giancarlo.ruffo@unipo.it

@giaruffo

Giancarlo Ruffo

Studying Fake News Spreading, Polarization Dynamics, and Social Media Manipulation
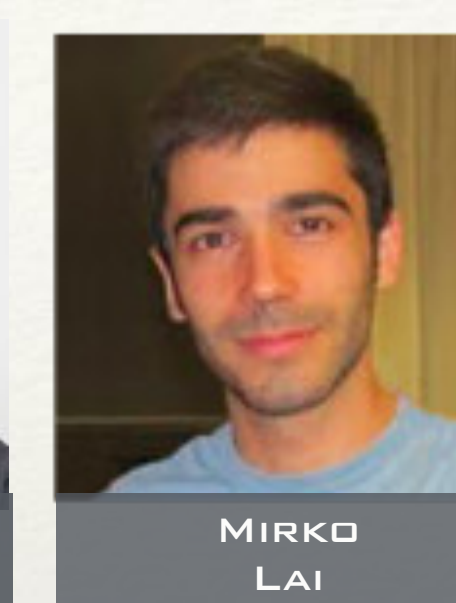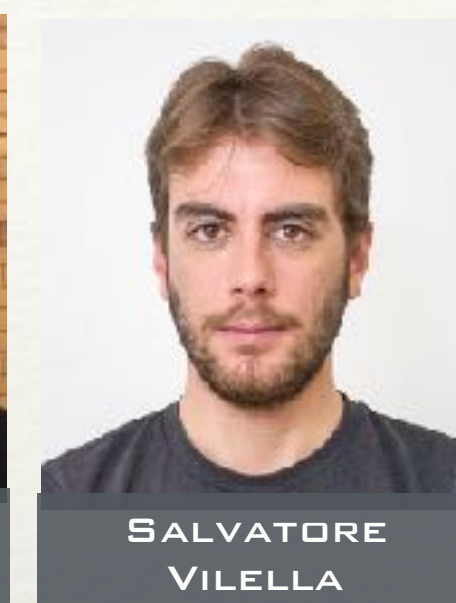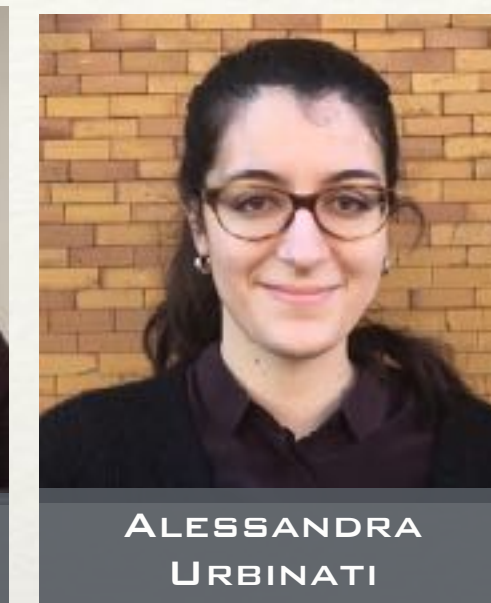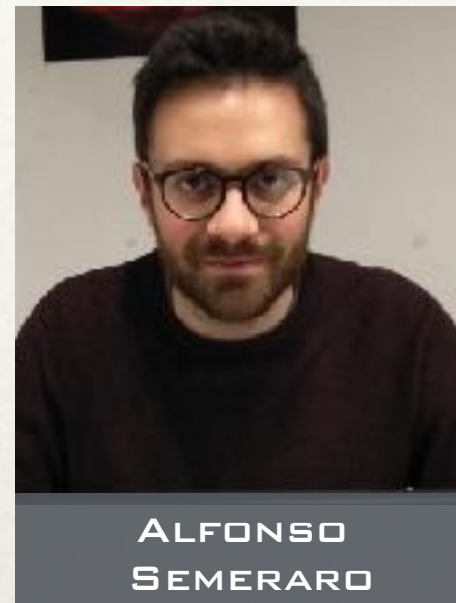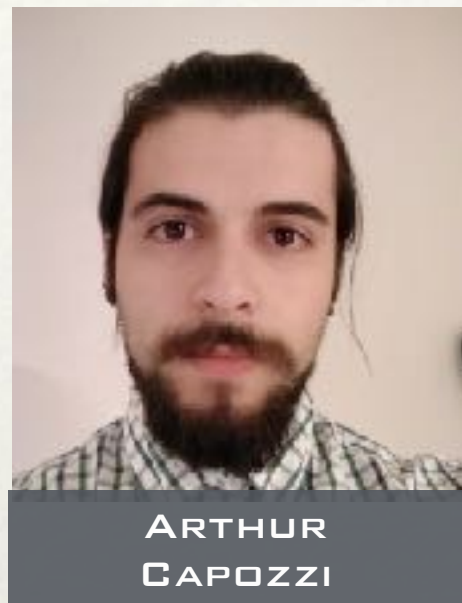
Tutorial

21/9/2022 - itaDATA 2022, Milan, Italy

-

Objectives of this overview

❖ Although fighting fake news spreading and social media manipulation is a multidisciplinary challenge, we want to focus on topics of interest for a data scientist

Us

ARTHUR CAPOZZI

ALFONSO SEMERARO

ALESSANDRA URBINATI

SALVATORE VILELLA

MIRKO LAI

GIANCARLO RUFFO

UNIVERSITÀ DI TORINO

UPO — UNIVERSITÀ DEL PIEMONTE ORIENTALE

ARC²S: Applied Research on Computational Complex Systems

❖ http://arcs.di.unito.it

❖ Our topics: data science, network science, information retrieval, computational social science

What this talk is

- ❖ Basic terminology

- ❖ The science of fake news main challenges

- ❖ The role of polarization

- ❖ Discussion and conclusion

What this talk is not

- fake news detection systems that use ML+NLP

- debunking a piece of information

- fully unbiased… :)

If you like this talk, you may want to read this

Information Disorders

is it just the same old story?

The science of fake news

2018 Manifesto

The science of fake news

David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, Jonathan L. Zittrain

The list of author affiliations is provided in the supplementary materials.

Email: d.lazer@northeastern.edu
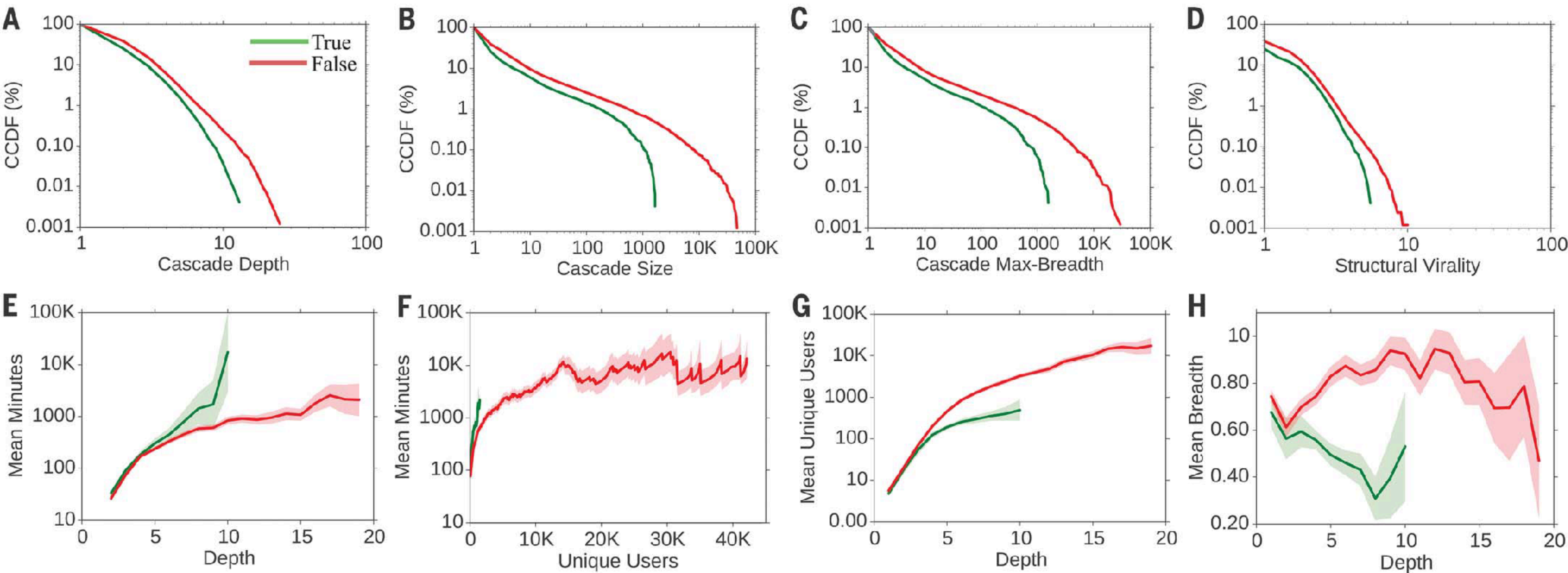– Hide authors and affiliations

"… much remains unknown regarding the vulnerabilities of individuals, institutions, and society to manipulations by malicious actors."

Prevalence

❖ Many observed that false stories in social media are more successful (in numbers and speed) than true stories

❖ *Which are the key factors?*

❖ *Who is to blame: bots or humans?*

Lies are faster than truth

- Dataset: ~126,000 stories tweeted by ~3 million people more than 4.5 million times.

- News classified as true or false using six independent fact-checking organizations that exhibited 95 to 98% agreement on the classifications.

S.Vosoughi, D. Roy, S. Aral, The spread of true and false news online, in Science, 09 mar 2018: 1146-1151

❖ Falsehood diffused significantly **farther**, **faster**, **deeper**, and **more broadly** than the truth in all categories of information
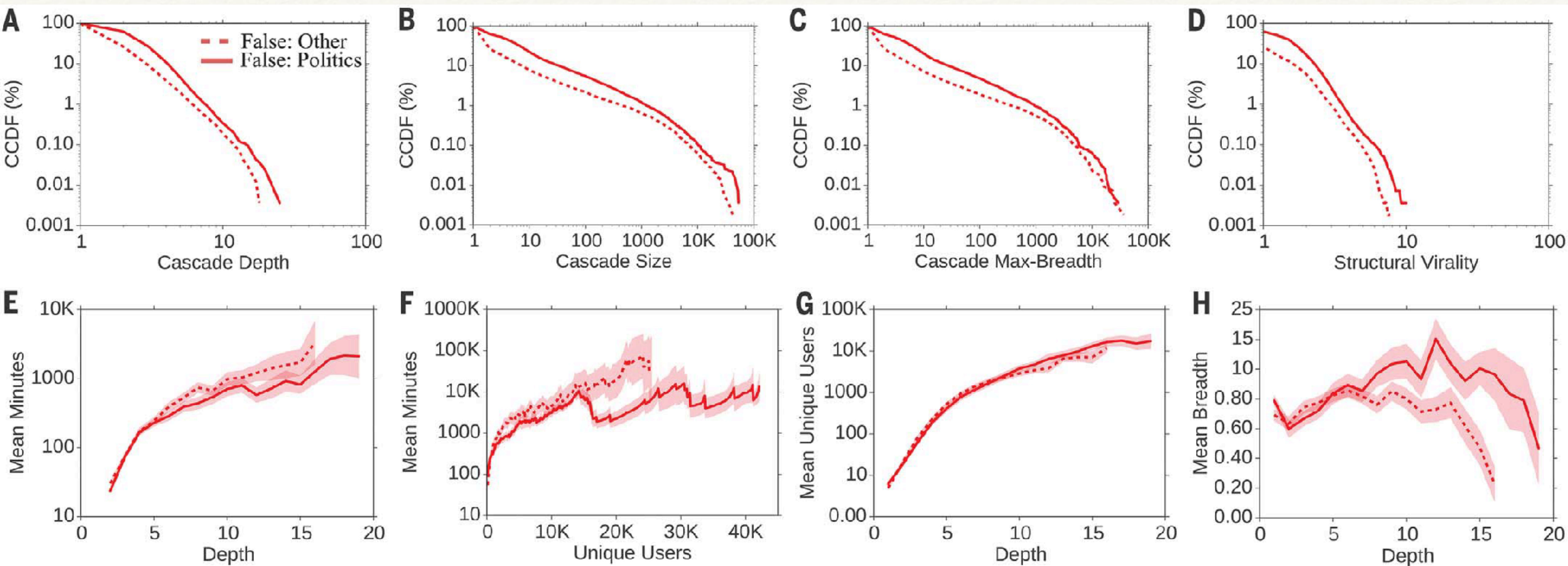
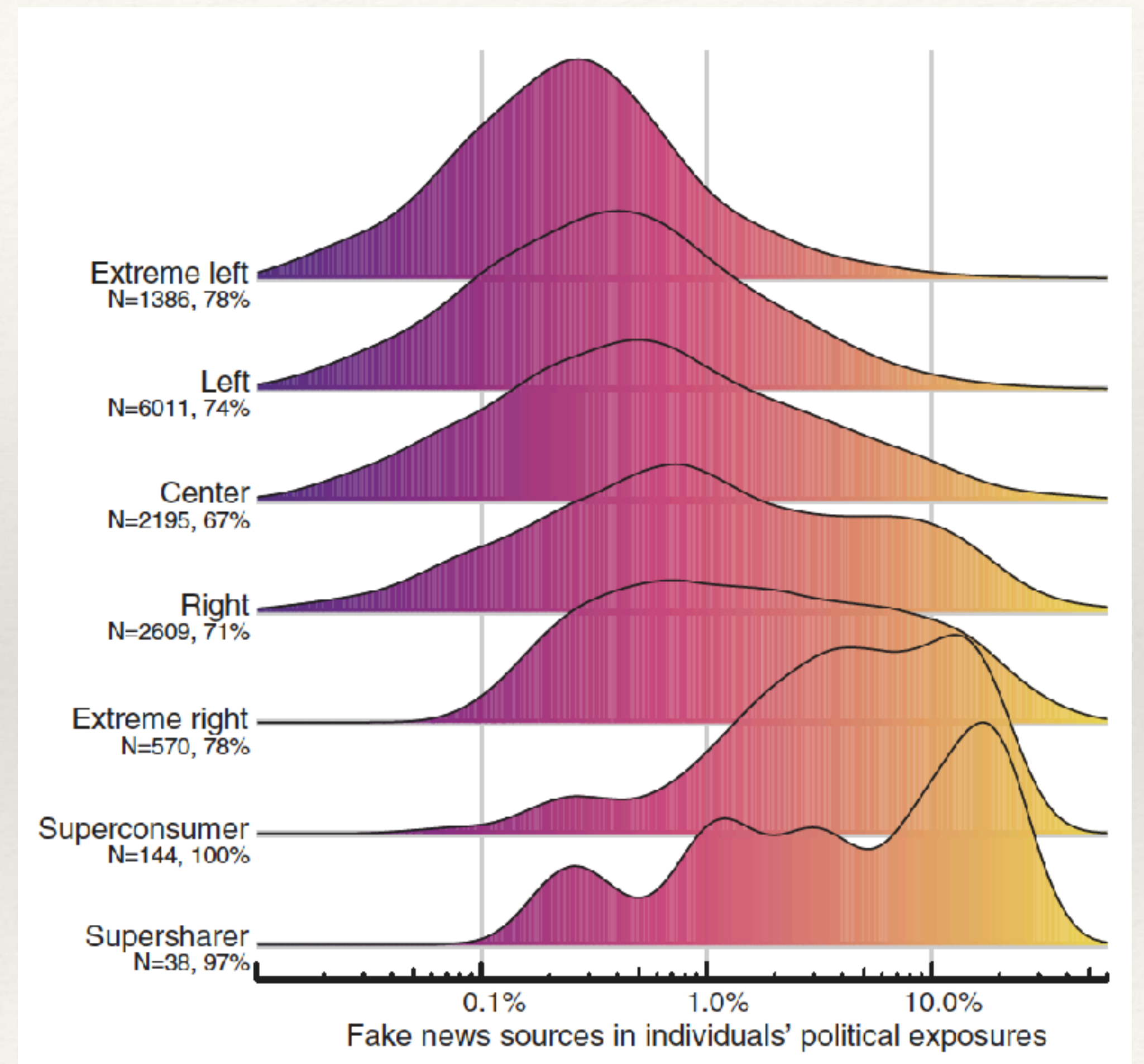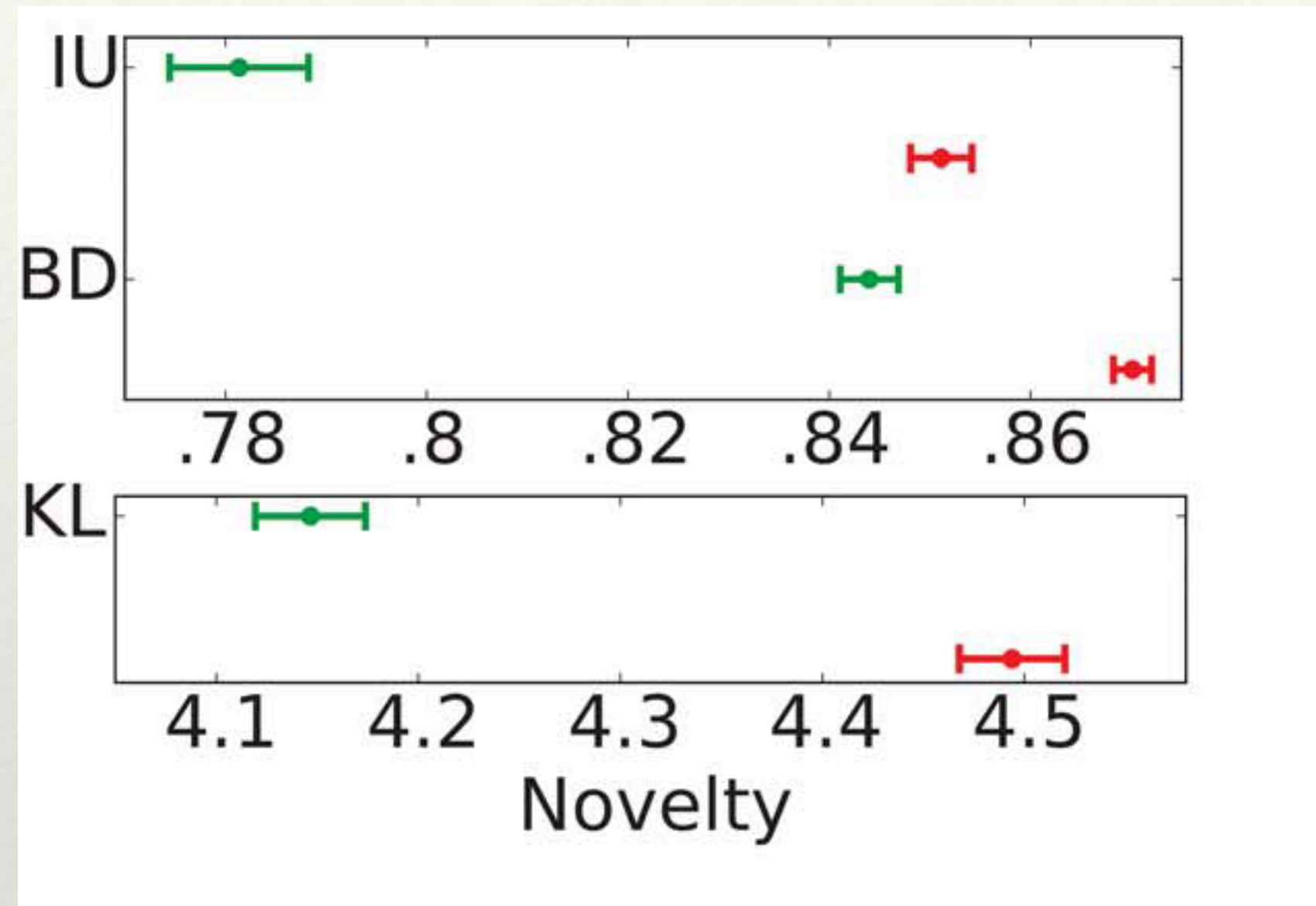❖ Effects were **more pronounced for false political news** than for false news about terrorism, natural disasters, science, urban legends, or financial information.

S.Vosoughi, D. Roy, S. Aral, The spread of true and false news online, in Science, 09 mar 2018: 1146-1151

Fake-News and elections

- Engagement with fake news sources extremely concentrated in 2016 US presidential elections

- Only 1% of individuals accounted for 80% of fake news source exposures, and 0.1% accounted for nearly 80% of fake news sources shared.

- Individuals most likely to engage with fake news sources were conservative leaning, older, and highly engaged with political news.
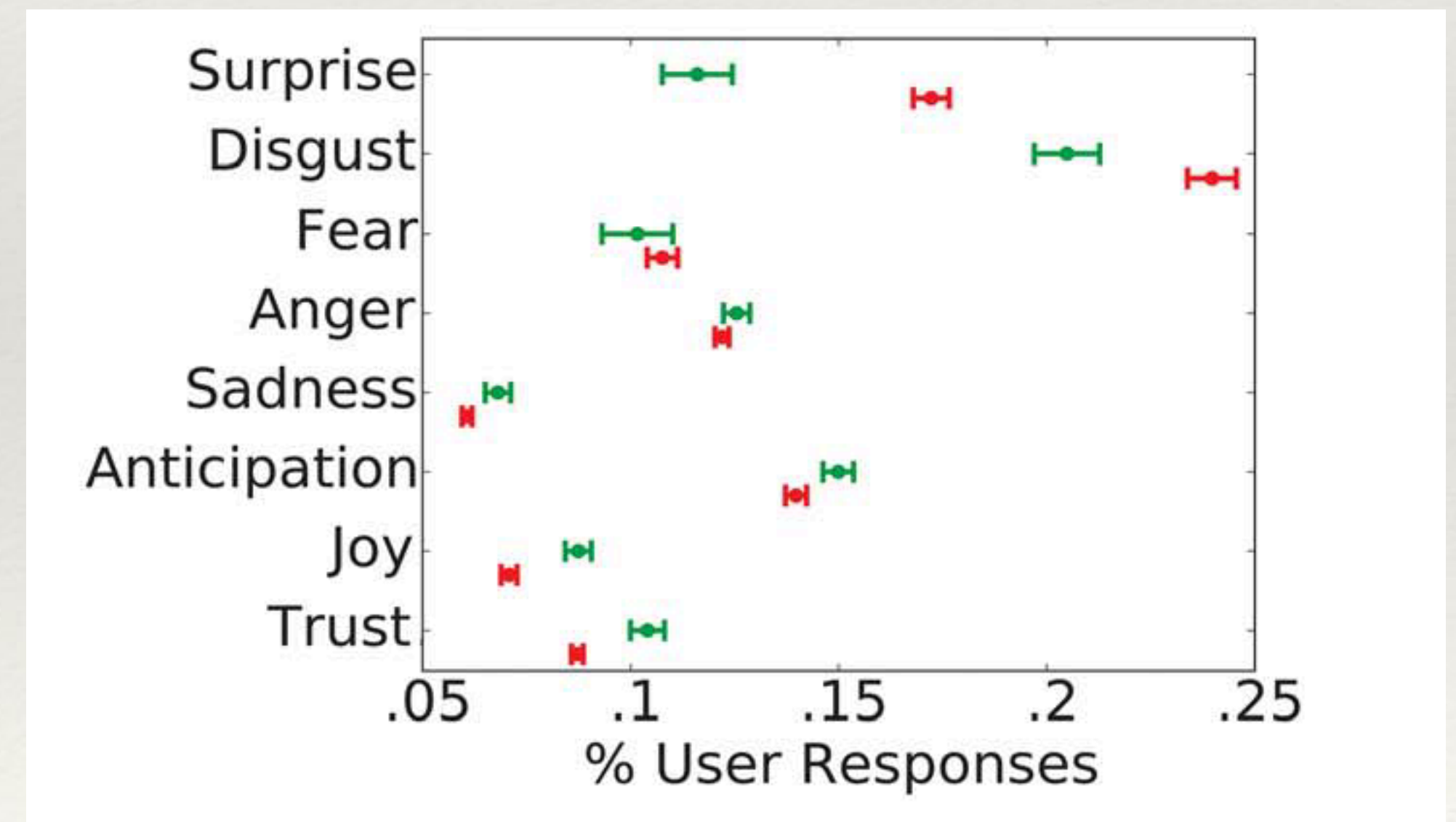

N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, D. Lazer, Fake news on Twitter during the 2016 U.S. presidential election, Science 25 Jan 2019:Vol. 363, Issue 6425, pp. 374-378 DOI: 10.1126/science.aau2706

Novelty and emotions
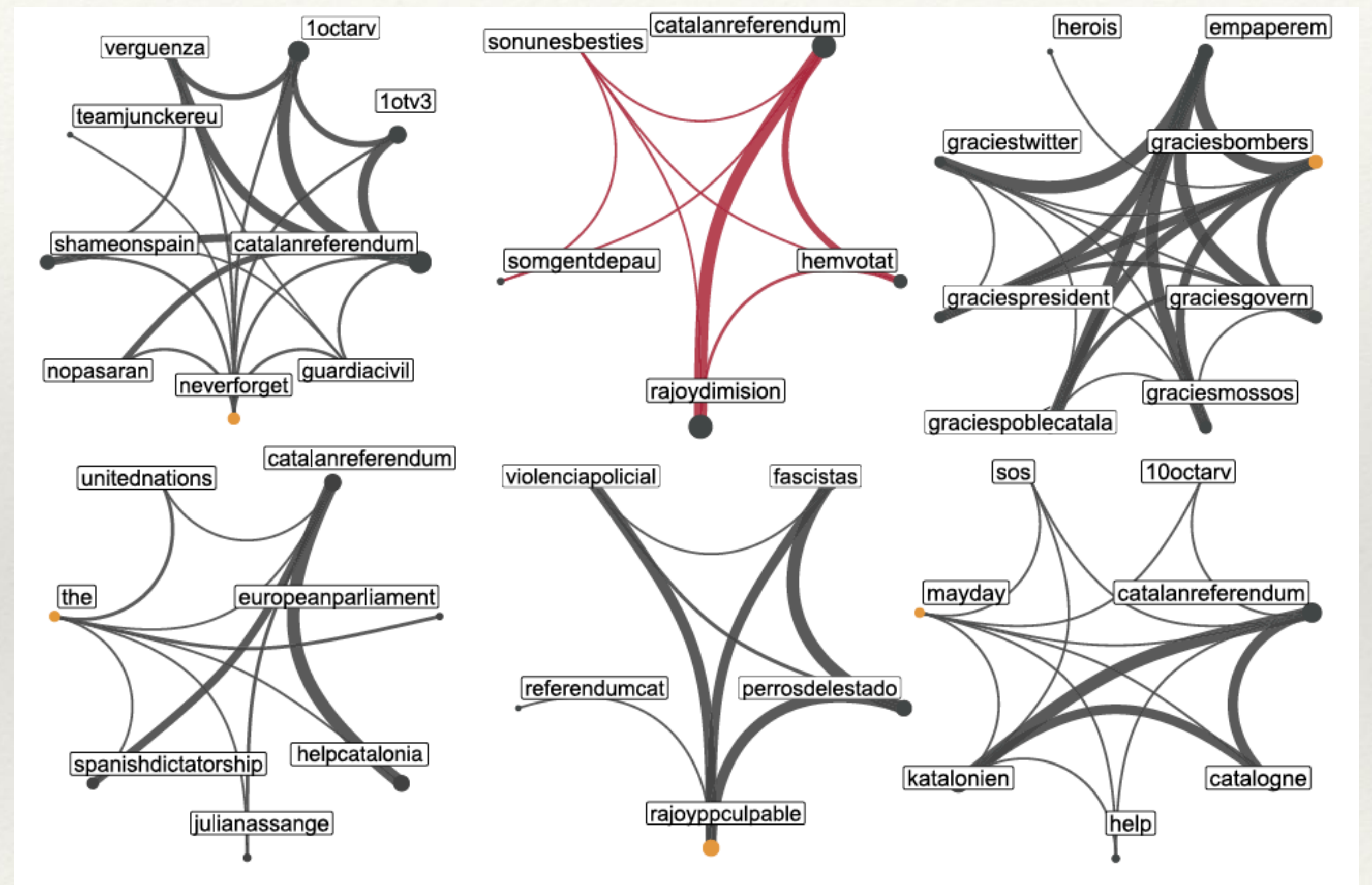

❖ False news **more novel** than true news, which suggests that people were more likely to share novel information

❖ False stories inspired **fear**, **disgust**, and **surprise** in replies, true stories inspired anticipation, sadness, joy, and trust.

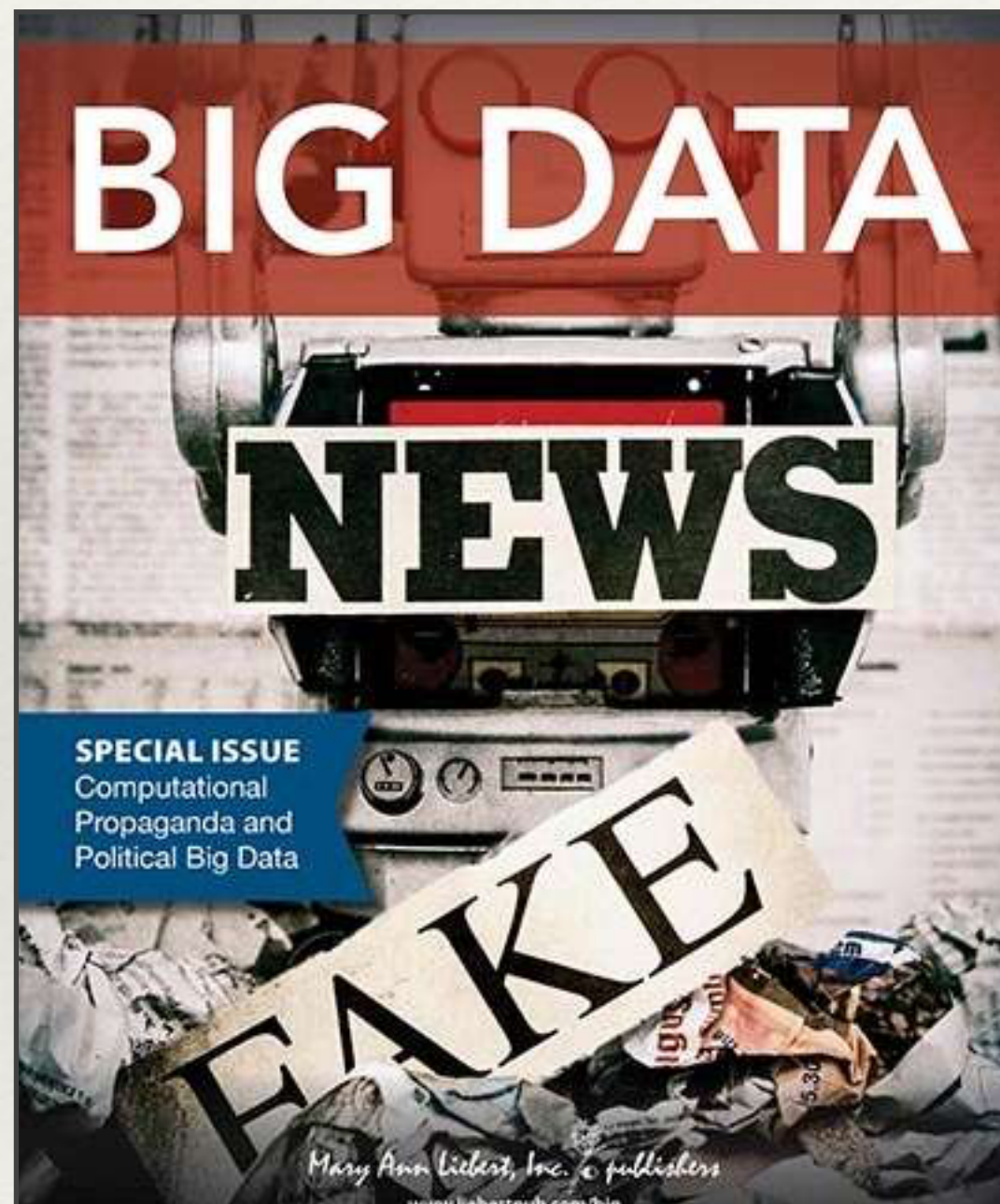S.Vosoughi, D. Roy, S. Aral, The spread of true and false news online, in Science, 09 mar 2018: 1146-1151

The role of emotions

- Large-scale social data collected during the **Catalan referendum for independence** on October 1, 2017, consisting of nearly 4 millions Twitter posts generated by almost 1 million users;

- Two polarized groups: **Independentists** vs **Constitutionalists**

- Structural and emotional roles played by **social bots**

  - Bots act from **peripheral areas** to target **influential humans** of both groups;

  - Bots bombard Independentists with **violent contents**, **increasing their exposure to negative and inflammatory narratives**, and exacerbating social conflict online.

M. Stella, E. Ferrara, M. Di Domenico, Bots increase exposure to negative and inflammatory content in online social systems, PNAS, Dec. 4, 2018, Vol. 115, no. 49, 12435–12440. www.pnas.org/cgi/doi/10.1073/pnas.1803470115

the of role social bots

The interplay between bots and polarization

❖ A simple **spambot can take power** in a social network

❖ A seed of **polarization** found in pre-existing network **structure**

❖ … also the structure changed after our experiment was run!

❖ What if the real identity and motivations of Lajello were fact-checked?

Analyzing the structure of a misinformation network

❖ *What are the structural and dynamic characteristics of the core of the misinformation diffusion network, and who are its main purveyors?*

❖ "As we move from the periphery to the core of the network, fact-checking nearly disappears, while social bots proliferate."

https://hoaxy.iuni.iu.edu

Shao C, Hui P-M, Wang L, Jiang X, Flammini A, Menczer F, et al. (2018) Anatomy of an online misinformation network. PLoS ONE 13(4):e0196087. https://doi.org/10.1371/journal.pone.0196087
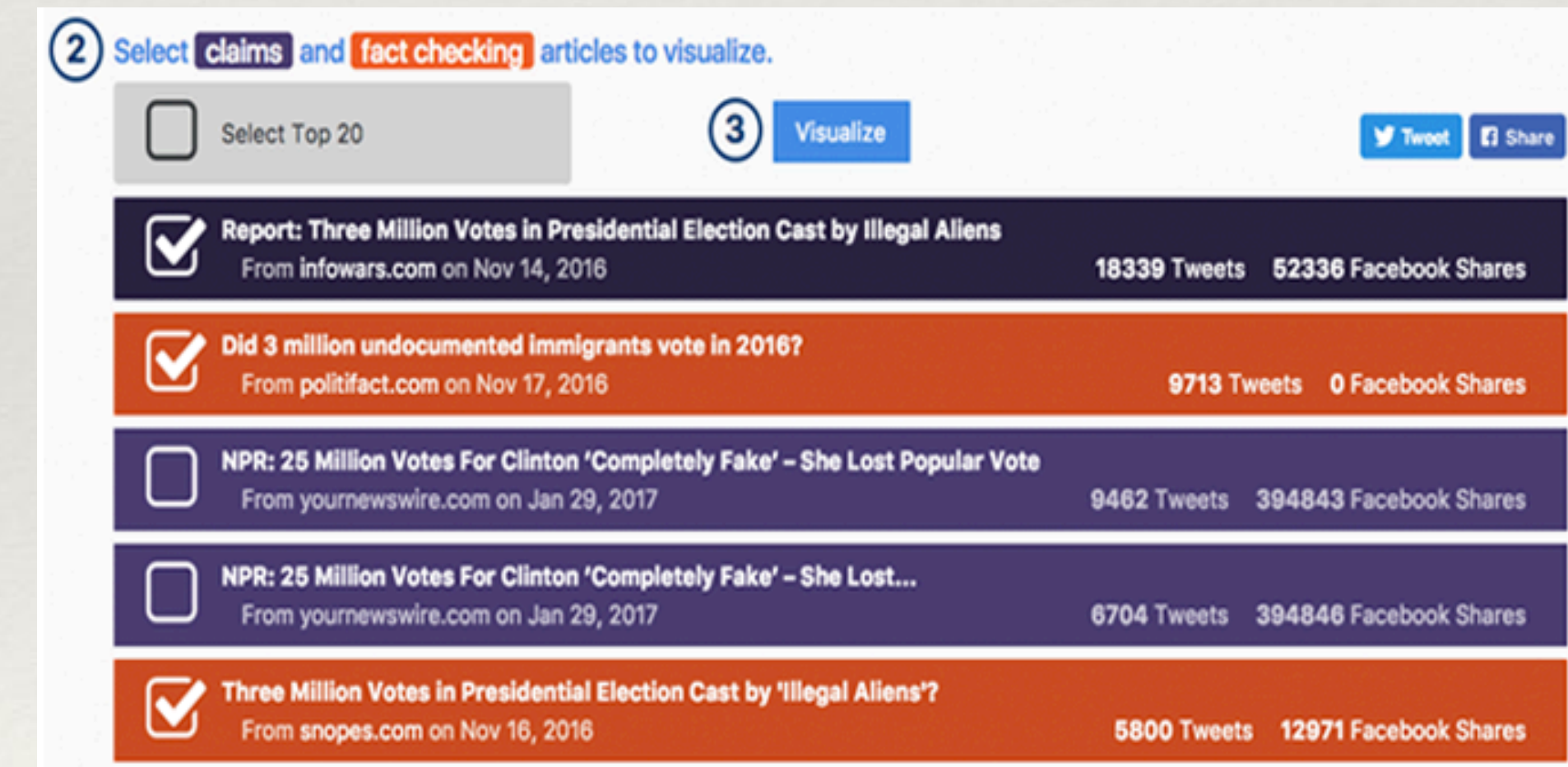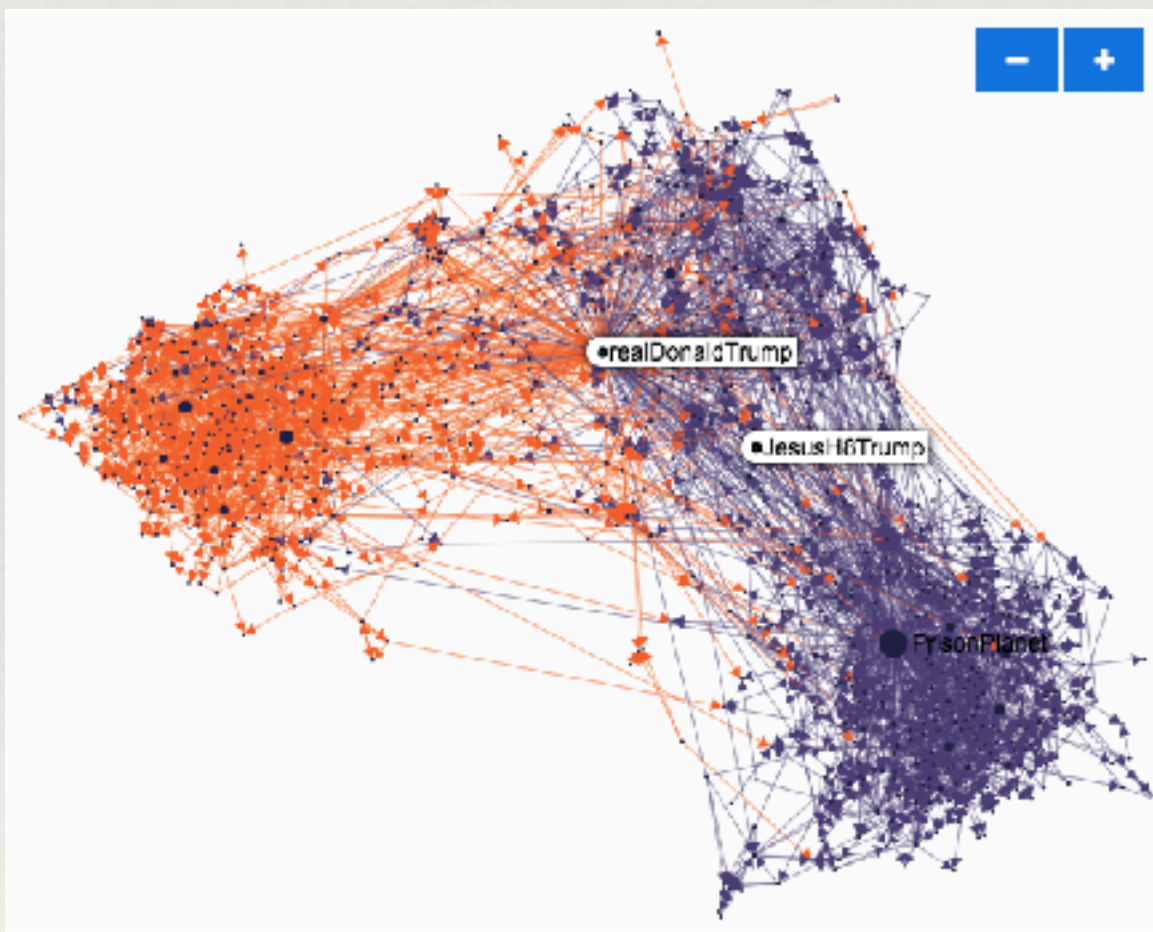
The spread of low-credibility content by social bots

❖ 14 million messages spreading 400 thousand articles on Twitter during ten months in 2016 and 2017

❖ Social bots played a disproportionate role in spreading articles from low-credibility sources.

❖ Bots amplify such content in the early spreading moments, before an article goes viral.

❖ They also target users with many followers through replies and mentions. Humans are vulnerable to this manipulation, resharing content posted by bots.

Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A., Menczer, F., The spread of low-credibility content by social bots. *Nat Commun* **9,** 4787 (2018). https://doi.org/10.1038/s41467-018-06930-7

BotSlayer and Botometer (IU)

❖ **BotSlayer**: it tracks and detect potential manipulation of information spreading on Twitter

https://osome.iuni.iu.edu/tools/botslayer/
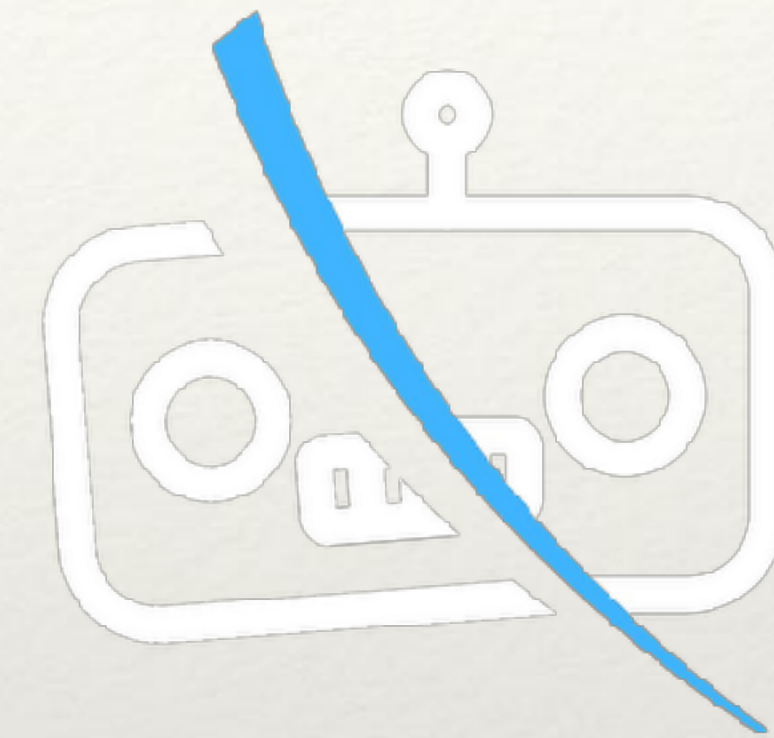
❖ **Botometer** (formerly known as BotOrNot) :checks the activity of a Twitter account and gives it a score. Higher scores mean more bot-like activity.

https://botometer.osome.iu.edu
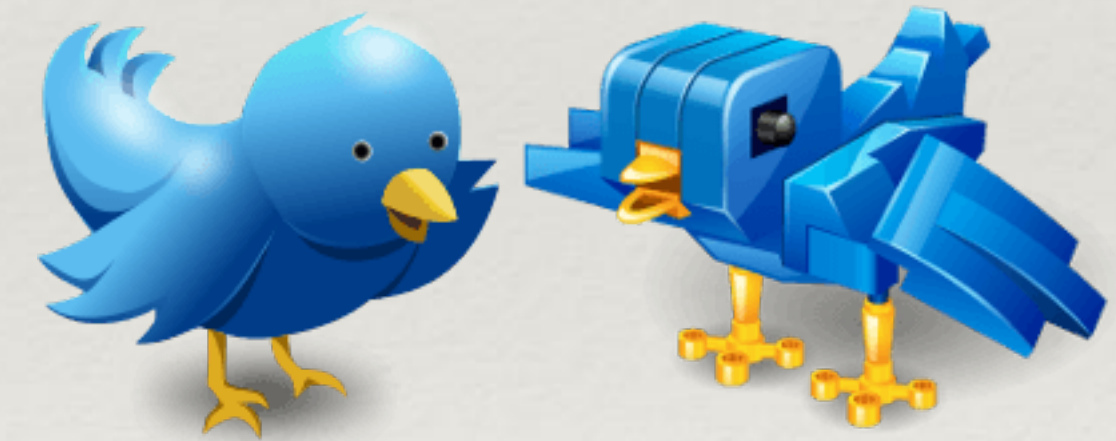
…but humans should be blamed the most

Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A., Menczer, F., The spread of low-credibility content by social bots. *Nat Commun* **9,** 4787 (2018). https://doi.org/10.1038/s41467-018-06930-7

The Hype Machine

- Prevalence of fake-news and role of social bots in spreading misinformation

- Bots share **novel** fake news and retweet it broadly
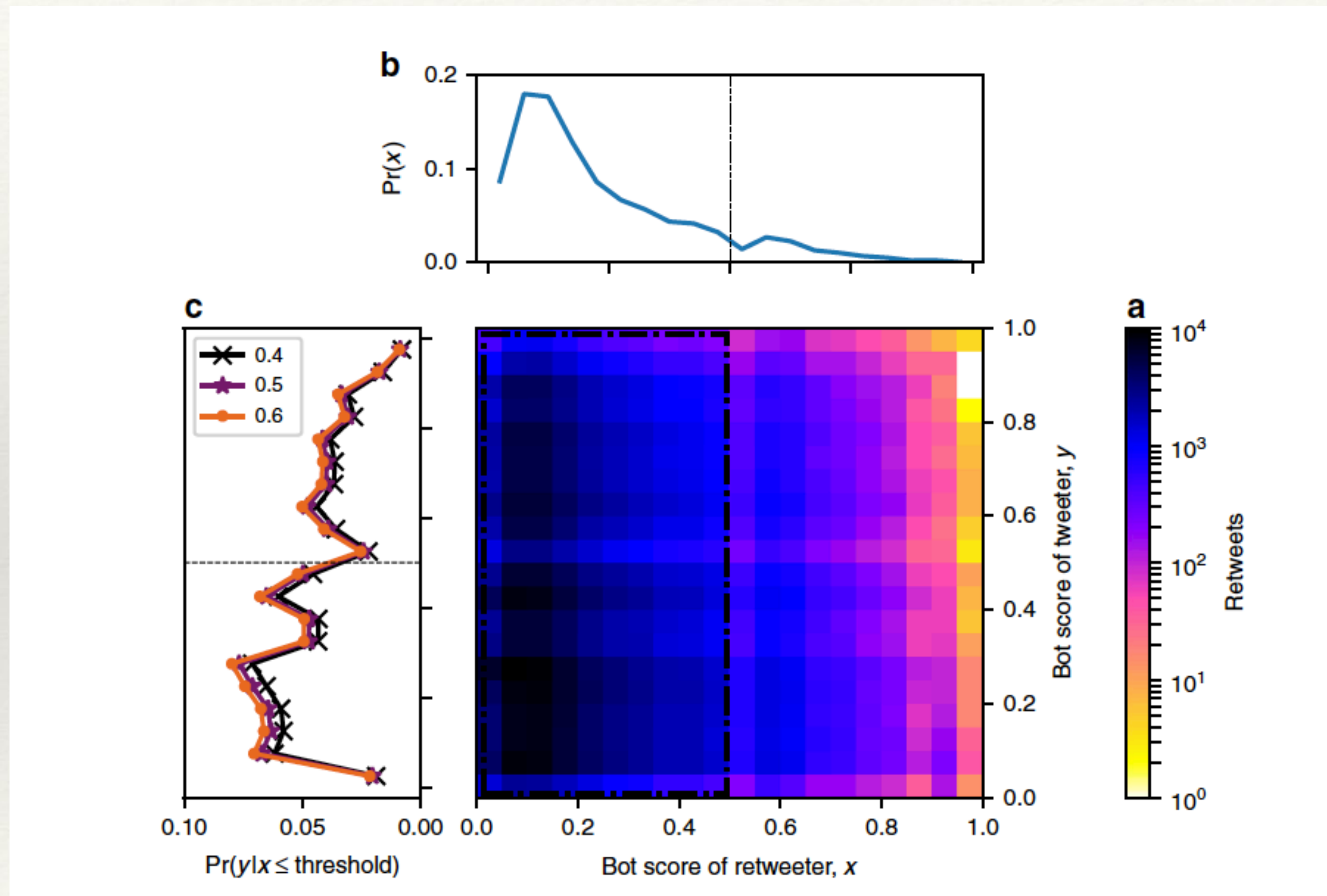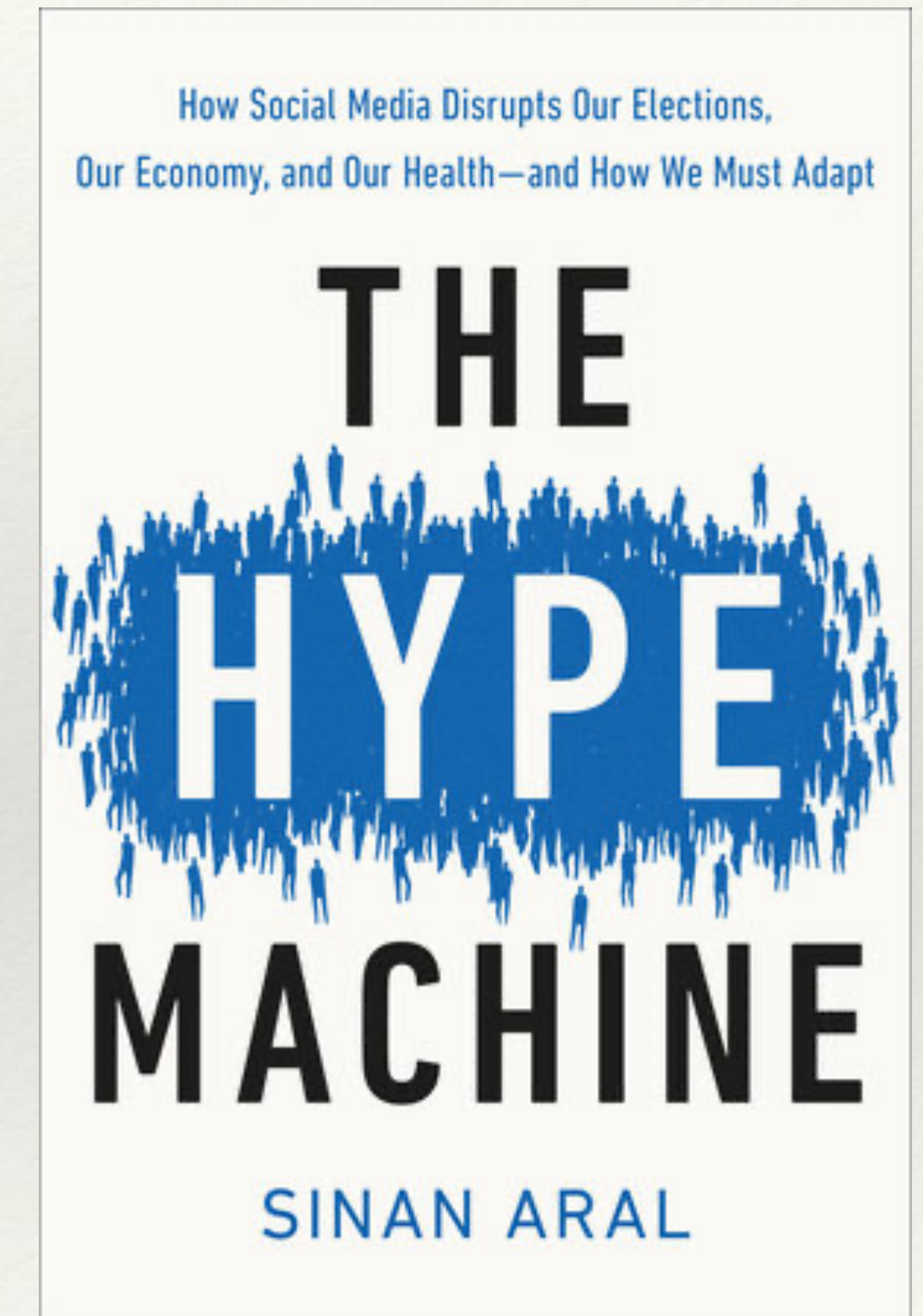
- Bots **mention influential humans** incessantly

- The strategy works when influential people are fooled into sharing the content.

- **Misleading humans is the ultimate goal of any misinformation campaign**

How Social Media Disrupts Our Elections, Our Economy, and Our Health—and How We Must Adapt

THE HYPE MACHINE

SINAN ARAL

https://www.salon.com/2020/09/27/fake-news-bots-spreading-misinformation-2020-election-propaganda/

US MARKETS
OPEN
In the news

▲ **Dow Jones -1.05%**   ▲ **Nasdaq -2.47%**   ▲ **S&P 500 -1.63%**   ▼ **TSLA +3.27%**   ▼ **FB +1.36%**   ▼ **BABA +6.2%**

| @elonmusk | Check user | Check followers | Check friends |

@elonmusk ✔

1 / 5

Bot type scores ❓

Echo-chamber	0.6
Fake follower	0.4
Financial	0.0
Self declared	1.7
Spammer	0.0
Other	2.2

Bot score based on ❓

All features: **1.0**

31% of accounts with a bot score above 1.0 are labeled as humans.

Language-independent: **1.8**

Majority tweet language: **en**

| 👤 Profile | 🐦 Tweet | 📊 Details | 💬 Feedback |

Exclusive: Putin's 'chef,' the man behind the troll factory

By Tim Lister, Jim Sciutto and Mary Ilyushina, CNN

Updated 0057 GMT (0857 HKT) October 18, 2017

The role of polarization

Exploitations of echo chambers

❖ "Echo chambers" methaphor, by Cass Sunstein in 2001

❖ Groups of like minded individuals that keep reinforcing their own beliefs (group-thinking)

❖ Many aspects:

  ❖ homophily (selection & influence)

  ❖ confirmation bias

  ❖ back-fire effect

  ❖ hypercorrection effect

  ❖ bandwagon effect

  ❖ …

CASS R. SUNSTEIN

#re public

La democrazia nell'epoca dei social media

il Mulino

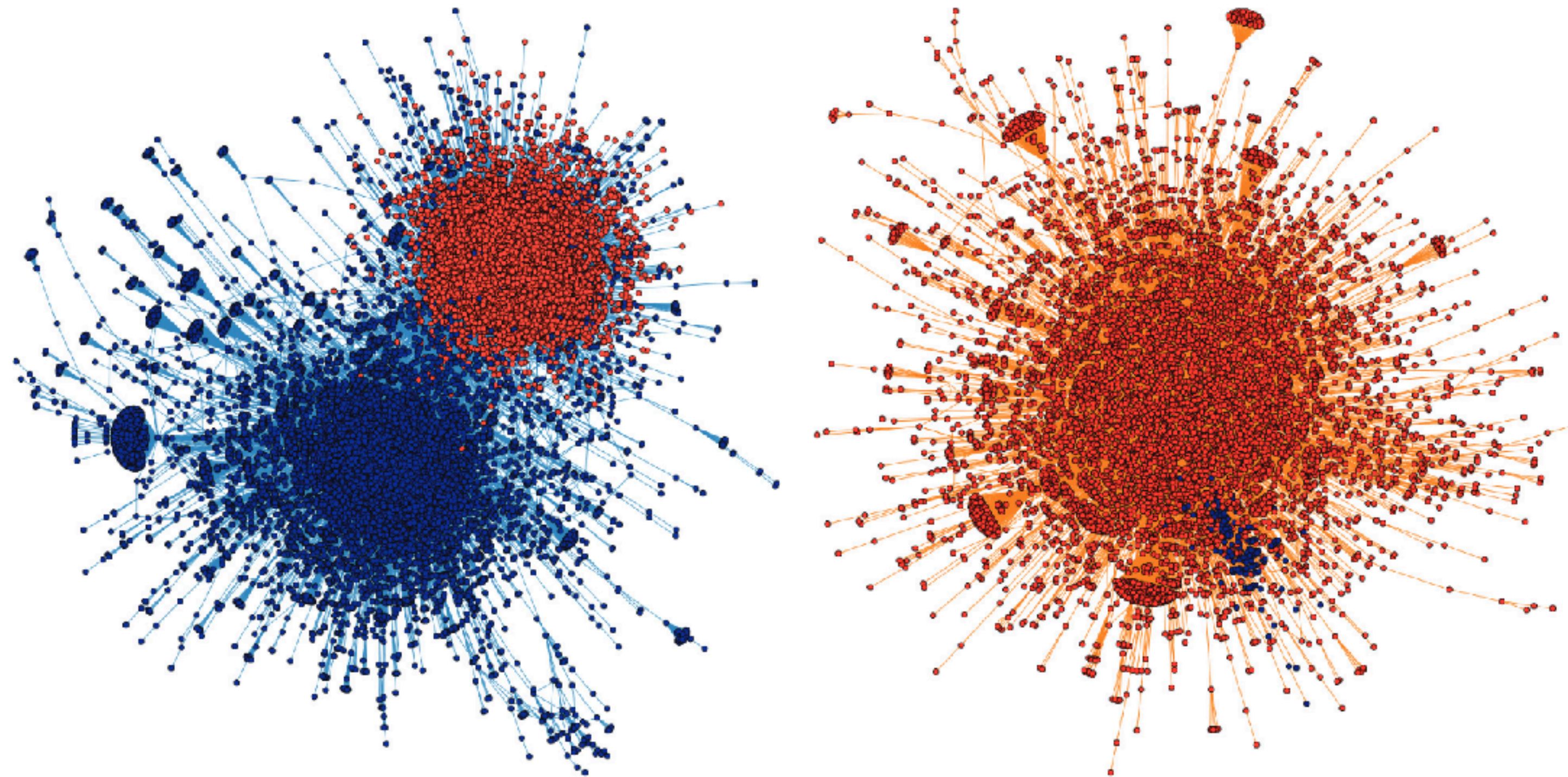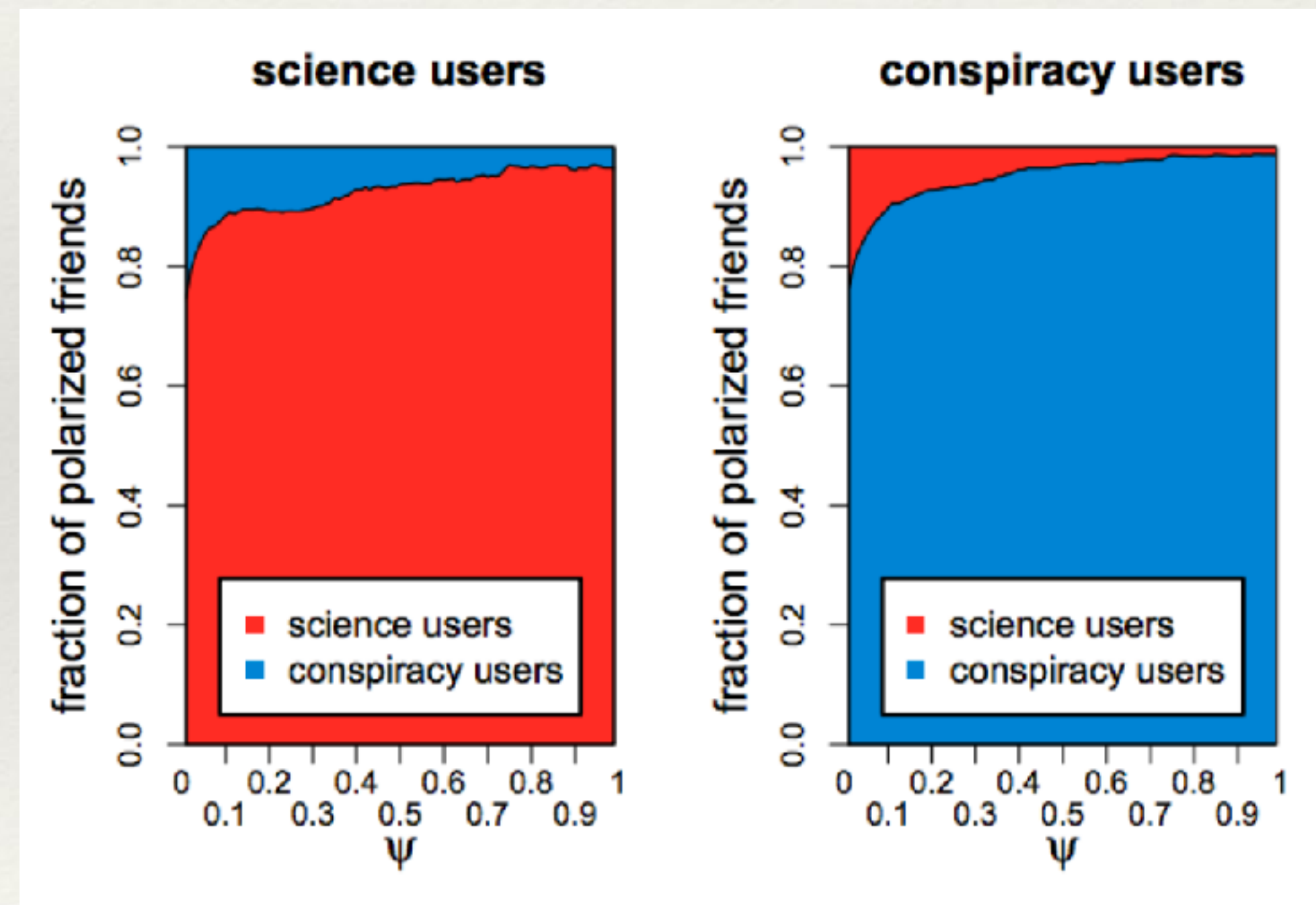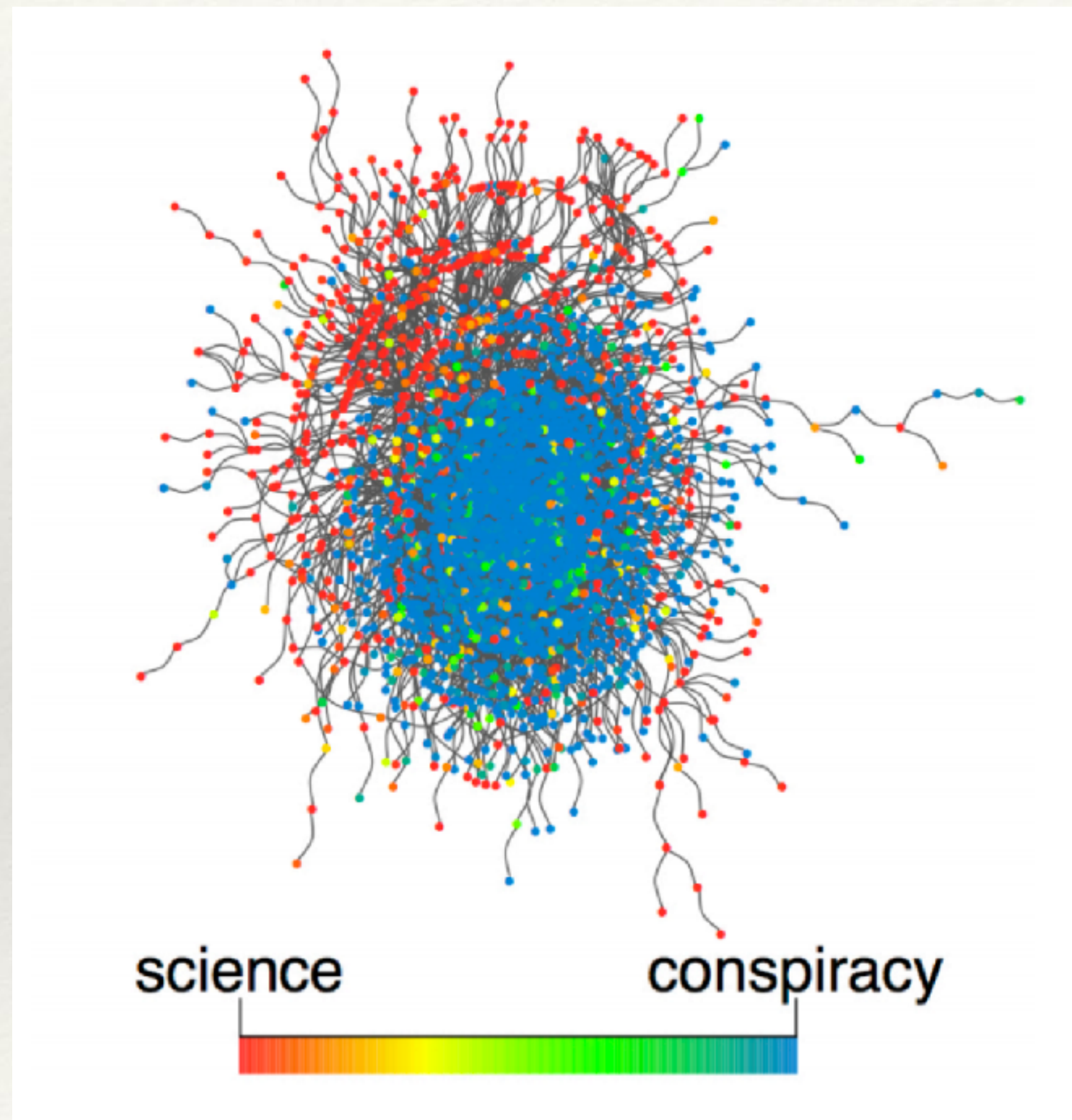# (political) polarization on social media

Figure 1: The political retweet (left) and mention (right) networks, laid out using a force-directed algorithm. Node colors reflect cluster assignments (see § 3.1). Community structure is evident in the retweet network, but less so in the mention network. We show in § 3.3 that in the retweet network, the red cluster A is made of 93% right-leaning users, while the blue cluster B is made of 80% left-leaning users.
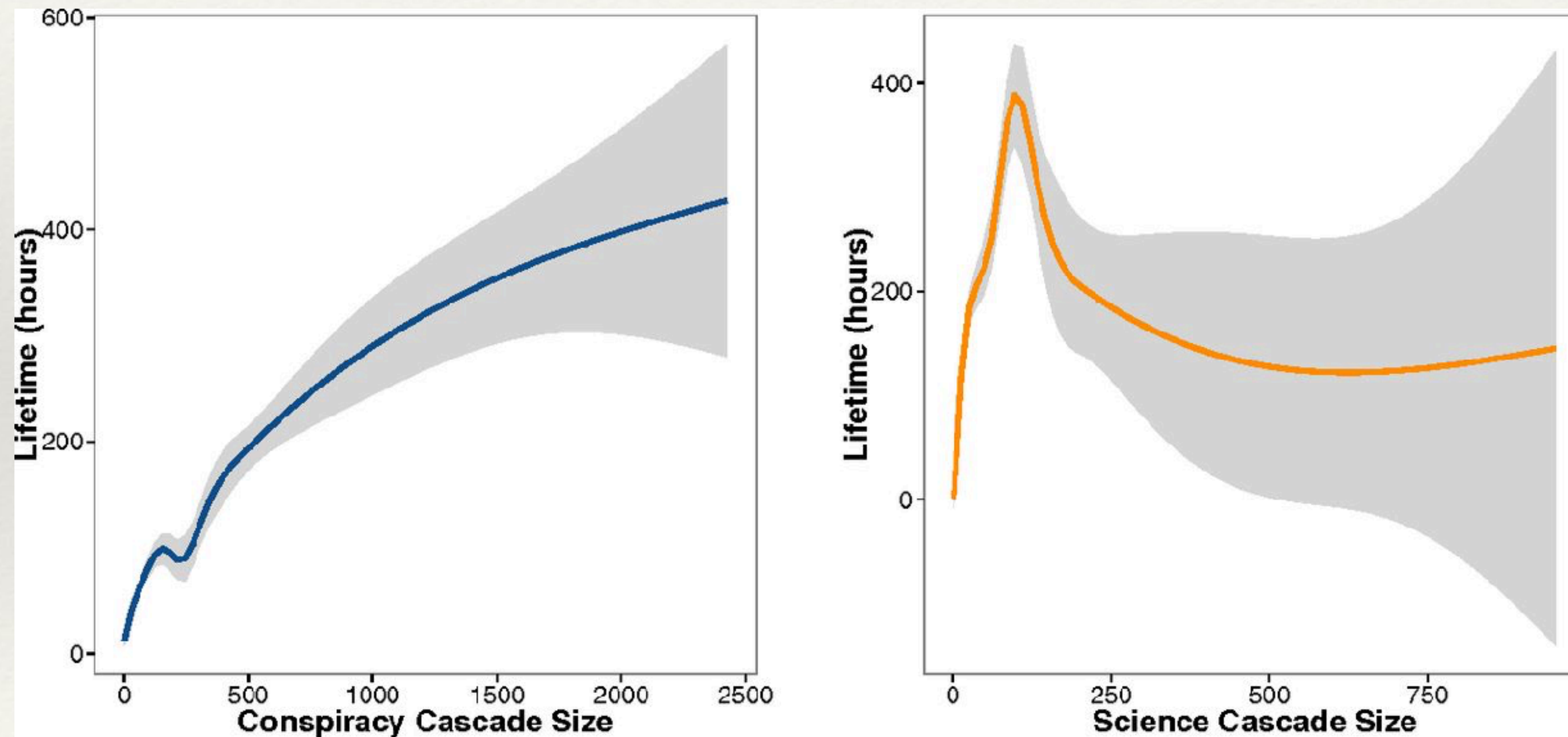
Conover, M. & Ratkiewicz, J. & Francisco, M. & Gonçalves, B. & Menczer, F. & Flammini, A. (2011). Political Polarization on Twitter. Fifth International AAAI Conference on Weblogs and Social Media

Misinformation tends to polarize

Users engagement correlates with the number of friends having similar consumption patterns

homophily!

A. Bessi, …., G. Caldarelli, W. Quattrociocchi, Viral Misinformation: The Role of Homophily and Polarization, WWW 2015 Companion, May 18–22, 2015, Florence, Italy.
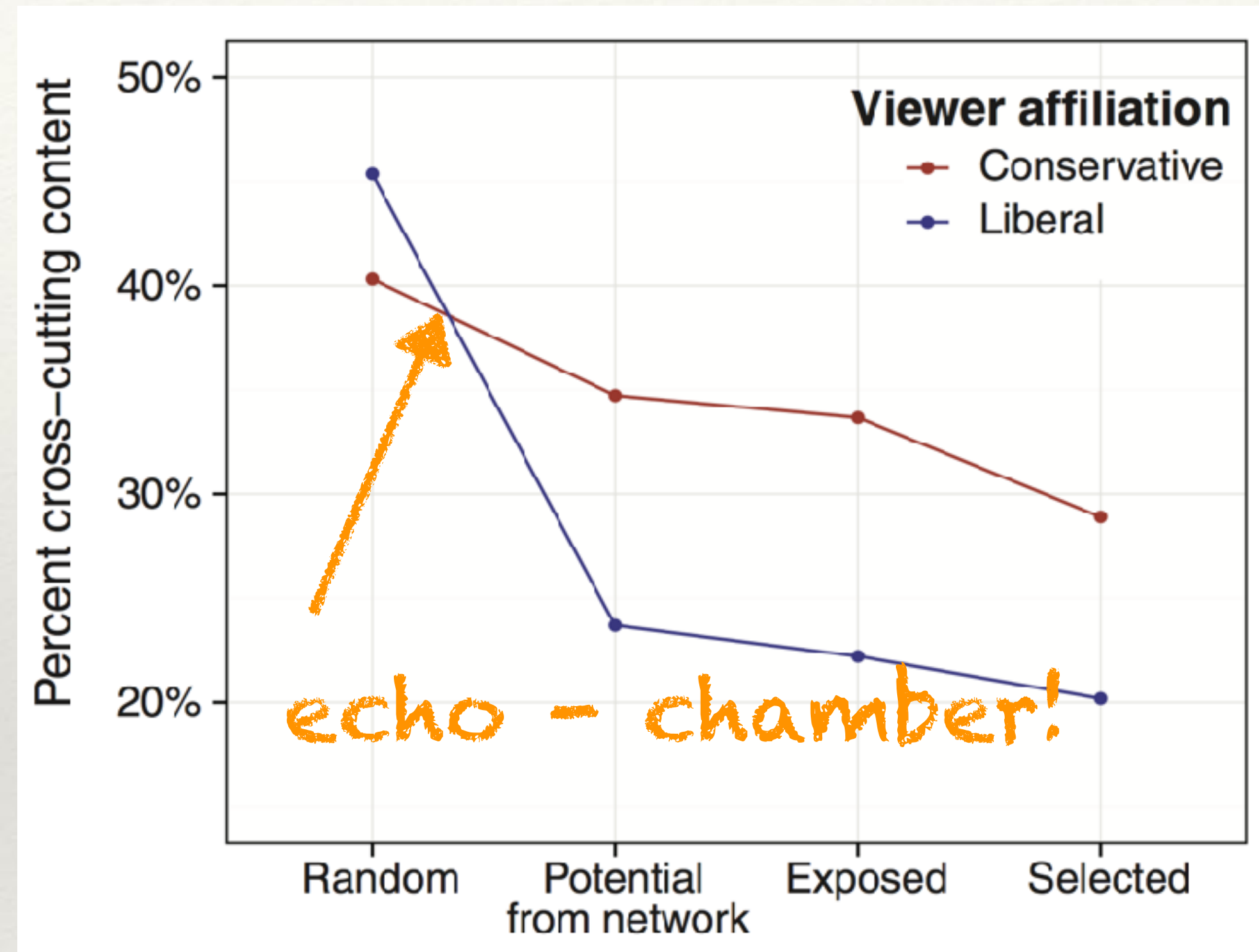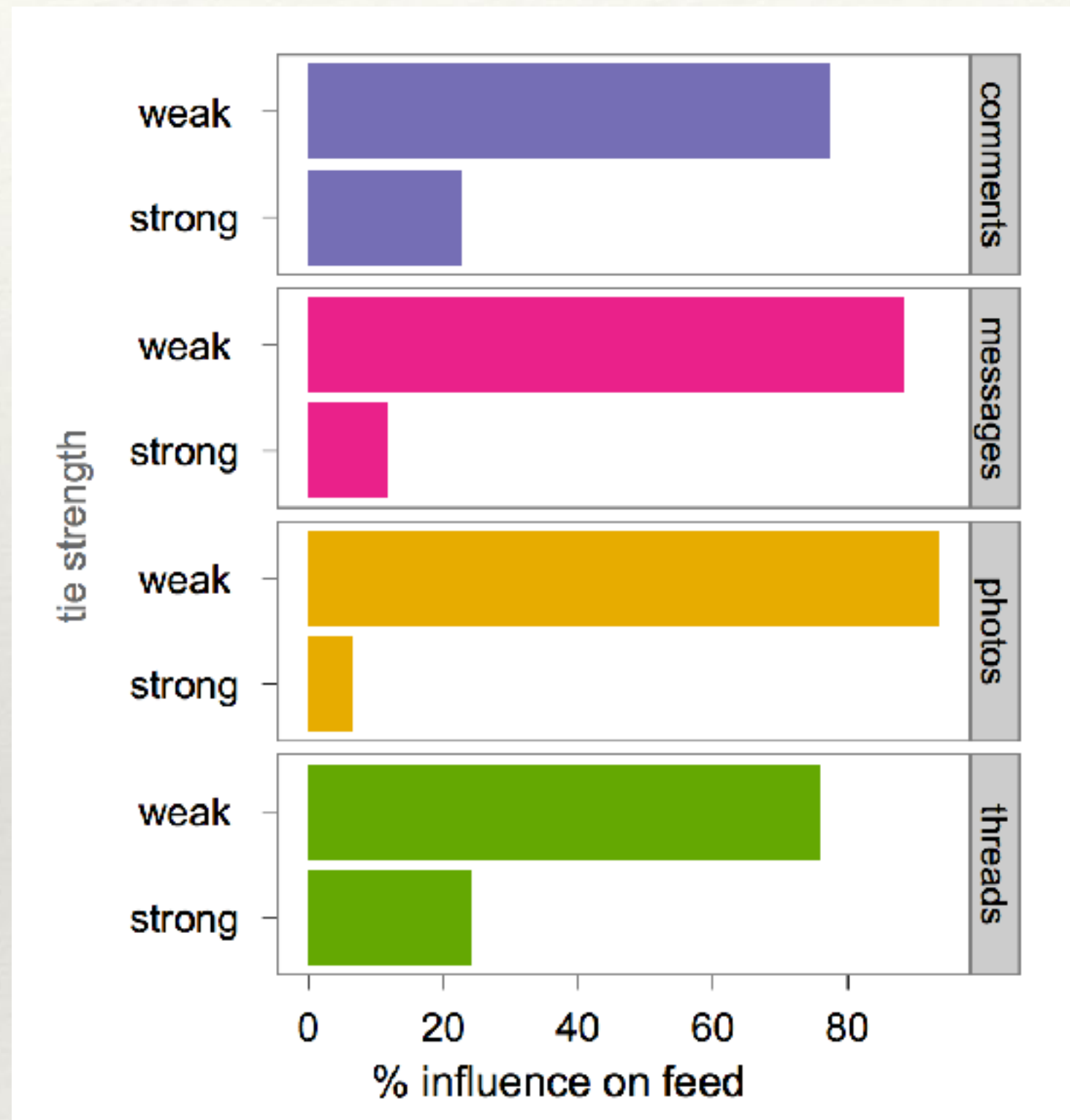
... and polarization fuels misinformation spread

A data-driven percolation model of rumor spreading that demonstrates that homogeneity and polarization are the main determinants for predicting cascades' size
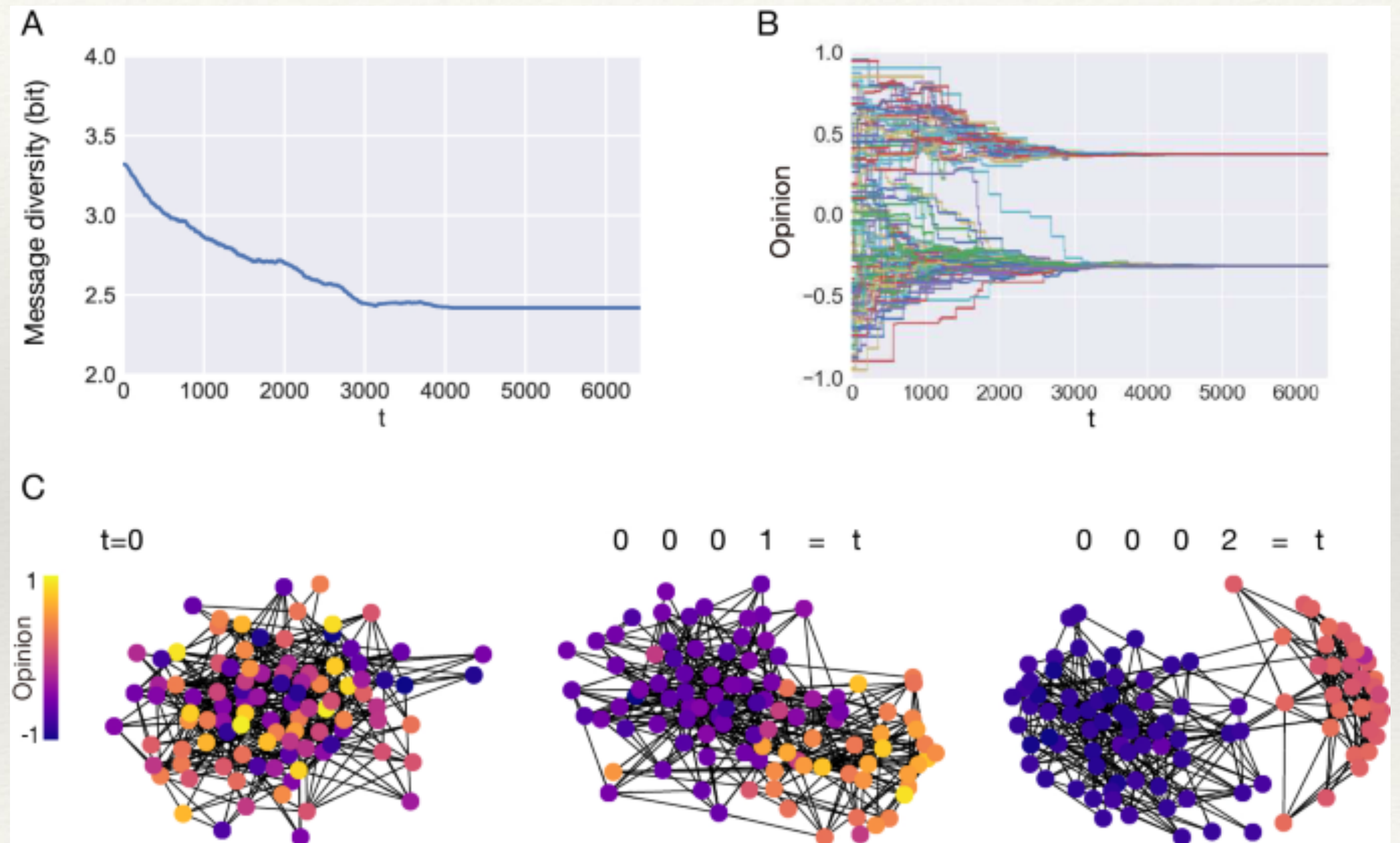

M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, W. Quattrociocchi, Echo chambers in the age of misinformation, PNAS, Jan 2016, 113 (3) 554-559; DOI: 10.1073/pnas.1517441113

"Weak ties" are important, too

E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic. 2012. The role of social networks in information diffusion. In Proc of the 21st Int. Conf. on World Wide Web (WWW '12). ACM, New York, NY, USA, 519–528. DOI:https://doi.org/10.1145/2187836.2187907
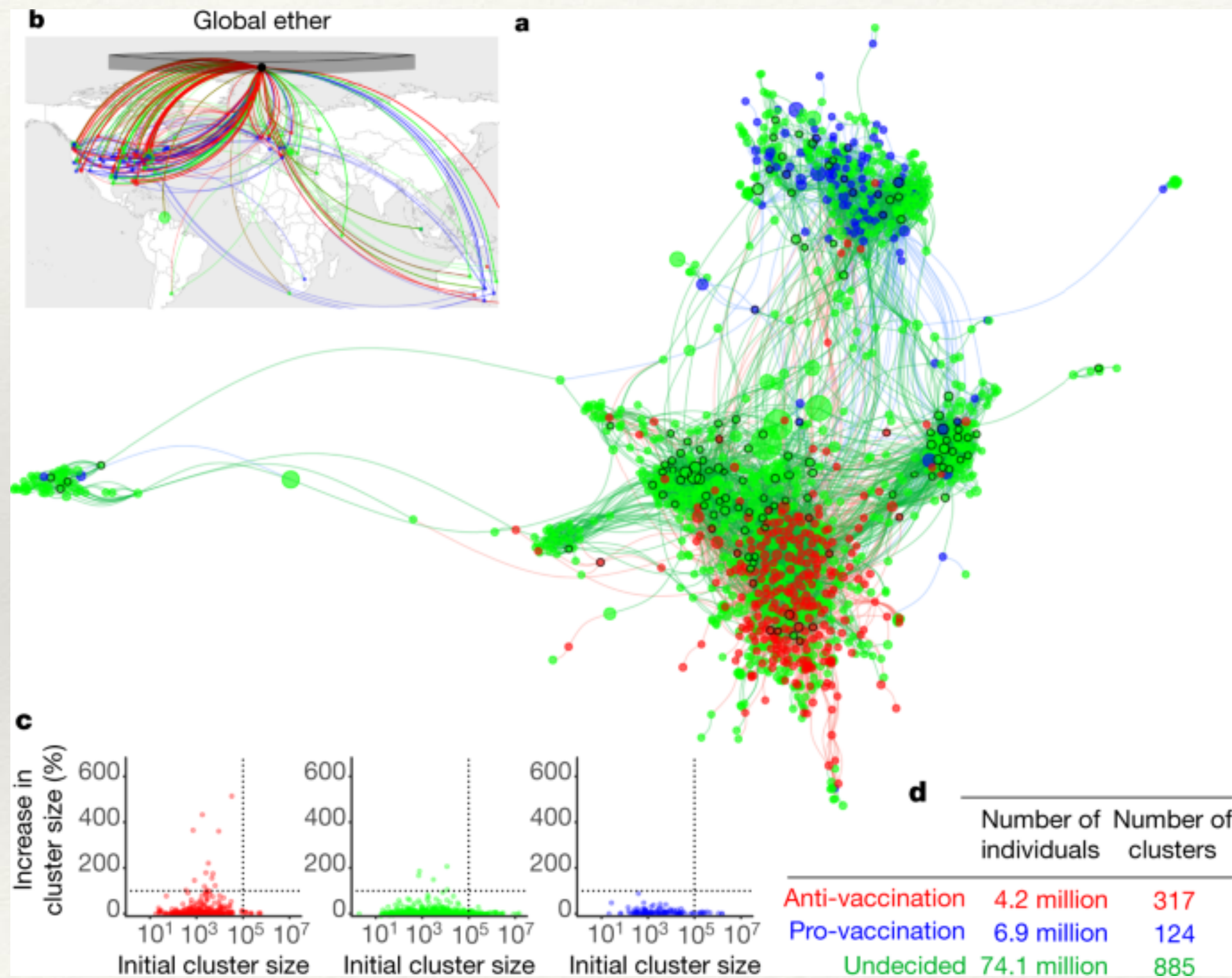
E. Bakshy, S. Messing, L. Adamic, Exposure to ideologically diverse news and opinion on Facebook, Science 05 Jun 2015: Vol. 348, Issue 6239, p. 1130-1132, DOI: 10.1126/science.aaa1160(Bakshy et al. 2015)
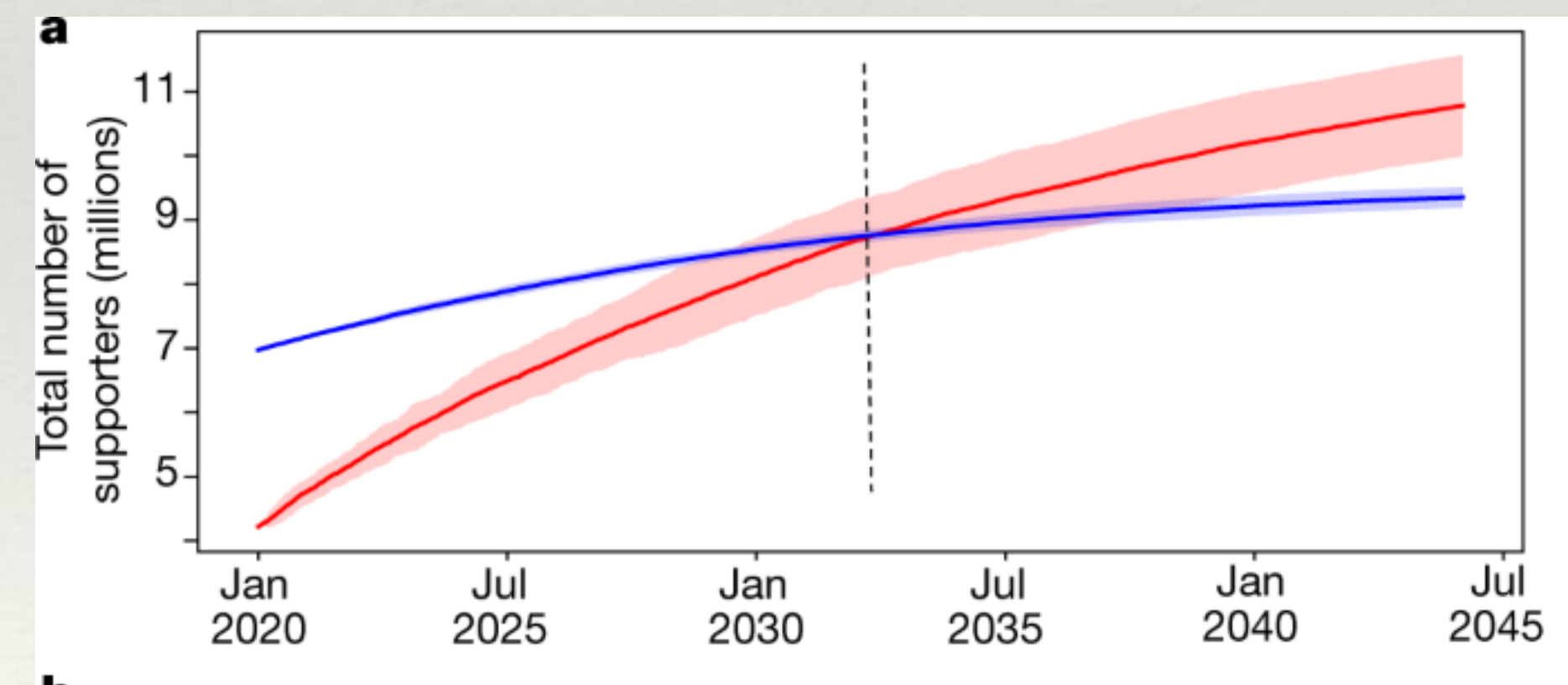
The role of unfollowing

❖ The **model dynamics** show that even with minimal amounts of **influence** and **unfriending**, the social network rapidly devolves into polarized communities

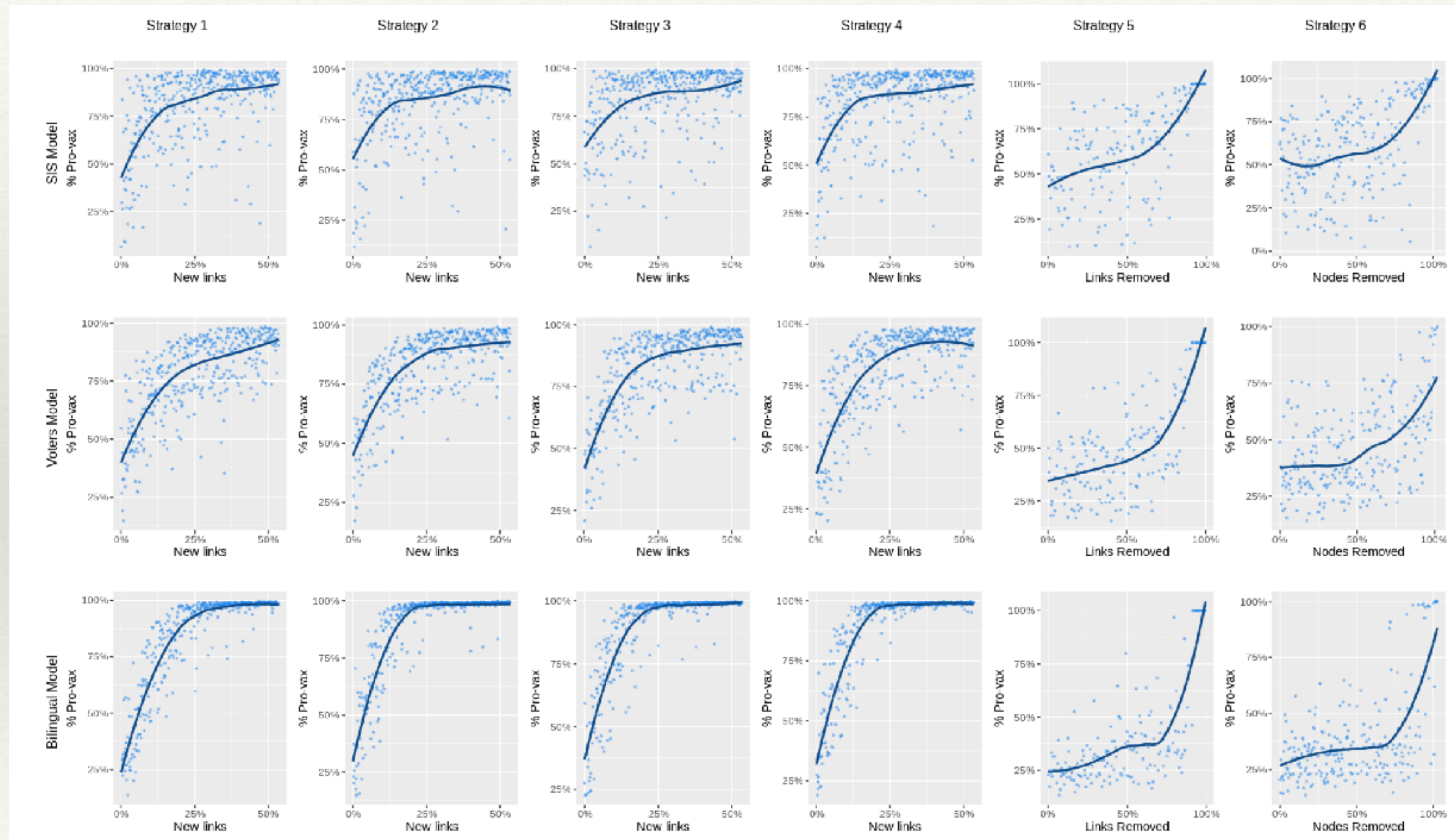❖ Predictions are consistent with **empirical data** from Twitter

Sasahara, K., Chen, W., Peng, H. Ciampaglia, G. L., Flammini, A., Menczer, F. Social influence and unfollowing accelerate the emergence of echo chambers. *J Comput Soc Sc* (2020).

The role of the undecided

- ❖ Theoretical prediction for the future total size of anti-vaccination and pro-vaccination support

- ❖ Under the present conditions, it predicts that total anti-vaccination support reaches dominance in around 10 years
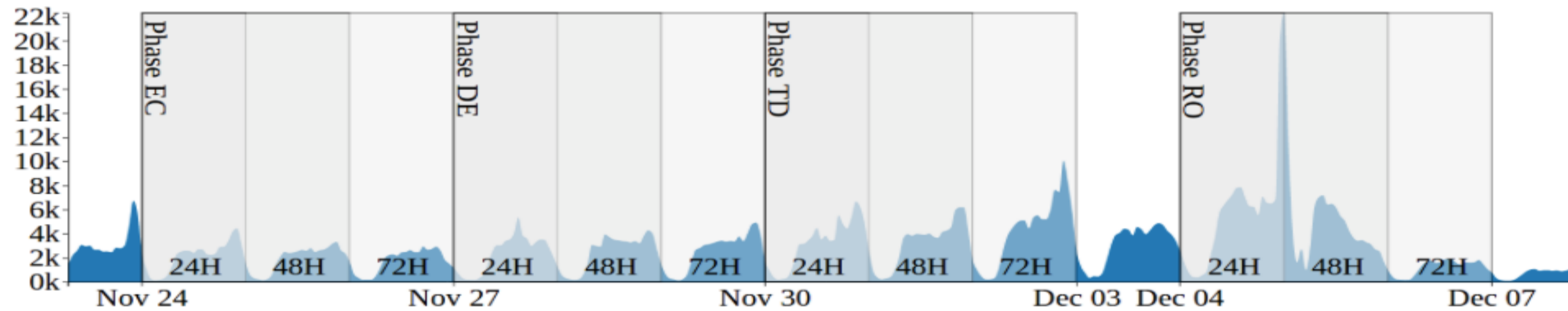
Johnson, N.F., Velásquez, N., Restrepo, N.J. *et al.* The online competition between pro- and anti-vaccination views. *Nature* **582,** 230–233 (2020).

The role of the undecided – cont'd

Lenti, J., Ruffo, G. Ensemble of opinion dynamics models to understand the role of the undecided about vaccines, *Journal of Complex Networks*, Volume 10, Issue 3 (2022).

Italian 2016 Constitutional Referendum

M Lai, M Tambuscio, V Patti, P Rosso, G. Ruffo, Stance Polarity in Political Debates: a Diachronic Perspective of Network Homophily and Conversations on Twitter, Data & Knowledge Engineering Journal, online: September 2019

Italian 2016 Constitutional Referendum

Collected Tweets

Reply-to Network

signal of **inverse homophily**

stance detected as **AGAINST**

stance detected as **IN FAVOR**

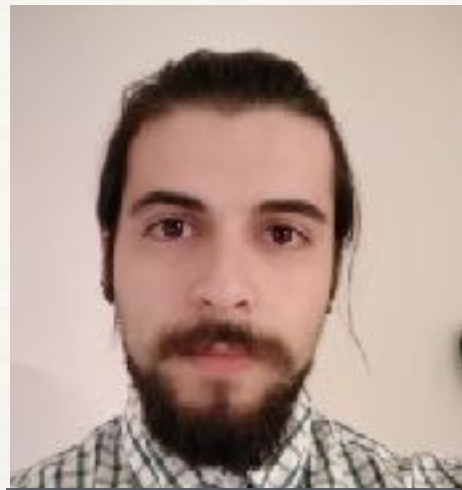stance detected as **NONE**

The interplay between polarization and misinformation spread

❖ Misinformation is fueled by group-thinking

❖ Unfriending accelerates polarization

❖ Building bridges is more effective than censoring controversial pages or unfriending

❖ In some domains, (the observation of) a polarizing debate can be extremely volatile - we live a dynamical world!

Some lessons learned

Take home messages

❖ The science of fake news is inherently multidisciplinary

❖ Data scientists have a central role in studying, modeling, analyzing

  ❖ ML, NLP, network analysis, statistics, diffusion models, …

❖ Real solutions are in education (long term), bot/troll detection and removal (short term), re-thinking of social media and Web outlets business models (desiderata…)

❖ Memento for professional journalists: **with great power comes great responsibility!**

  ❖ Is the design of novel (AI based) **editor assistants** a solution?

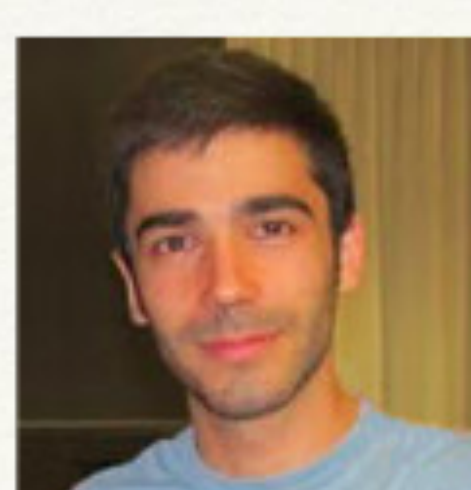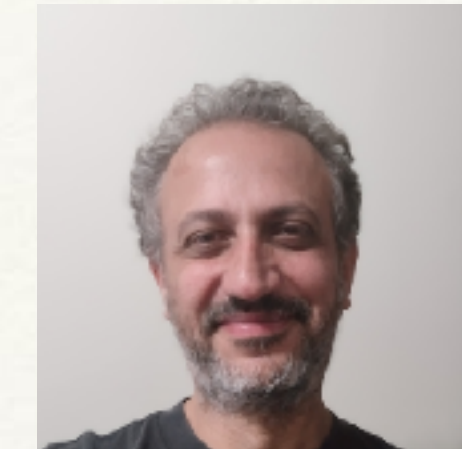❖ There are thousands of relevant papers published from 2015 on… good luck :)

ARC²S: Applied Research on Computational Complex Systems

Questions?