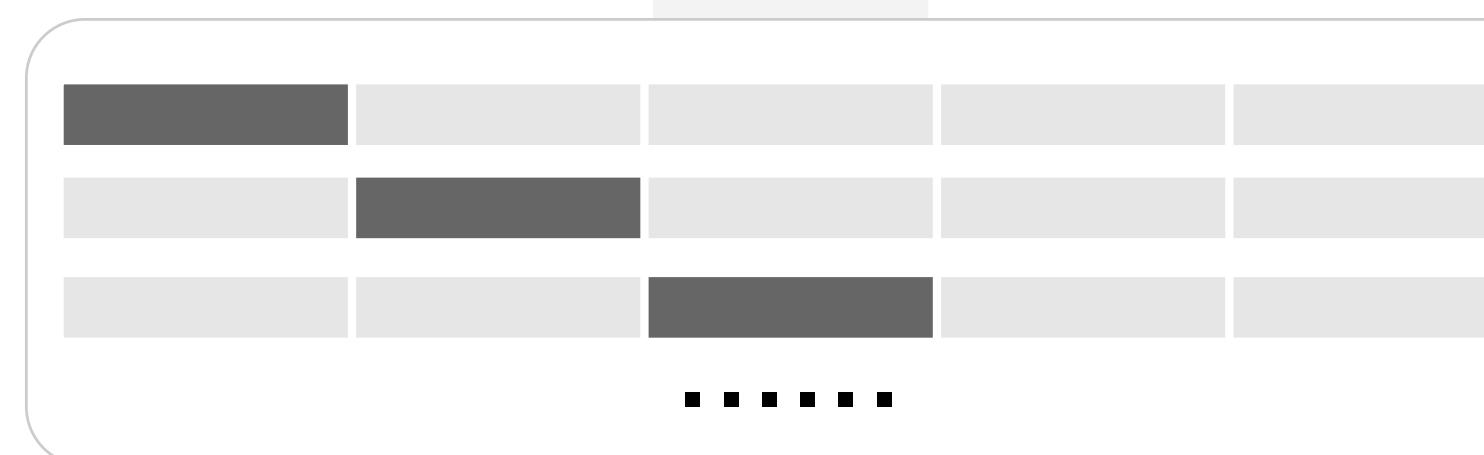
Filippo Menczer receives funding from Knight Foundation, Craig Newmark Philanthropies, Open Technology Fund, and DoD. He owns a Tesla.

Kai-Cheng Yang does not work for, consult, own shares in or receive funding from any company or organization that would benefit from this article, and has disclosed no relevant affiliations beyond their academic appointment.

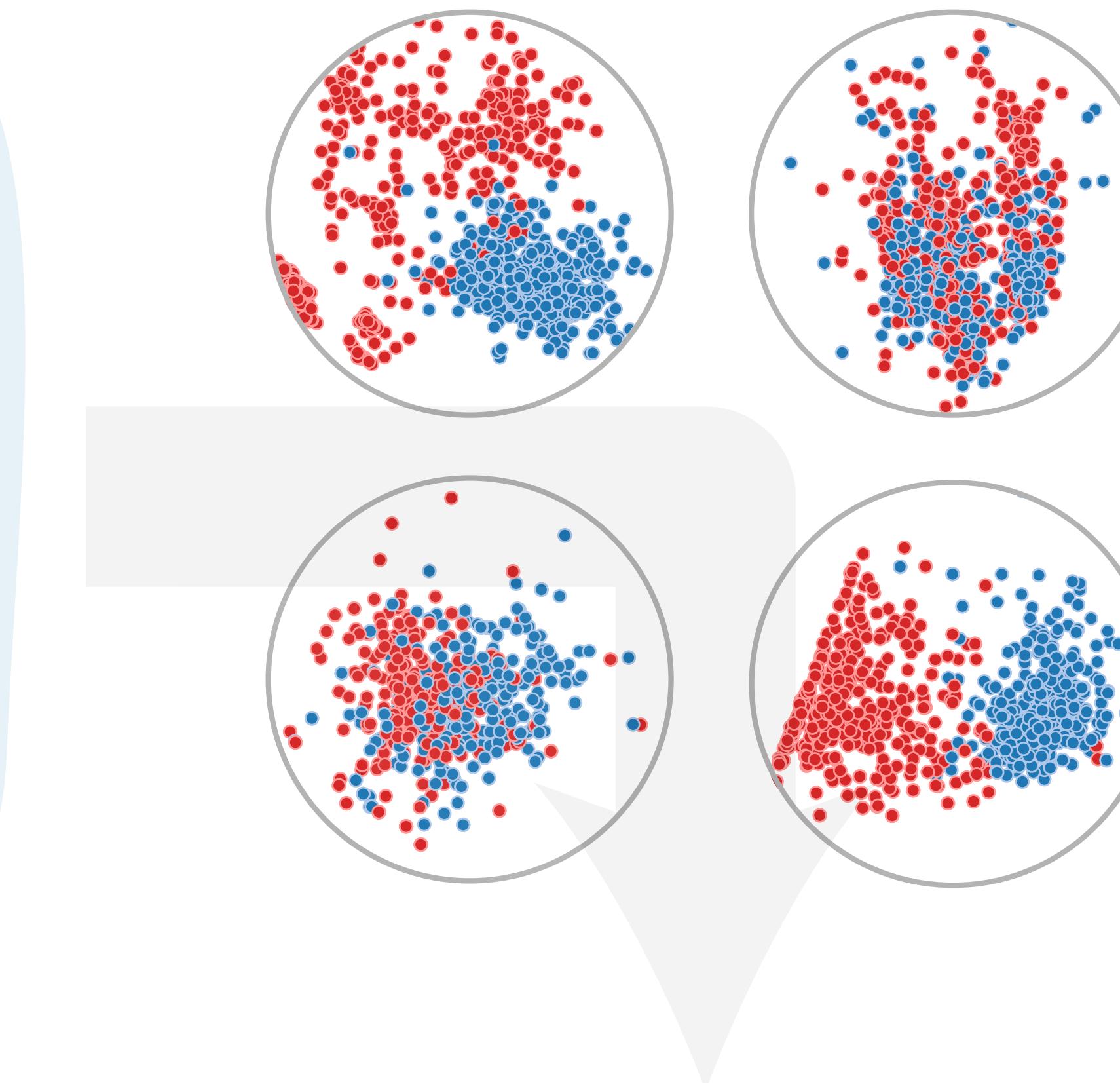
Candidate training sets



● bot ● human



Holdout test sets



Model leading in all
tests/validation wins

Dataset selection

Datasets used in different models

Dataset	Botometer	M196
caverlee	✓	
varol-icwsm	✓	✓
cresci-17	✓	✓
pronbots	✓	
celebrity	✓	✓
vendor-purchased	✓	
botometer-feedback	✓	✓
political-bots	✓	✓

Performances of different models

Metric	Botometer	M196
botwiki-verified	0.92	0.99
midterm-18	0.96	0.99
gilani-17	0.67	0.68
cresci-rtbust	0.71	0.60
5-fold cross-validation	0.97	0.98
Spearman's r	1.00	0.60

Identification of misinformation at scale

Links	Low-credibility?
https://beforeitsnews.com/prophecy/2020/05/5g-whistleblower-coronavirus-test-spreading-virus-you-cannot-catch-the-coronavirus-unless-you-take-the-covid-19-test-or-covid-19-vaccines-exposing-the-covid-19-lie-great-videos-2511687.html	Yes
https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html	No
https://humansarefree.com/2020/04/covid-19-and-5g-whats-the-connection.html	Yes
https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html	No

Identification of misinformation at scale



HOME THE GATEWAY PUNDIT

The Gateway Pundit

Share:

f t p e t s in f G +

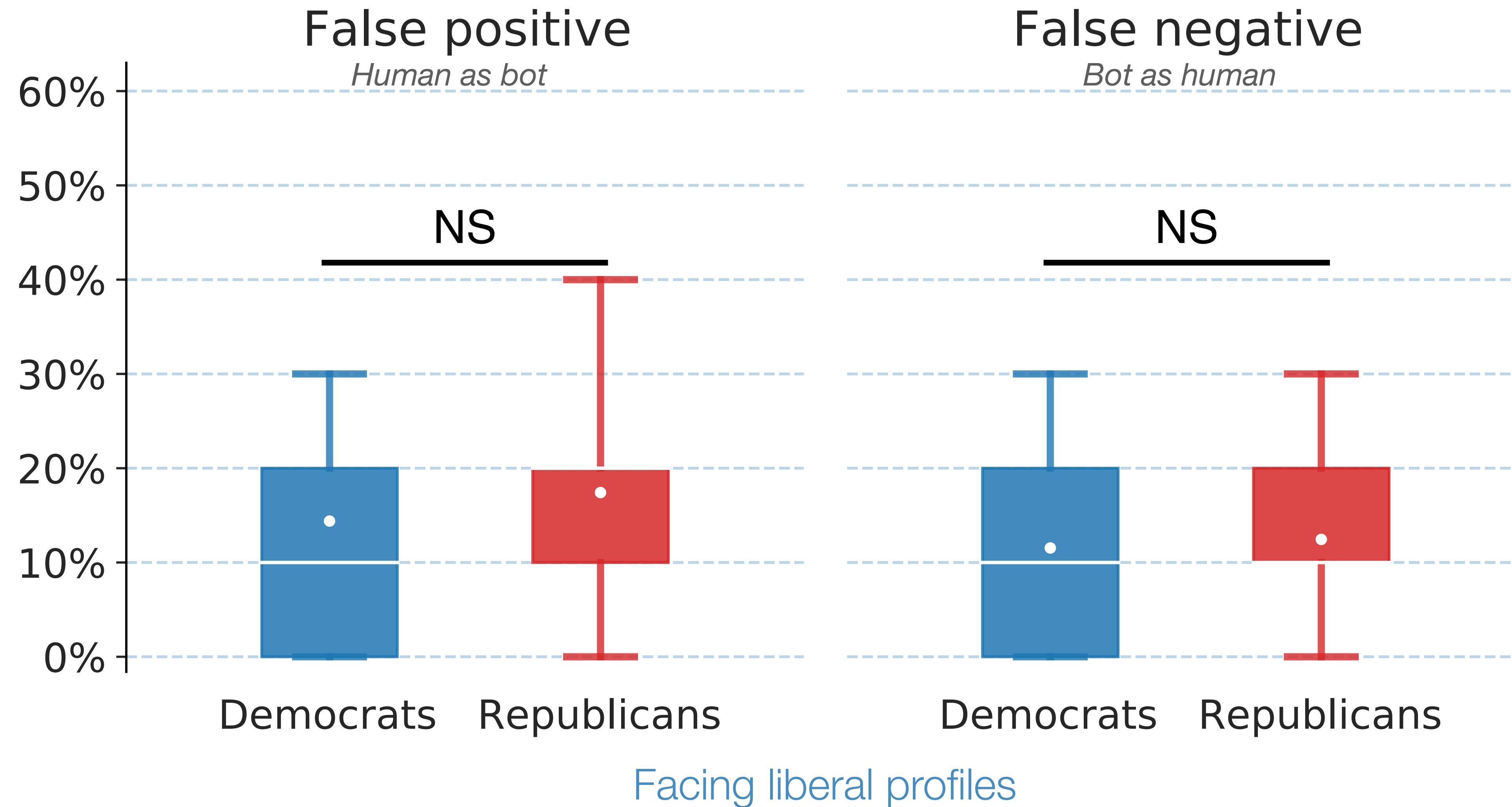
Extreme Left Left-Center Least Biased Right-Center Right Extreme

QUESTIONABLE SOURCE

The screenshot shows the header of the website with "HOME" and "THE GATEWAY PUNDIT" buttons. Below the header, the title "The Gateway Pundit" is displayed. A "Share" button with icons for various social media platforms follows. Below that is a horizontal color scale with labels: "Extreme" (blue arrow), "Left" (light blue), "Left-Center" (white), "Least Biased" (light green), "Right-Center" (yellow), "Right" (orange), and "Extreme" (red arrow). A yellow dot is positioned on the "Right-Center" label. At the bottom, there is a vertical color scale for "Factual Reporting" with labels: "Very High" (light blue), "High" (medium blue), "Mostly Factual" (white), "Mixed" (light orange), "LOW" (orange), and "Very Low" (red). A red bar covers the "Mixed", "LOW", and "Very Low" categories. A horizontal line with the text "QUESTIONABLE SOURCE" is at the bottom.

<https://mediabiasfactcheck.com/the-gateway-pundit/>

Why did they make mistakes?



Policy preference

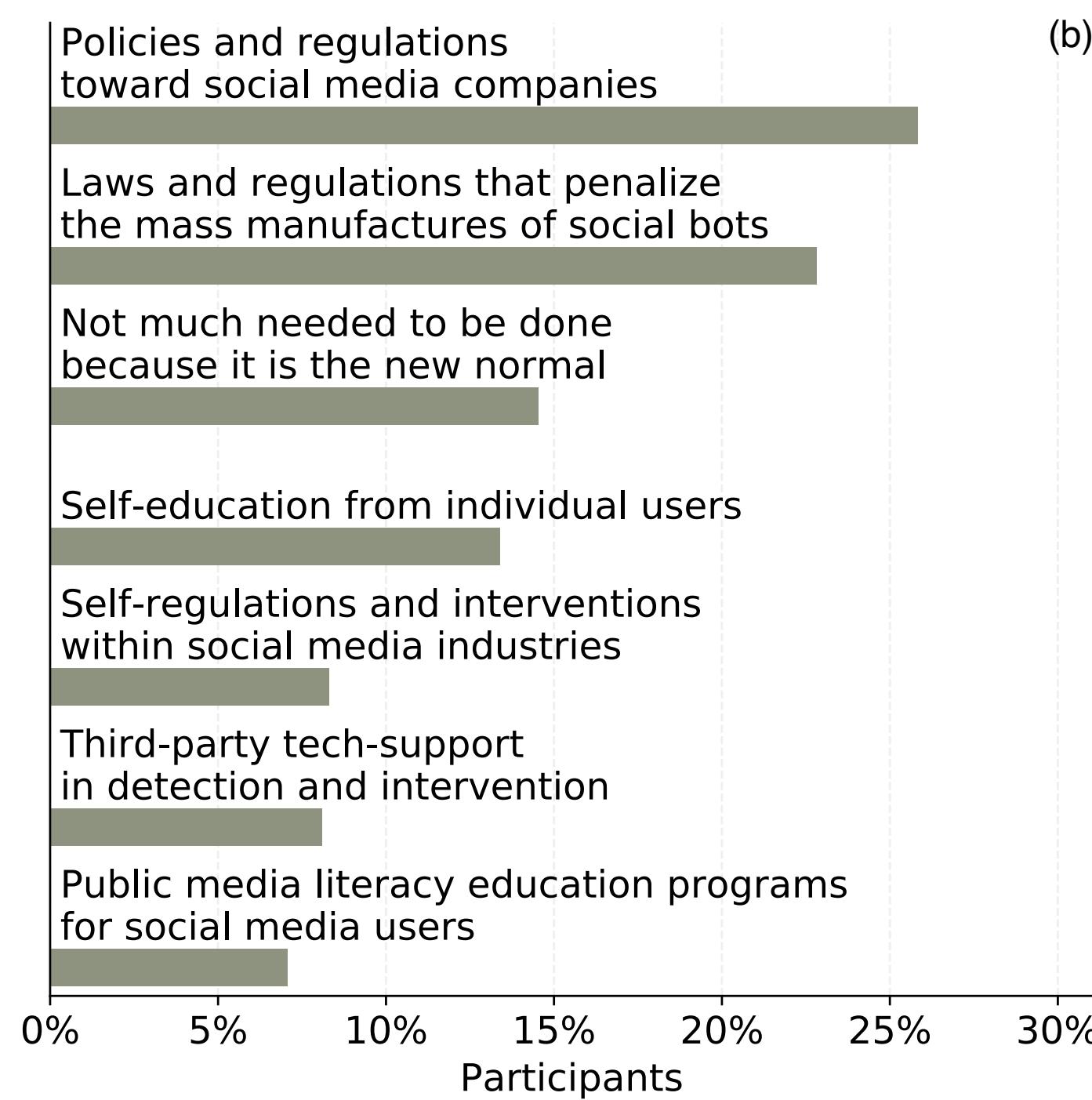
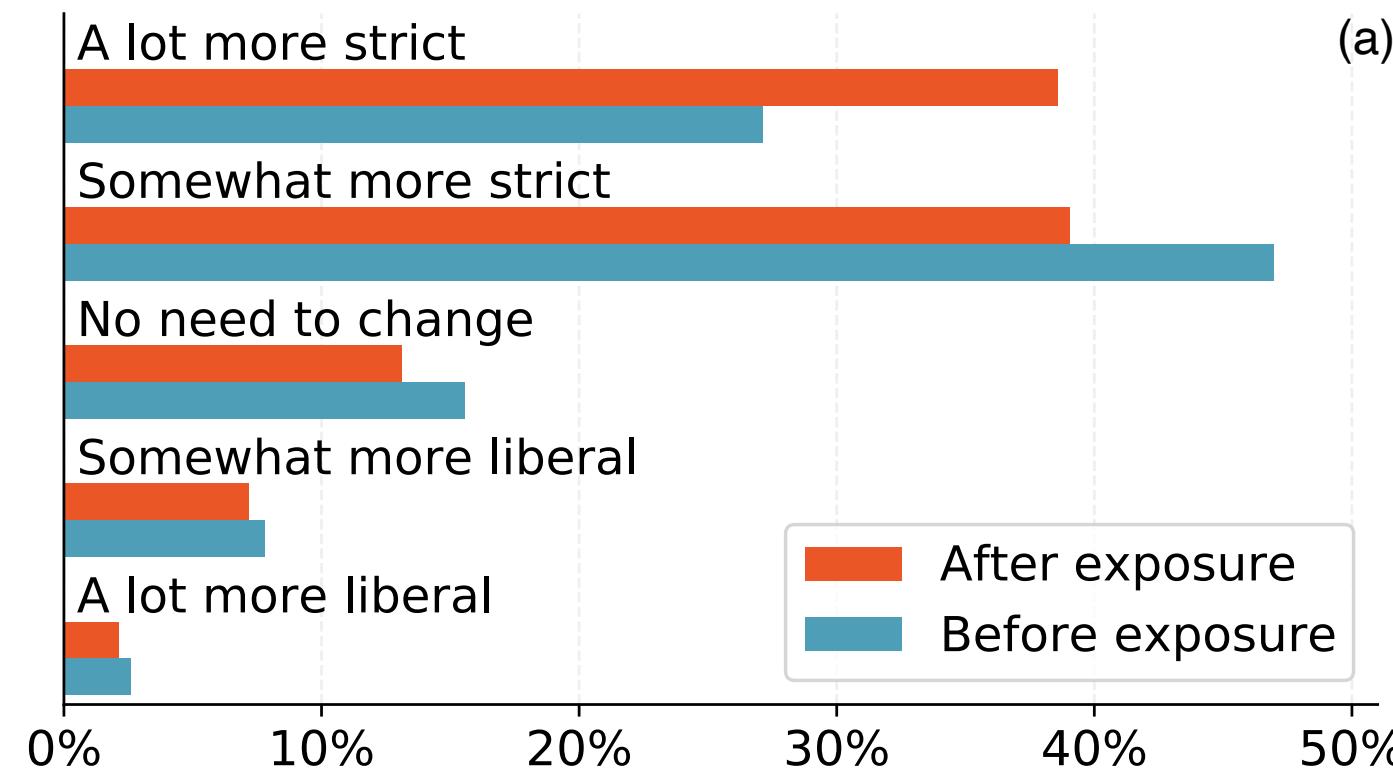


Table 2: Ordinary least-square models predicting changes of policy restrictions

Predictors	Model I ¹		Model II ²	
	B	p	B	p
(Intercept)	.05	.443	.04	.478
Age	-.04	.33	-.04	.338
Education	-.06	.112	-.06	.108
Independent ^a	-.06	.535	-.03	.56
Republican ^a	-.12	.259	-.11	.302
<i>Change scores^b</i>				
Perceived prevalence	.06	.153	.04	.308
Self-efficacy	-.11	.005	-.1	.008
Third-person perceptions	.07	.073		
Perceived influence on self			-.00	.988
Perceived influence on others			.10	.015
R ² / R ² adjusted	.03/.02		.034/.022	

Notes:

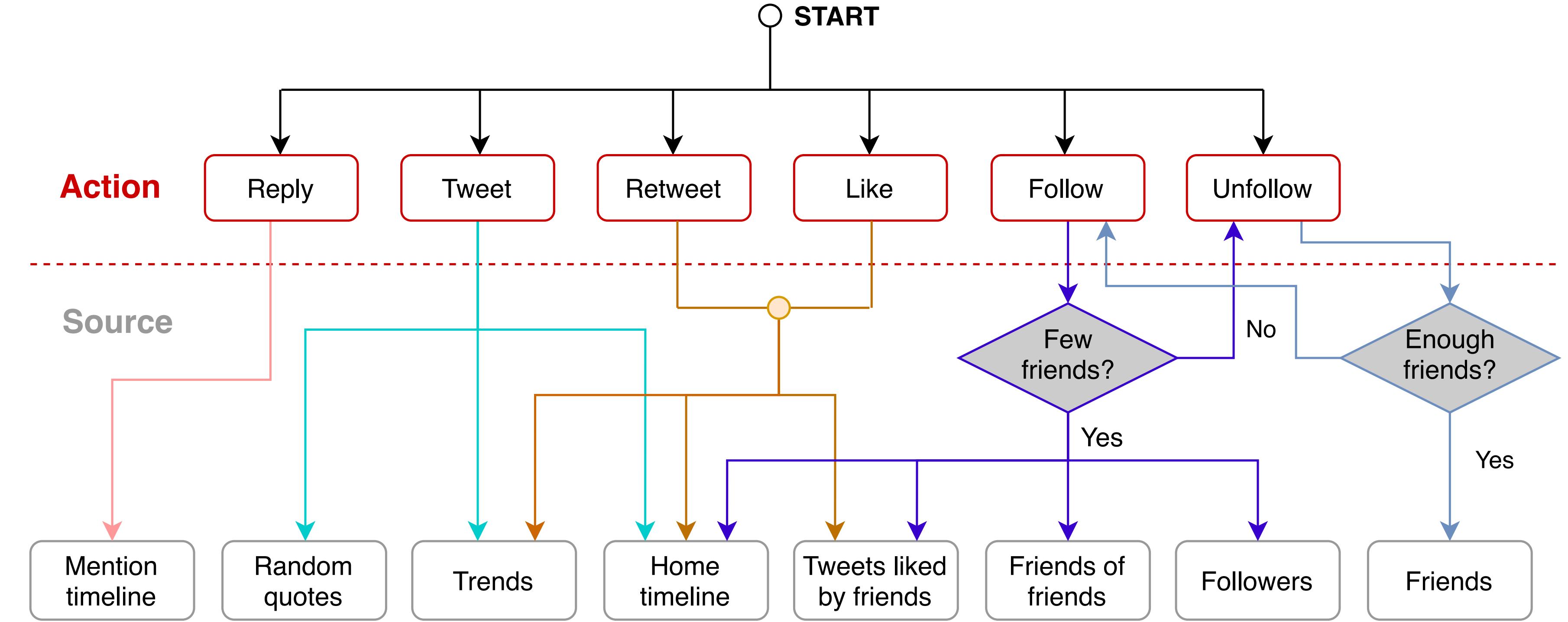
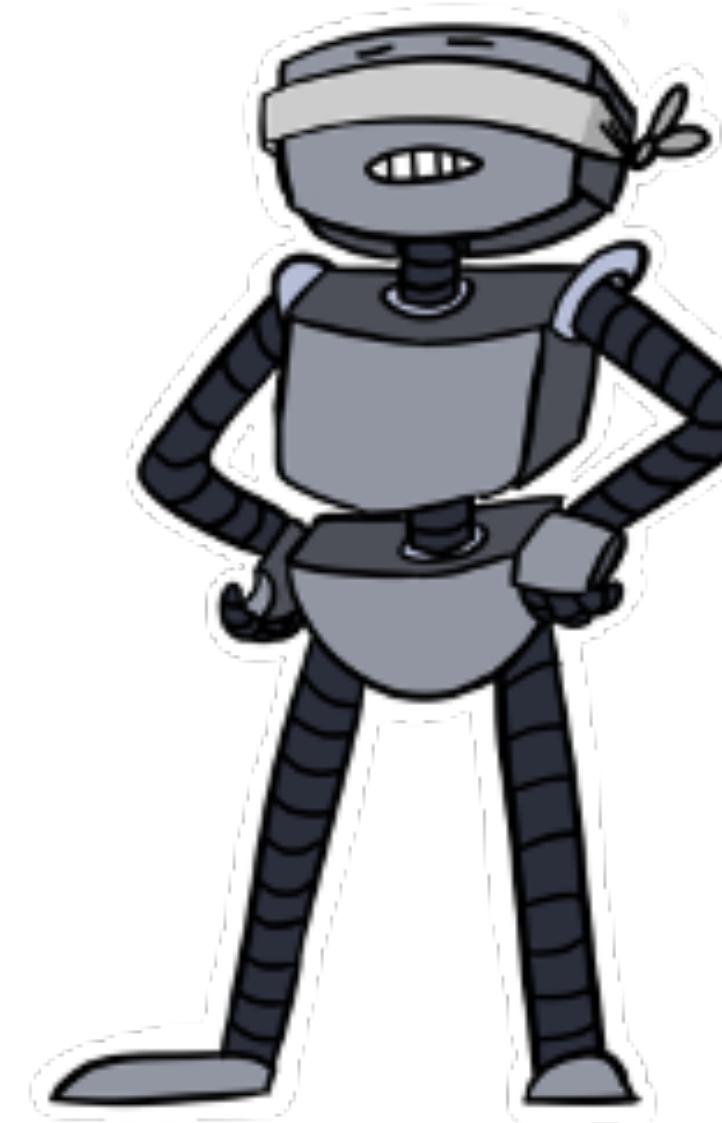
¹ Model I includes the discrepancy between perceived influences on others and on selves as a single independent variable to measure the effect of third-person perceptions.

² Model II includes the perceived influence of bots on others and on selves as separate independent variables.

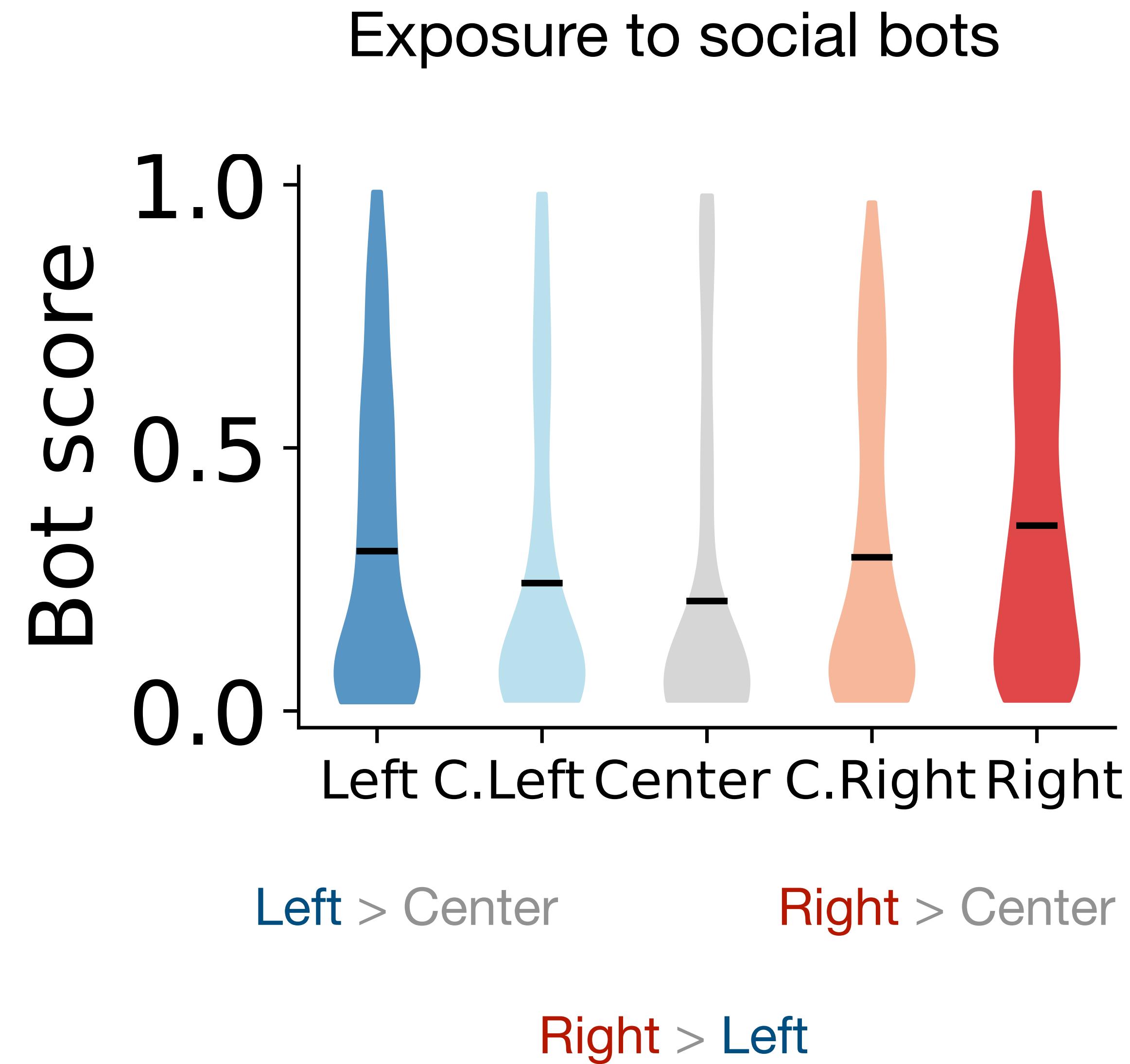
^a Democrat is the reference group.

^b Change scores are the differences between pre- and post-exposure measures.

Drifter models

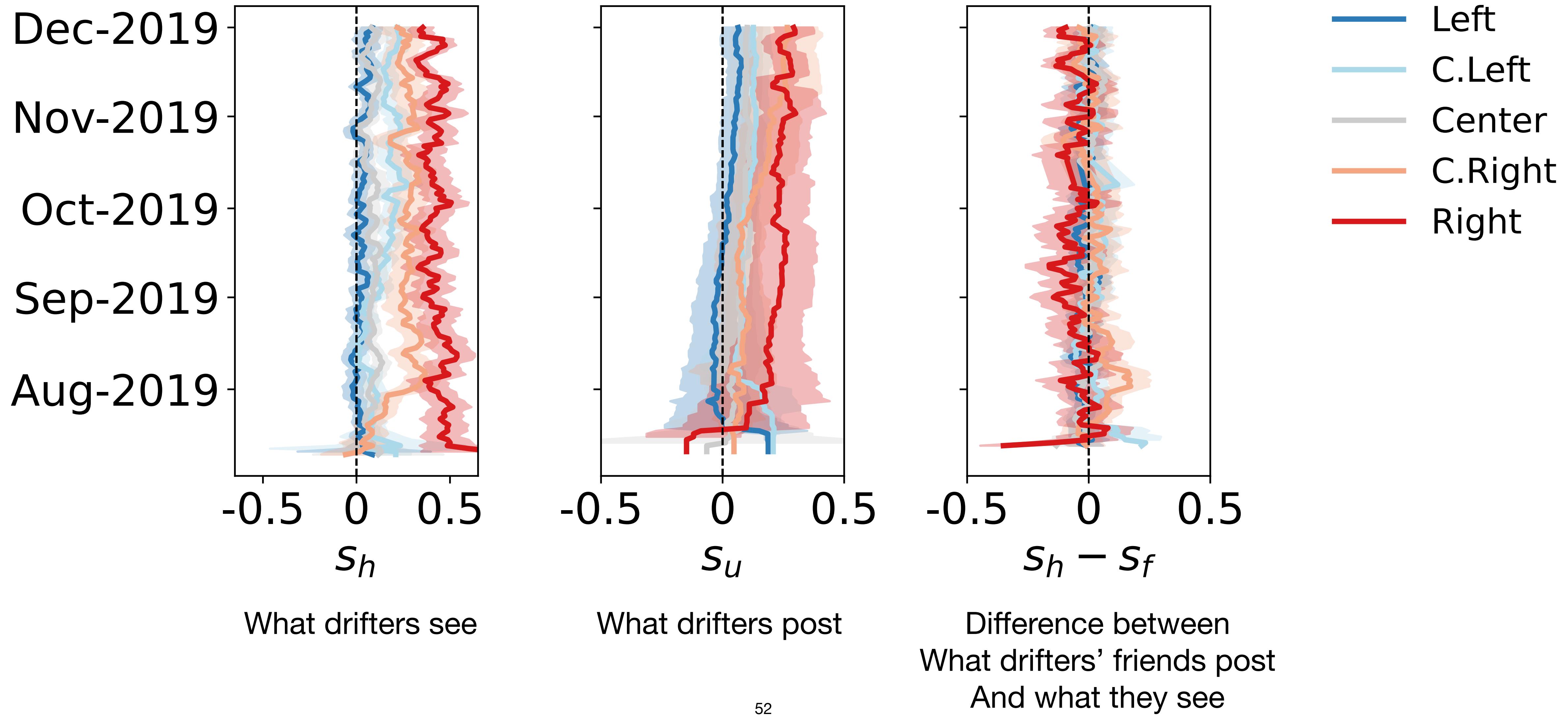


Bots as research instruments



Bots as research instruments

Political valence and algorithmic bias



Third-person perception

