Investigating (dis)information diffusion in online social networks

Francesco Pierri
Data Science Lab, Department of Electronics, Information and Bioengineering, Politecnico di Milano

Politecnico di Milano
16/12/2024

ABOUT ME

Ph.D. in Data Analytics and Decision Sciences

Assistant Professor (Data Science Group)

Visiting scholar at IUNI and USC

Research interests: computational social science, online misinformation, generative AI

Salerno

THE RISE OF COMPUTATIONAL SOCIAL SCIENCE

- Unprecedented volume of data on human behavior and societal trends
- Rise of computational techniques to analyze and interpret digital trace data
- Social media platforms have become a primary source of data through their APIs
- Recent concerns about the limited availability and quality of such data

SOCIAL SCIENCE

Computational Social Science

David Lazer,[1] Alex Pentland,[2] Lada Adamic,[3] Sinan Aral,[2,4] Albert-László Barabási,[5] Devon Brewer,[6] Nicholas Christakis,[1] Noshir Contractor,[7] James Fowler,[8] Myron Gutmann,[3] Tony Jebara,[9] Gary King,[1] Michael Macy,[10] Deb Roy,[2] Marshall Van Alstyne[2,11]

A field is emerging that leverages the capacity to collect and analyze data at a scale that may reveal patterns of individual and group behaviors.

Source: Lazer et al. (2009) Science

POLICY FORUM

SOCIAL SCIENCE

Computational social science: Obstacles and opportunities

Data sharing, research ethics, and incentives must improve

By David M. J. Lazer[1,2], Alex Pentland[3], Duncan J. Watts[4], Sinan Aral[3], Susan Athey[5], Noshir Contractor[6], Deen Freelon[7], Sandra Gonzalez-Bailon[4], Gary King[2], Helen Margetts[8,9], Alondra Nelson[10,11], Matthew J. Salganik[12], Markus Strohmaier[13,14], Alessandro Vespignani[1], Claudia Wagner[14,15]

dependencies within data. A loosely connected intellectual community of social scientists, computer scientists, statistical physicists, and others has coalesced under this umbrella phrase.

MISALIGNMENT OF UNIVERSITIES

Source: Lazer et al. (2020) Science

MISINFORMATION ABOUT MISINFORMATION

Misinformation

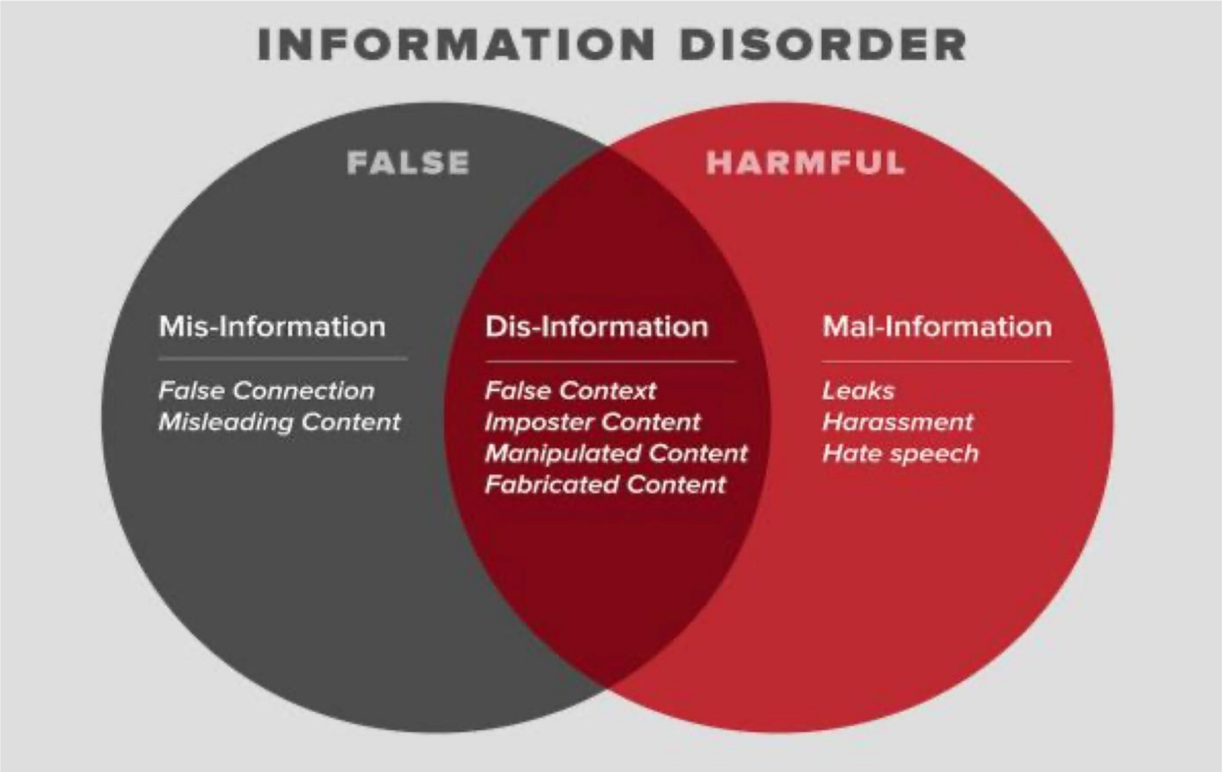
Disinformation

Propaganda

Clickbait

Junk News

Fake News



Source: Wardle et al. Council of Europe (2017)

SOCIAL MEDIA AND MISINFORMATION

No quality control

Superspreaders

Echo chambers

Fake accounts and social bots

Cognitive biases

Source: PlaygroundAI

RECENT CONCERNS

No access to relevant social media data

Less and less content moderation

New generative AI

Researchers being (politically) harassed

Source: NYTIMES

The New York Times

G.O.P. Targets Researchers Who Study Disinformation Ahead of 2024 Election

The Washington Post

Stanford's top disinformation research group collapses under pressure

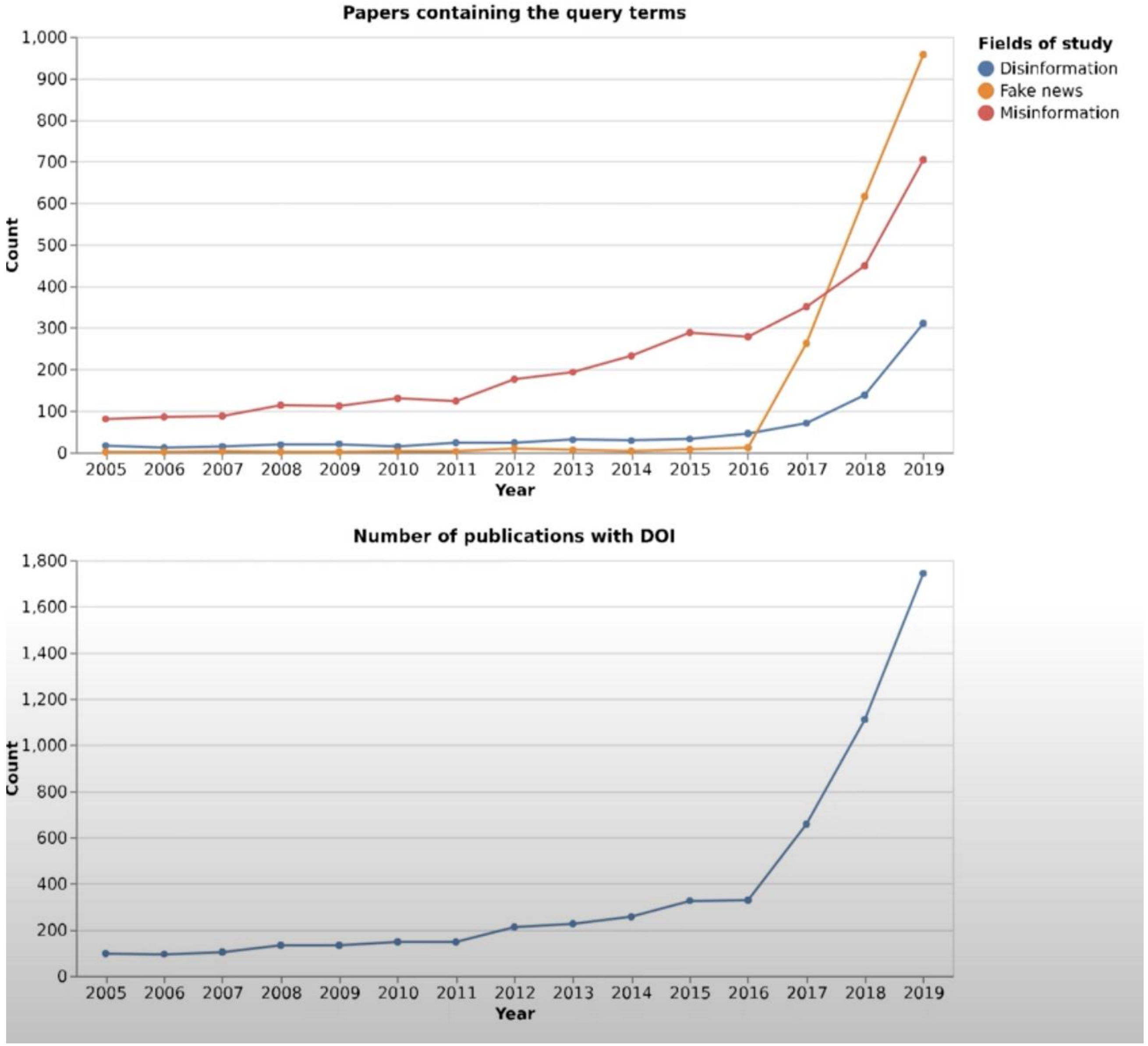
The Stanford Internet Observatory provided real-time analysis on viral election falsehoods but has struggled amid attacks from conservative politicians and activists.

A NEW FIELD OF STUDY

"Fake News" appeared as a field of study after the 2016 US elections

Exponential growth of papers over time

Similar patterns with COVID-19 and LLMs



Source: Stathoulopoulos et al. (2020)

THE SPREAD OF TRUE AND FALSE NEWS ONLINE

Rumors spreading on Twitter (2006-2017)

Network science and statistical analysis

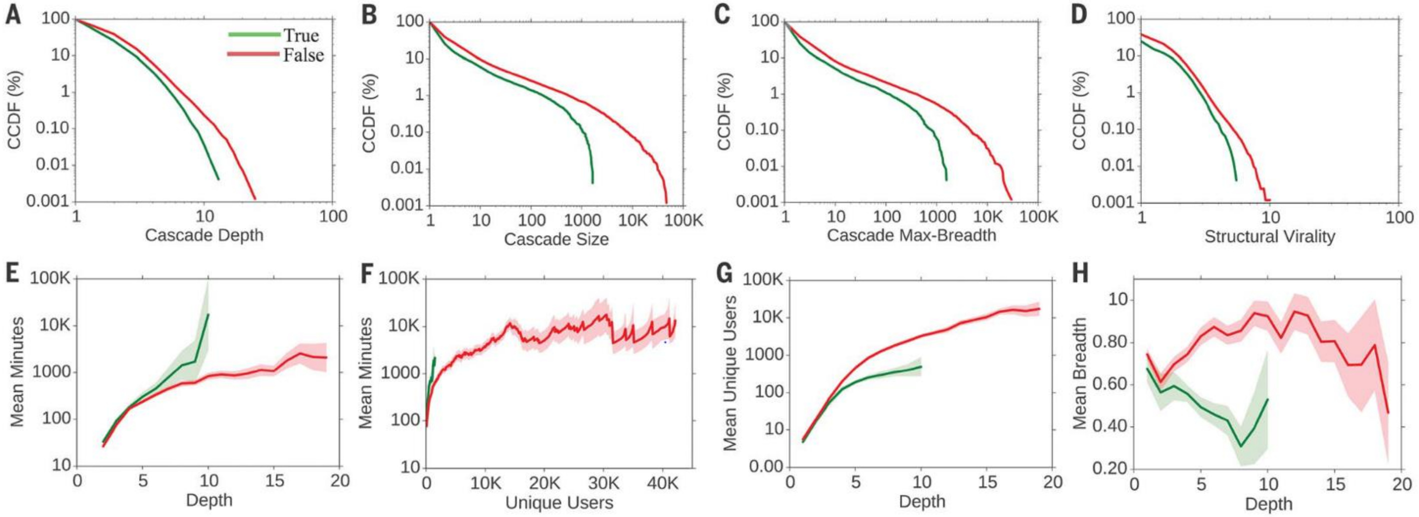
False rumors spread faster, deeper and broader than true rumors

RESEARCH

SOCIAL SCIENCE

The spread of true and false news online

Soroush Vosoughi,[1] Deb Roy,[1] Sinan Aral[2]*



SPREADING DYNAMICS BY MATCHING CASCADE SIZE

Statistical differences in cascade properties are challenging due to dependencies among these properties

Important to control for cascade size

No more differences between true and false rumors!

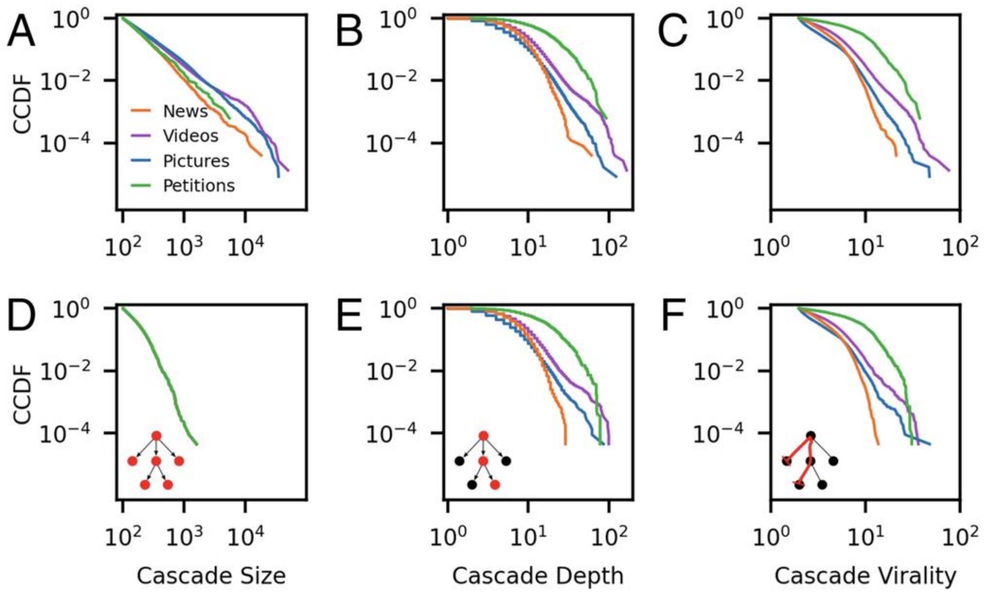
PNAS RESEARCH ARTICLE | APPLIED MATHEMATICS

Comparing information diffusion mechanisms by matching on cascade size

Jonas L. Juul[a,1] and Johan Ugander[b,1]

Edited by Duncan Watts, University of Pennsylvania, and accepted by Editorial Board Member Adrian E. Raftery August 30, 2021 (received for review January 14, 2021)

November 8, 2021 | 118 (46) e2100786118 | https://doi.org/10.1073/pnas.2100786118



FAKE NEWS AND SOCIAL BOTS

Twitter conversations around 2016 US elections

Bot detection and network dismantling

Most central users are likely to be bots

Very active in the early spread of misinformation news

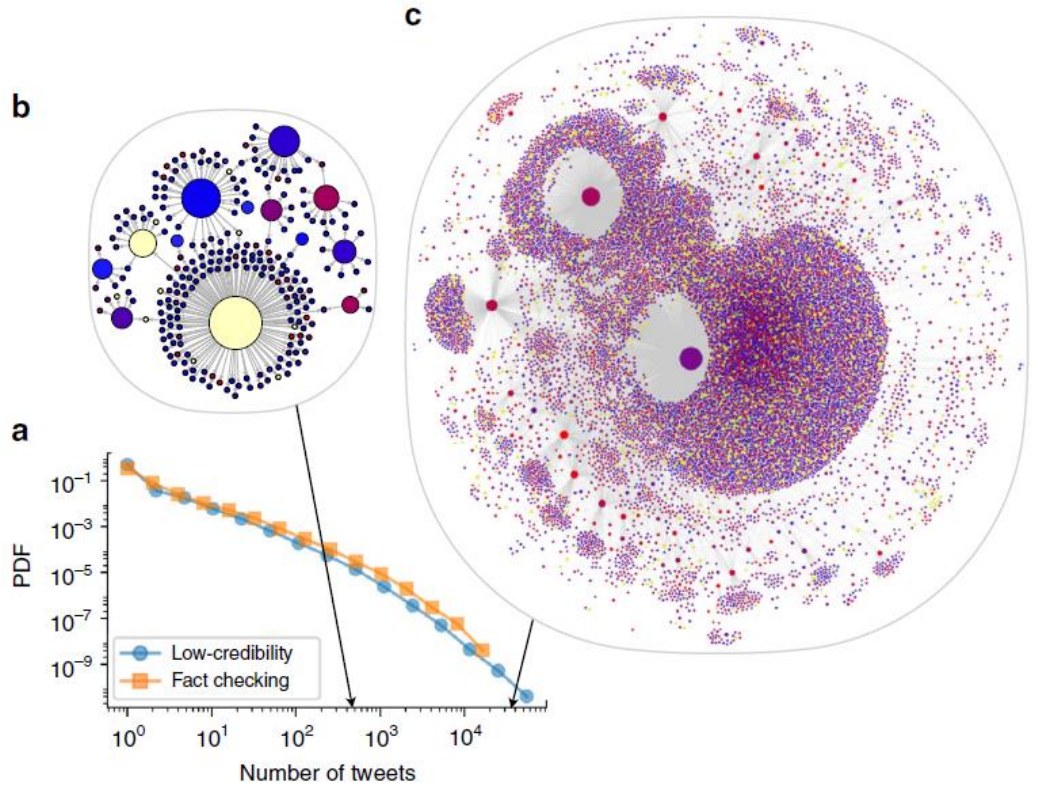
NATURE COMMUNICATIONS | (2018)

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 Yang[1], Alessandro Flammini[1,3] & Filippo Menczer[1,3]



FAKE NEWS AND POLITICAL SUPPORTERS

Twitter conversations around 2016 US elections

Influence maximization and causality analysis

Clinton supporters drive mainstream news

Trump supporters drive fake news spreaders

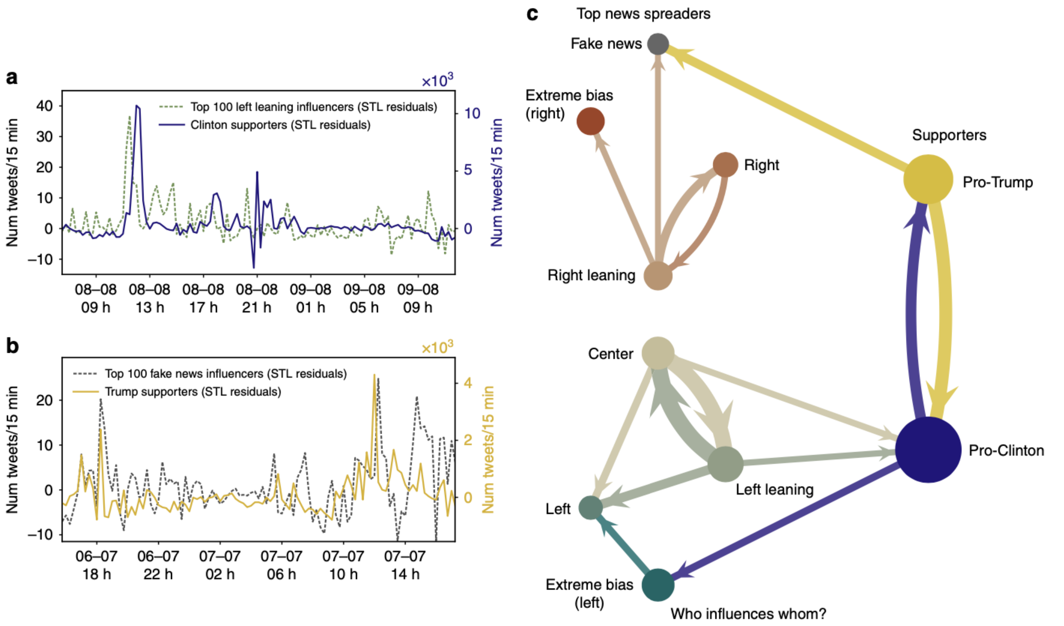
NATURE COMMUNICATIONS | (2019)

ARTICLE

https://doi.org/10.1038/s41467-018-07761-2 OPEN

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

Alexandre Bovet[1,2,3] & Hernán A. Makse[1]



CHALLENGES OF RESEARCH ABOUT MISINFORMATION

Adversarial setting

Massive volumes and velocity

Low-availability of data by platforms

Censorship concerns about potential intervention

Source: chatGPT

ITALIAN MISINFORMATION BEFORE 2019 EU ELECTION

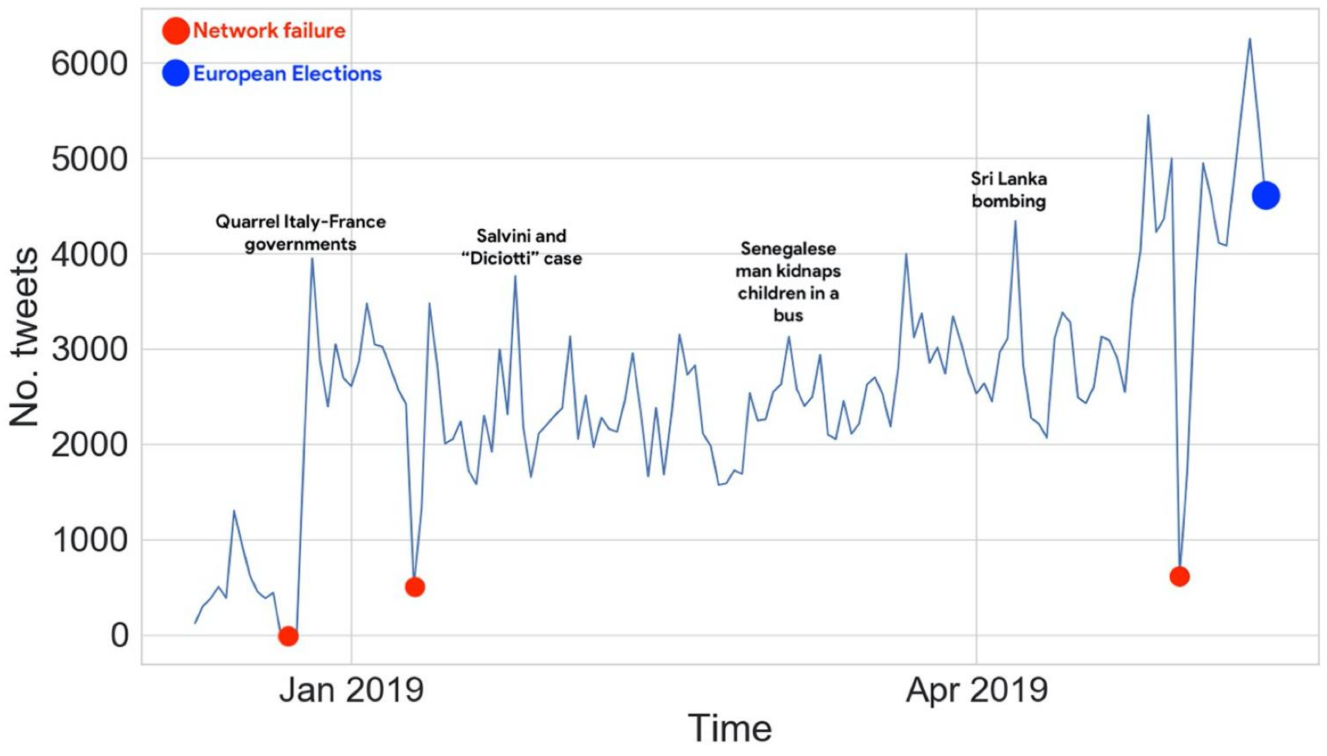
Network-based analysis

Monitoring low-credibility sources on Twitter

Controversial topics

Links to the far-right community

Limited volume



PLOS ONE

RESEARCH ARTICLE

Investigating Italian disinformation spreading on Twitter in the context of 2019 European elections

Francesco Pierri*, Alessandro Artoni, Stefano Ceri

IDENTIFYING MISINFORMATION AT SCALE

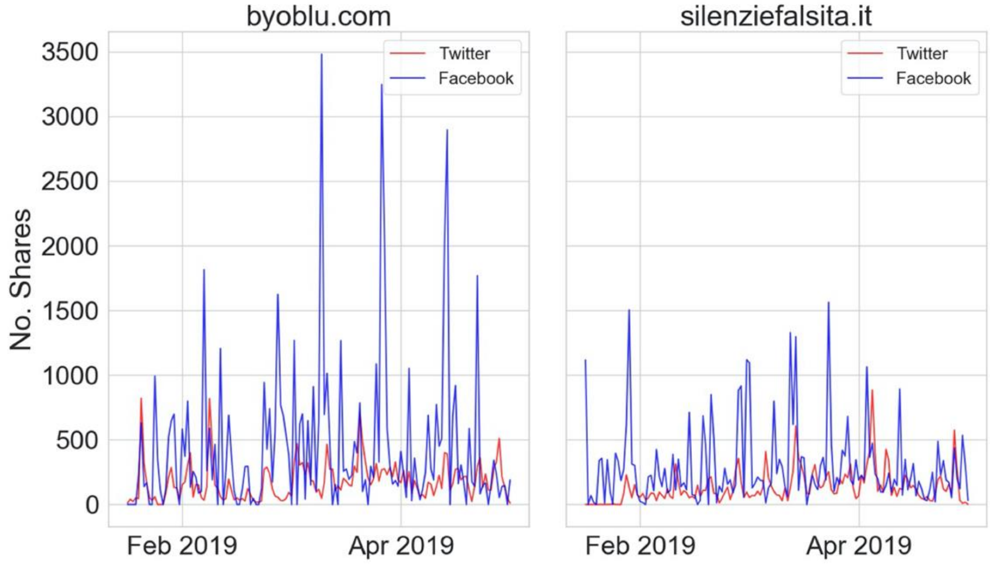
Disinformation, misinformation, false news, click-bait, conspiracy theories, unverified rumours, etc.

63 Italian news websites flagged by fact-checkers and journalists

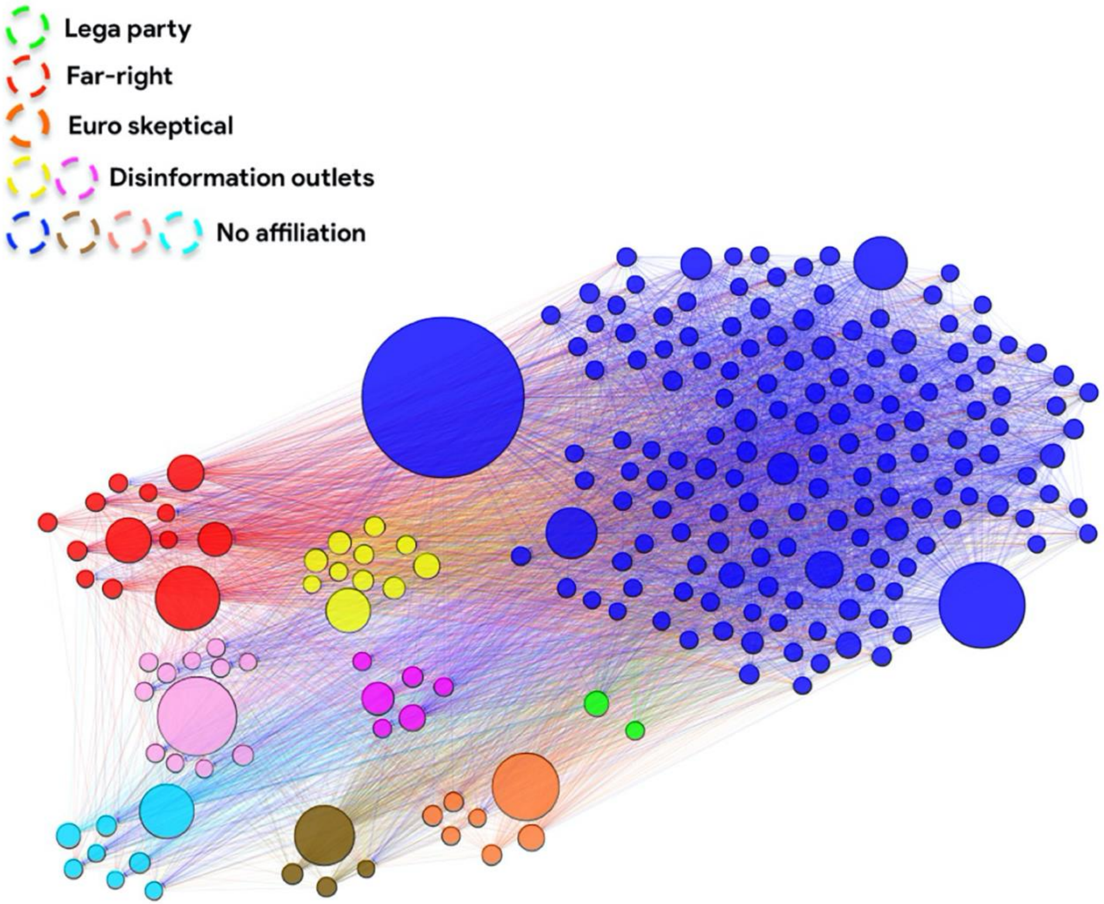
300k tweets, 16k articles, 20k unique users

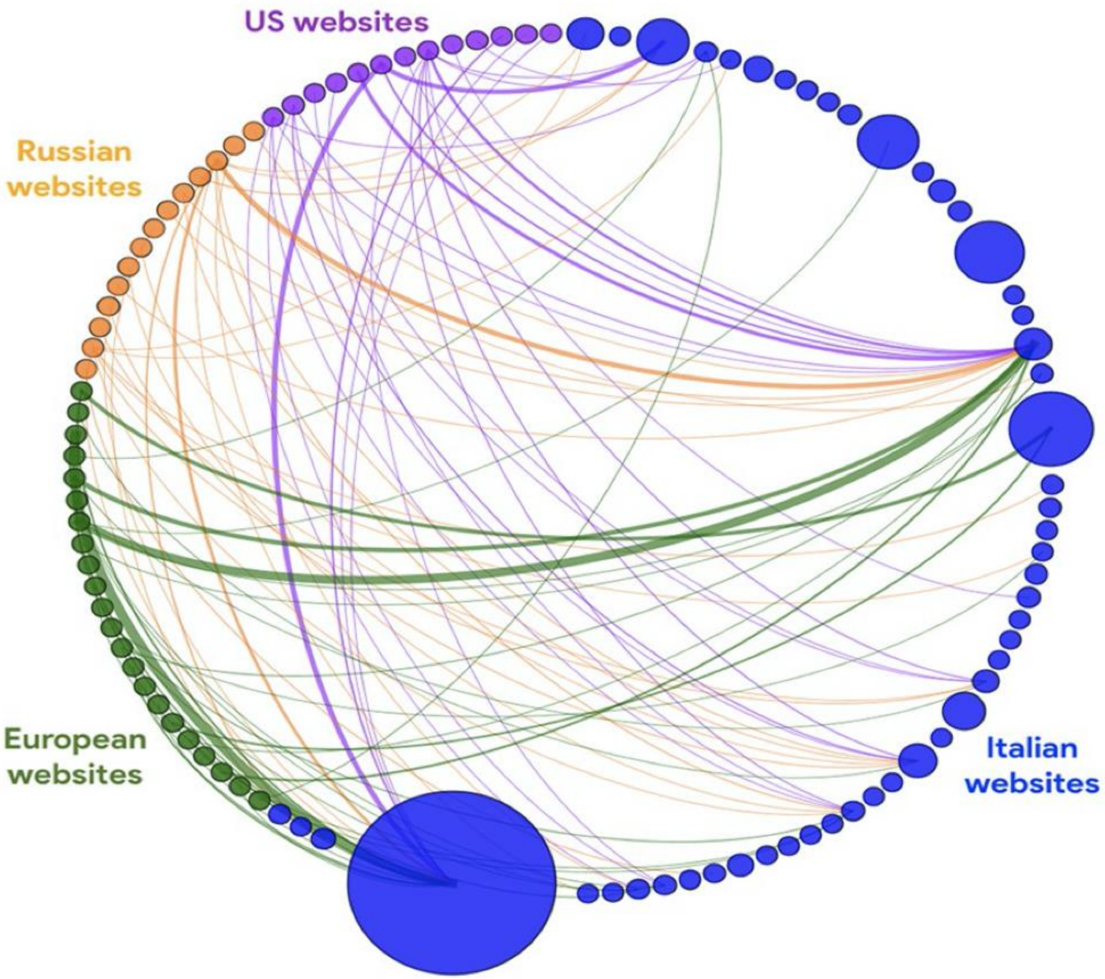
TOP SOCIAL MEDIA AND MESSAGING

Rank	Brand	For News	For All
1	Facebook	54% (+3)	77%
2	WhatsApp	27% (+2)	78%
3	YouTube	25% (-)	69%
4	Instagram	13% (+6)	41%
5	Facebook Messenger	8% (-)	40%
6	Twitter	8% (-2)	19%



NETWORKS OF MISINFORMATION USERS AND WEBSITES





IDENTIFYING DISINFORMATION USING TWITTER NETWORKS

Data Collection

Mainstream and Disinformation news

→

Data Representation

Twitter diffusion networks

→

Data Analysis

Network comparison

→

Final Results

Binary classification

SCIENTIFIC REPORTS
natureresearch

Topology comparison of Twitter diffusion networks effectively reveals misleading information

Francesco Pierri, Carlo Piccardi & Stefano Ceri*

HOW TO COMPARE NETWORKS?

1. **Global Network Properties:**

   Connected components, average clustering coefficient, main K-core number, etc…

2. **Node Centrality Measures Distributions**:

   Degree, eigenvector, betweenness, etc…

3. **Network Distances**:

   Portrait divergence, directed graphlets correlation

Low Complexity

High Complexity

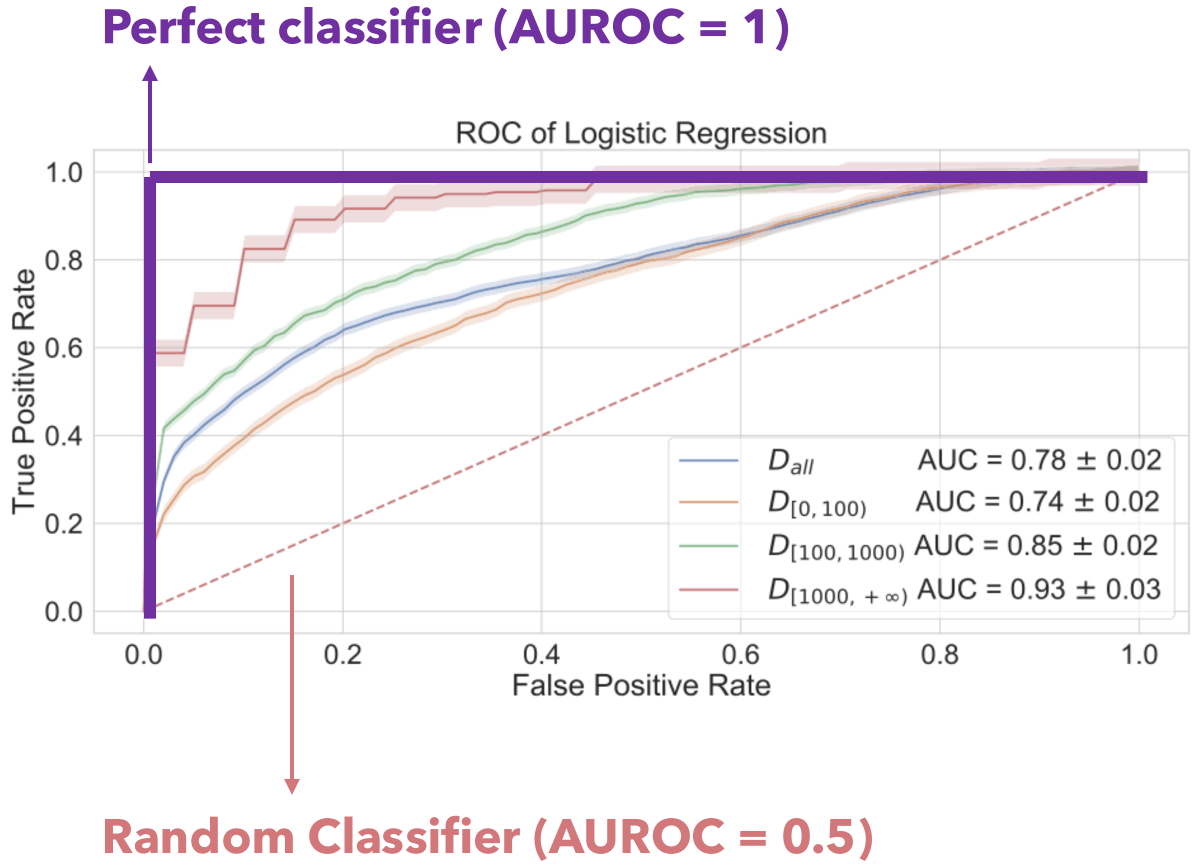
CLASSIFICATION RESULTS

High accuracy, especially for larger networks

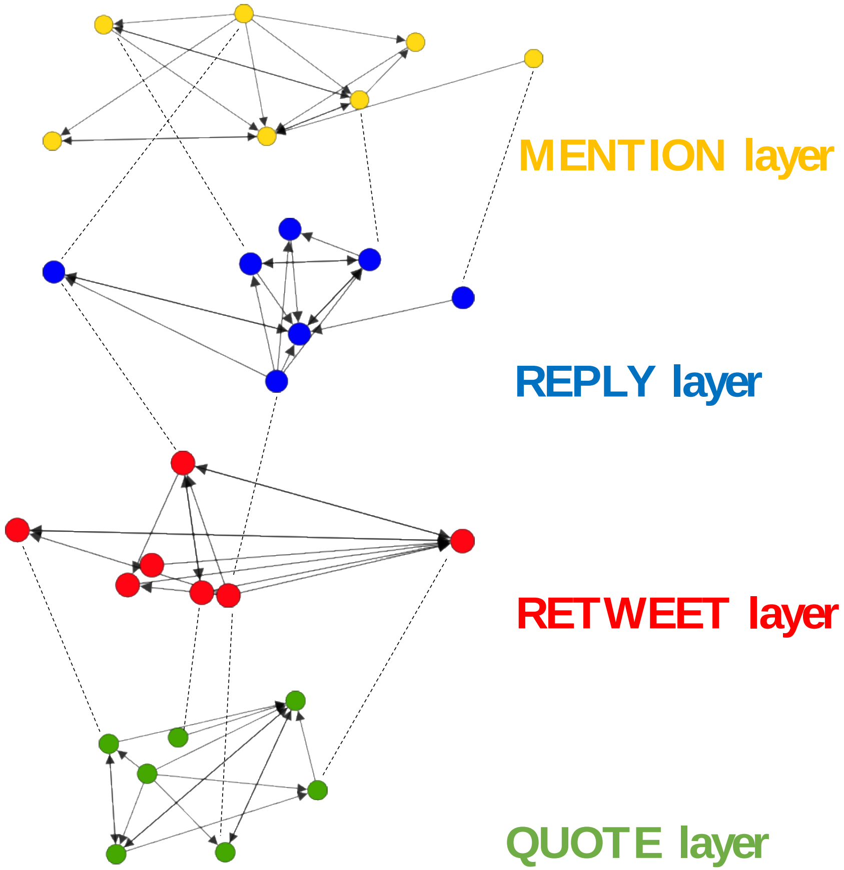
Topological features are interpretable

Disinformation is clustered

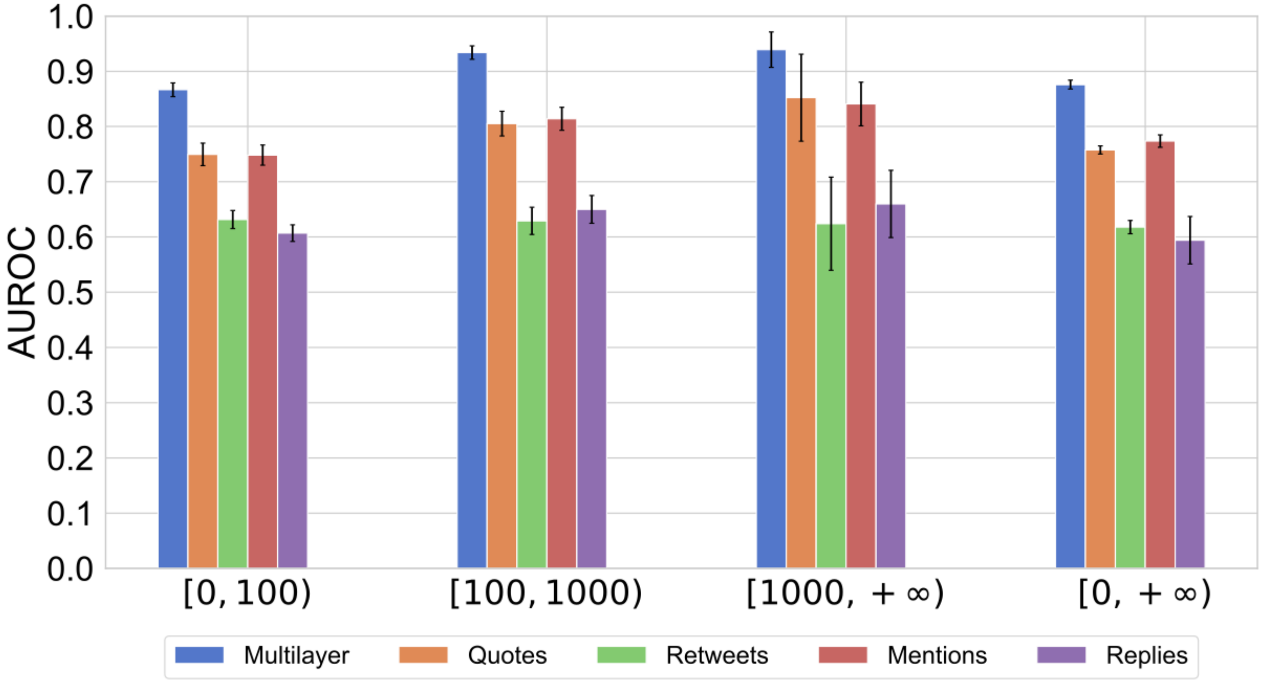
Mainstream news are broadcasted



IDENTIFYING DISINFORMATION USING MULTILAYER TWITTER NETWORKS




Pierri et al. EPJ Data Science (2020) 9:35
https://doi.org/10.1140/epjds/s13688-020-00253-8
EPJ.org
EPJ Data Science
a SpringerOpen Journal
REGULAR ARTICLE
Open Access
A multi-layer approach to disinformation detection in US and Italian news spreading on Twitter
Francesco Pierri[1*], Carlo Piccardi[1] and Stefano Ceri[1]



QANON CONSPIRACY

Melting pot of conspiracy theories (Pizzagate, Obamagate)

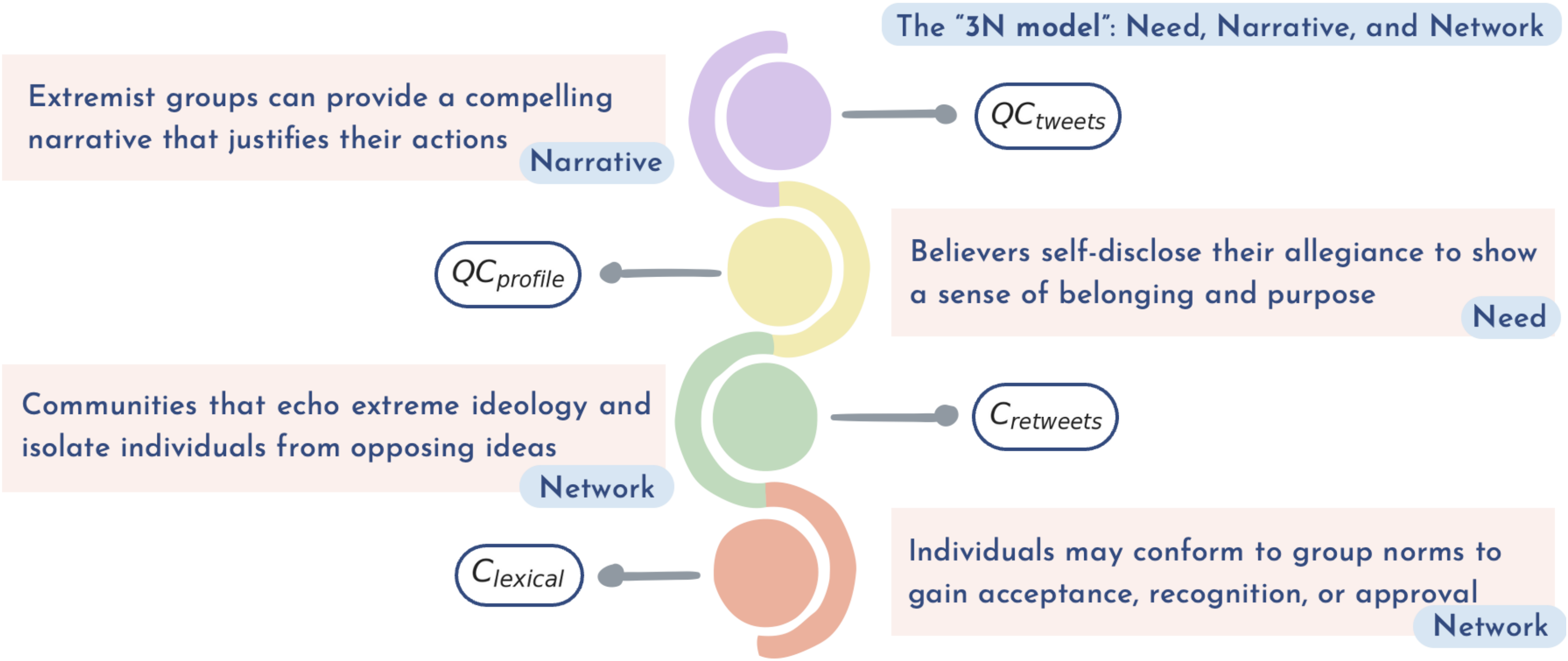
From 4chan/8kun to Twitter, Facebook and YouTube

106 US candidates endorsed QAnon in 2020

Real-world violence

Source: CNN

A MULTIVARIATE METRIC FOR QANON ENGAGEMENT



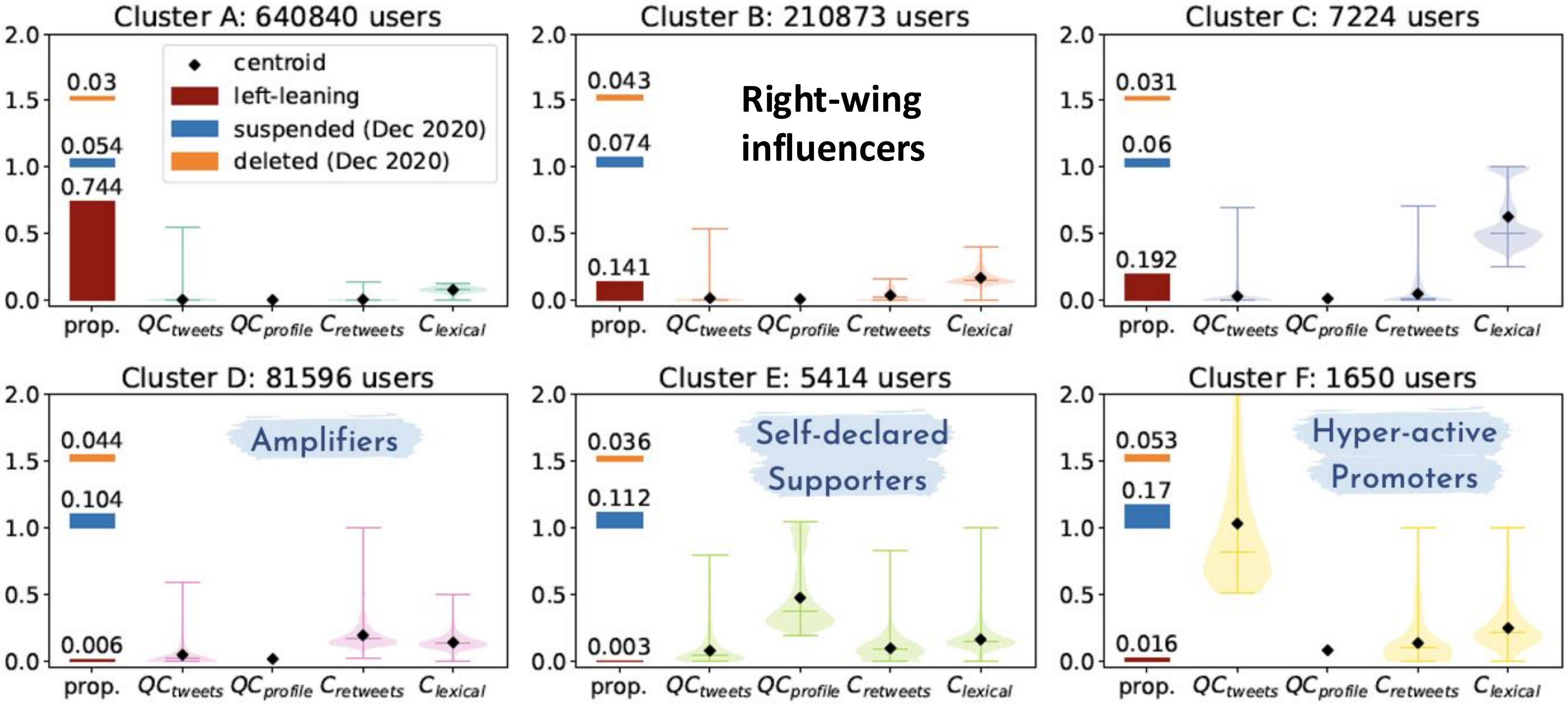
THE 17TH INTERNATIONAL AAAI CONFERENCE ON WEB AND SOCIAL MEDIA

Information Sciences Institute

Identifying and Characterizing Behavioral Classes of Radicalization within the QAnon Conspiracy on Twitter

Emily L. Wang[1,2], Luca Luceri[1], Francesco Pierri[1,3], Emilio Ferrara[1]

CLASSES OF RADICALIZATION ON TWITTER



THE 17TH INTERNATIONAL AAAI CONFERENCE ON WEB AND SOCIAL MEDIA

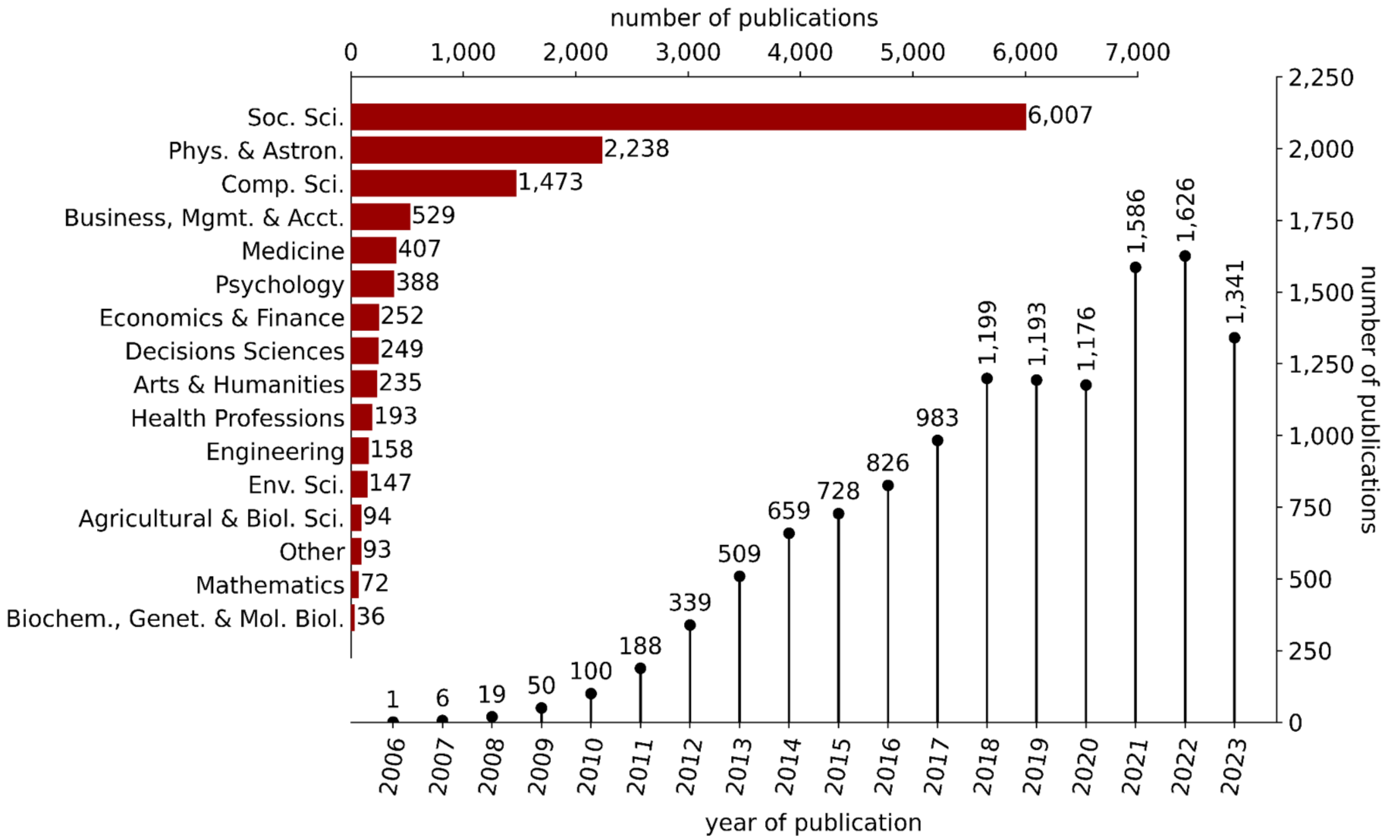
Information Sciences Institute

Identifying and Characterizing Behavioral Classes of Radicalization within the QAnon Conspiracy on Twitter

Emily L. Wang[1,2], Luca Luceri[1], Francesco Pierri[1,3], Emilio Ferrara[1]

INFORMATION DIFFUSION ON SOCIAL MEDIA

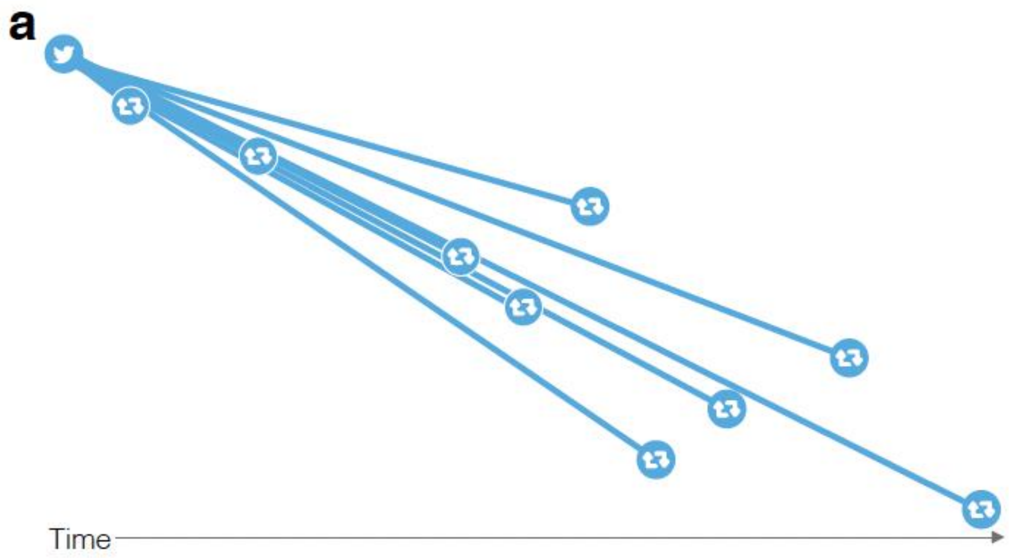
- The study of information diffusion using social media data has exploded over the last 20 years
- Many researchers employ simplifying assumptions about that data
- Can those assumptions affect our understanding of social networks?



Source: DeVerna et al. (2024)

RESHARING CASCADES IN MICROBLOGGING PLATFORMS

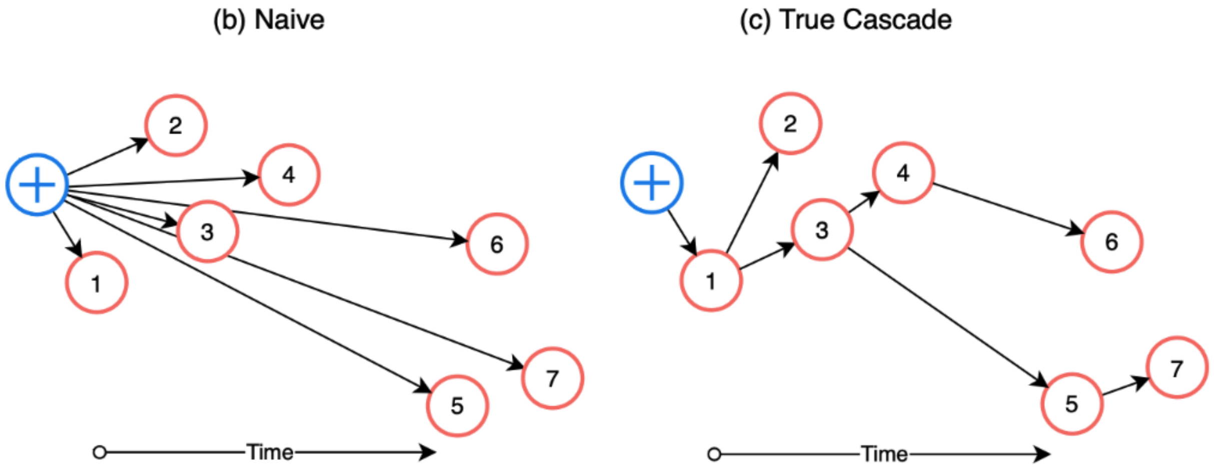
- Some platforms (e.g., Facebook) only provide aggregated data about cascade sizes
- Microblogging platforms (e.g., X) provide more data but attributes all resharing actions to the original poster
  - Star topology
- This misrepresentation hides the true dynamics of how information spreads



Source: Vosoughi et al. (2018) *Science*

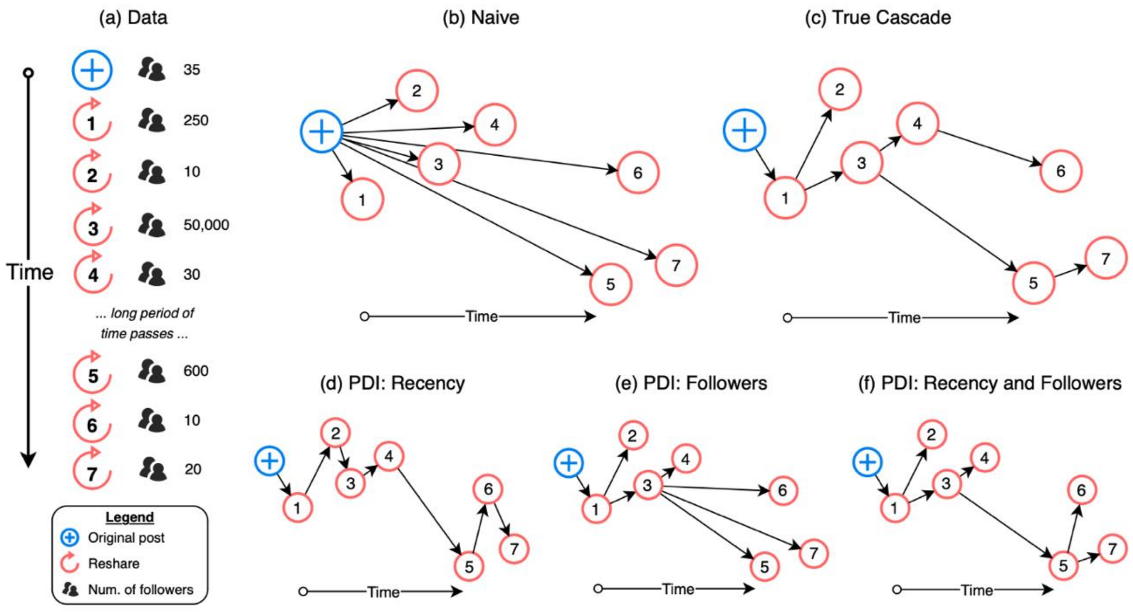
CONCERNS IN RECONSTRUCTING CASCADES

- We study how bypassing the cascade reconstruction process altogether impacts social influence analyses
  - Twitter and Bluesky resharing networks
- We then investigate the structural effects of different reconstruction approaches
  - A widely studied dataset of over 100,000 Twitter true and false rumor cascades



PROBABILISTIC DIFFUSION INFERENCE (PDI)

- A method to weigh the likelihood of potential parents being the true parent
- Stochastic approach: generate many versions
- We adopt two assumptions based on previous work to infer potential parents:
  - users with more followers (**followers** assumption)
  - users who are more recently active in the cascade (**recency** assumption)
  - Parameters alpha and gamma

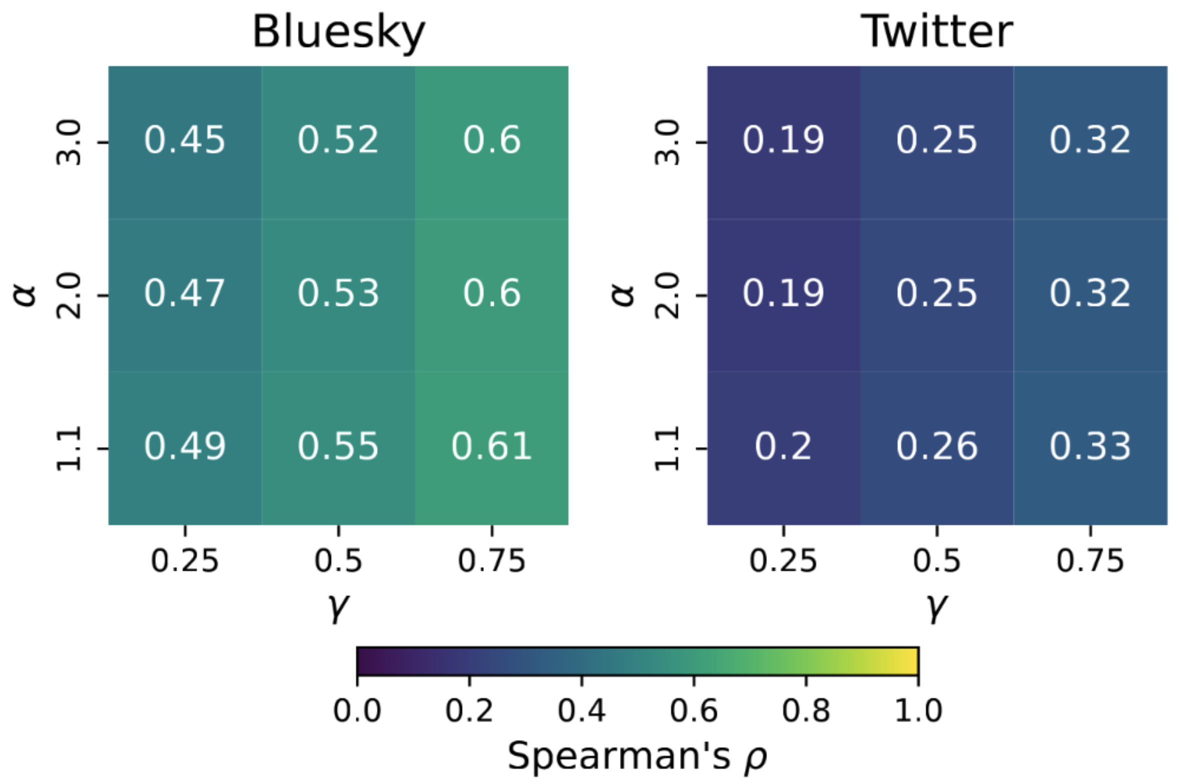


SOCIAL INFLUENCE ANALYSIS

- Pinpointing the most influential individuals within social networks
- Resharing networks
  - 2 case studies: Twitter, Bluesky
  - 10k random cascades
- **Naïve** vs **PDI-reconstructed** networks
- Node strength as proxy for influence
- If the two networks are very similar, the reconstruction process has minimal impact on these analyses

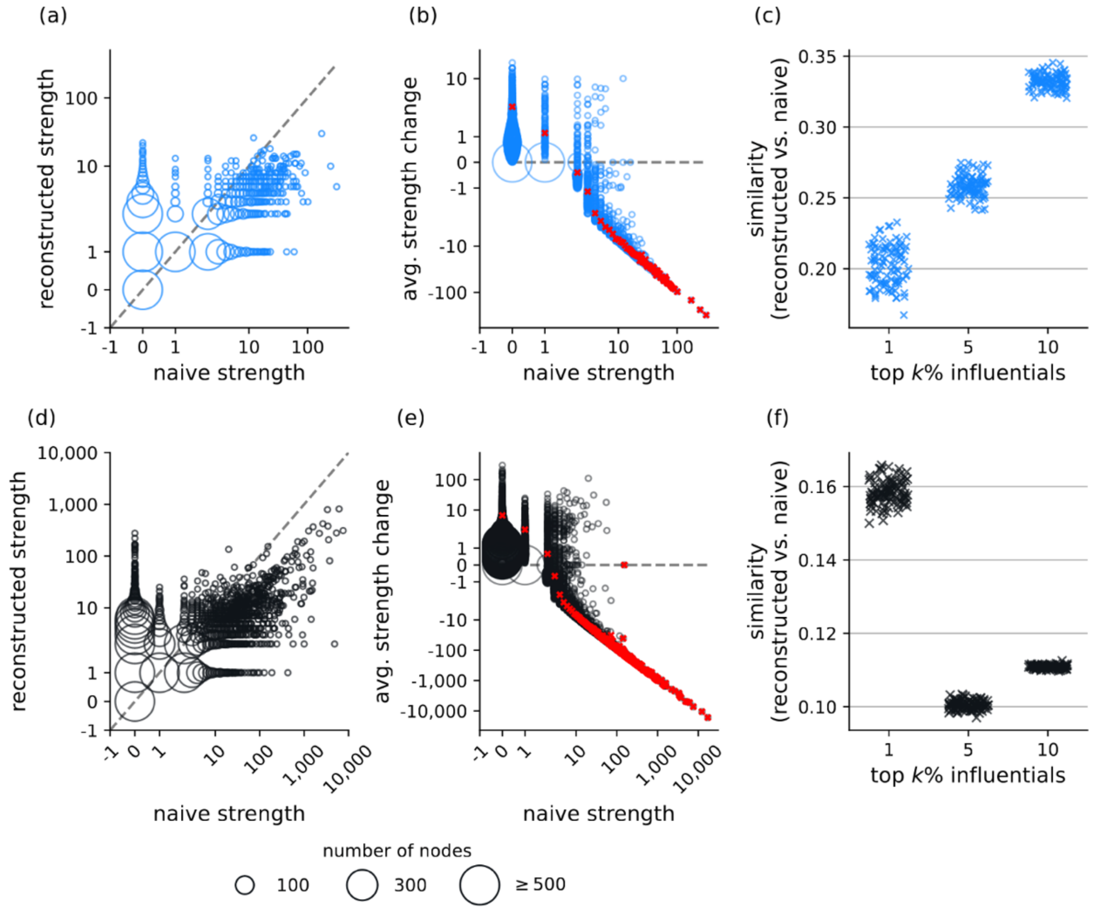
CORRELATION BETWEEN NAÏVE AND RECONSTRUCTED INFLUENCE

- Spearman's rank correlation between node strength in the naive and reconstructed networks
- Low values in both platforms (especially Twitter)
- Cascade reconstruction can considerably alter the perceived influence of nodes on both platforms



WHICH NODES ARE AFFECTED?

- Non-influential nodes in the naive network gain influence in the reconstruction
- Most influential nodes lose influence after reconstruction
- Considering the top K% (1, 5, 10) influentials yields very different results in the two networks
  - Careful when looking for "superspreaders"!



HOW DIFFERENT ARE RECONSTRUCTED CASCADES?

- Dataset of over 100k true and false rumor cascades
- Comparison with Time Inferred Diffusion method (by Vosoughi et al.)
  - You only reshare from accounts you follow
- Problematic assumption: 50% of the content you see comes from outside your network
- Topological analysis of reconstructed cascades: **PDI** vs **TID**

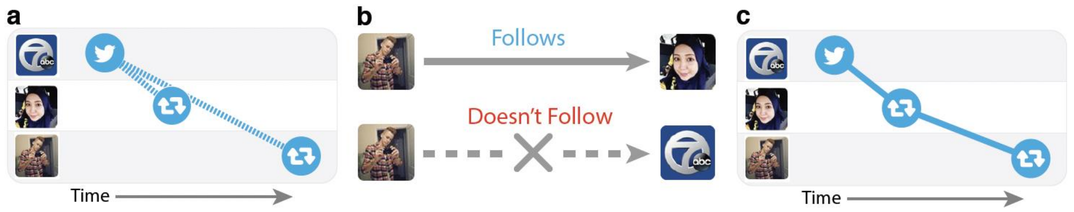
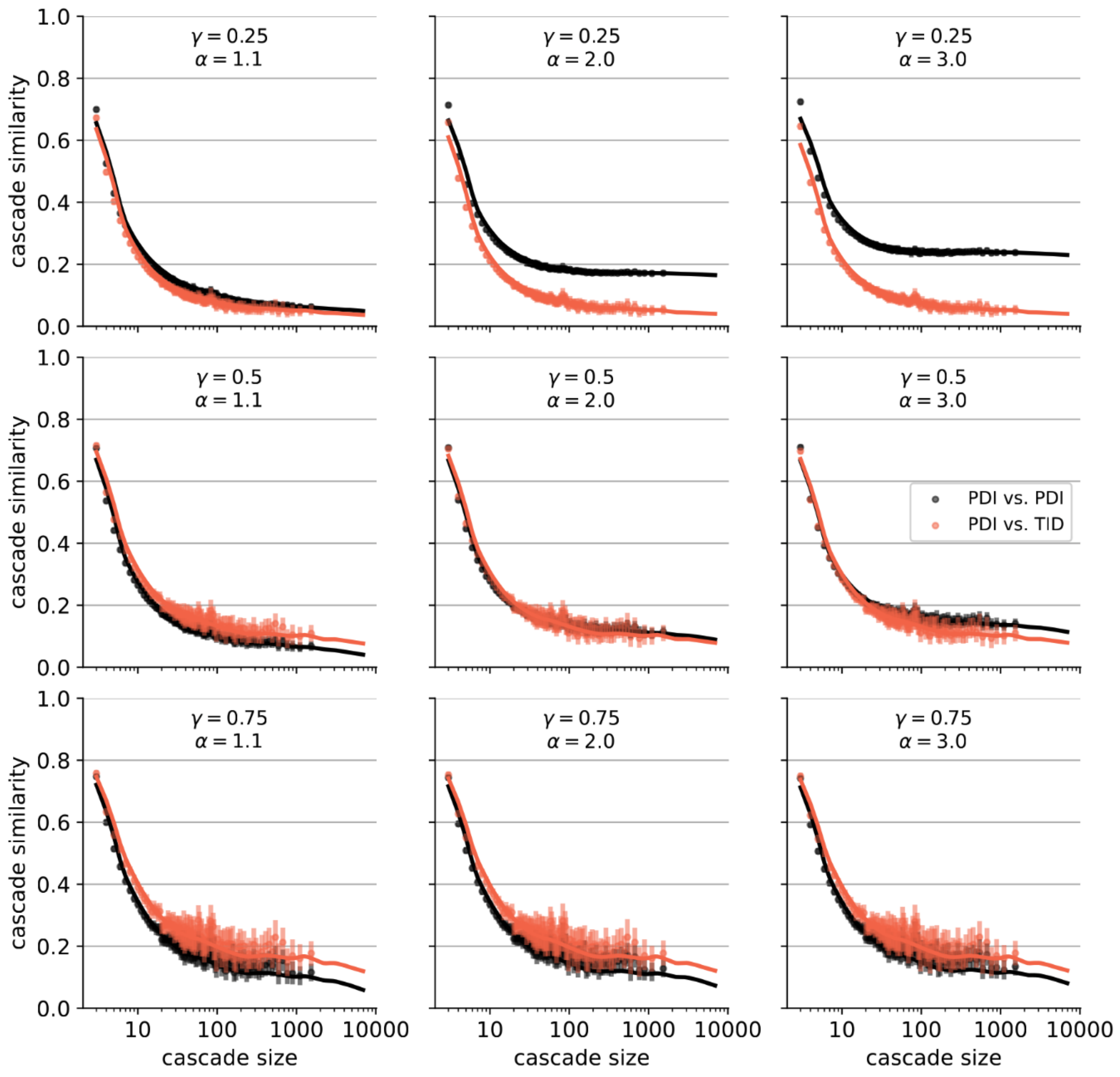


Figure S5: Using Twitter's follower graph to infer the correct retweet path of a tweet. Panel (a) shows the retweet path provided by the Twitter API. Panel (b) shows that the bottom user is a follower of the middle user but not that of the top user (the user who tweeted the original tweet). Panel (c) shows that using this information, and the fact that the bottom user retweeted after the middle user, we can infer that the bottom person retweeted the middle person and not the top person.

Source: Vosoughi et al. (2018) *Science*

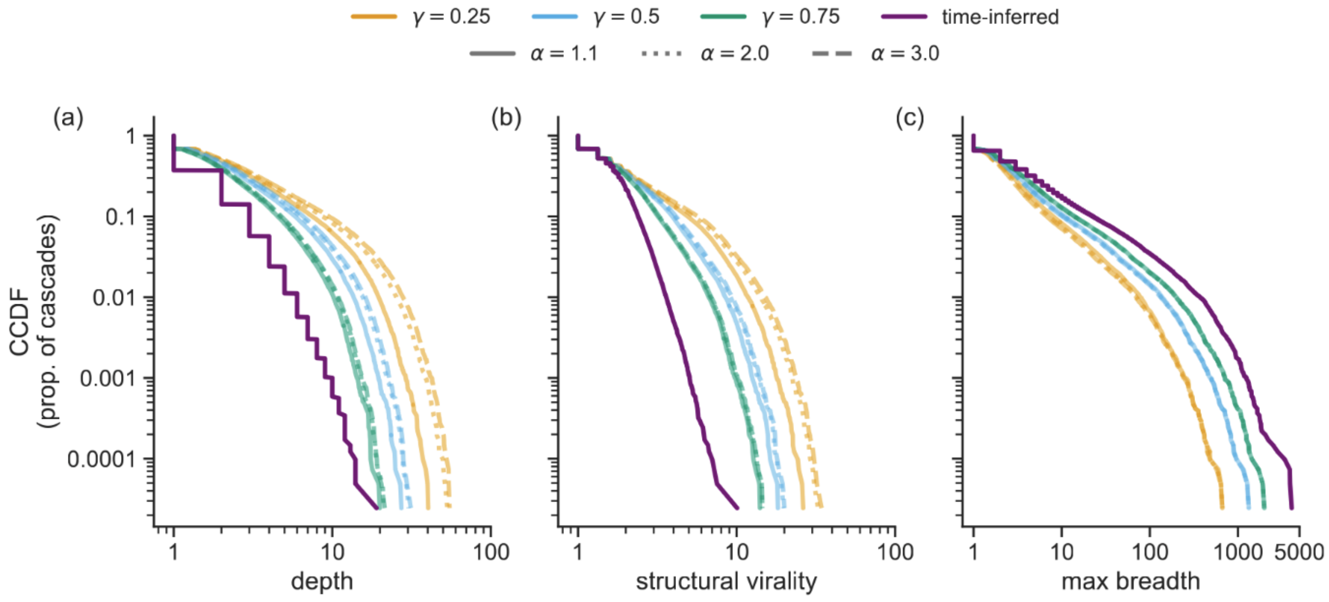
RECONSTRUCTED CASCADES ARE HIGHLY DISSIMILAR

- Cascade similarity = Jaccard index of edges
- 100 PDI reconstructions for each cascade
- PDI vs PDI = 100*99/2 = 4950 comparisons
- PDI vs TID = 100 comparisons
- Reconstructed cascades are highly dissimilar, especially for larger cascades



DIFFERENT DISTRIBUTIONS OF TOPOLOGICAL METRICS

- Analysis of topological metrics:
  - depth, structural virality and max breadth
- All distributions are statistically different according (2-sample Kolmogorov-Smirnov test)
- Example: more weight to the recency of a potential parent's post increases the depth and structural virality of cascades increase



VACCINE HESITANCY AND ONLINE MISINFORMATION

Vaccine hesitancy hinders vaccination campaigns

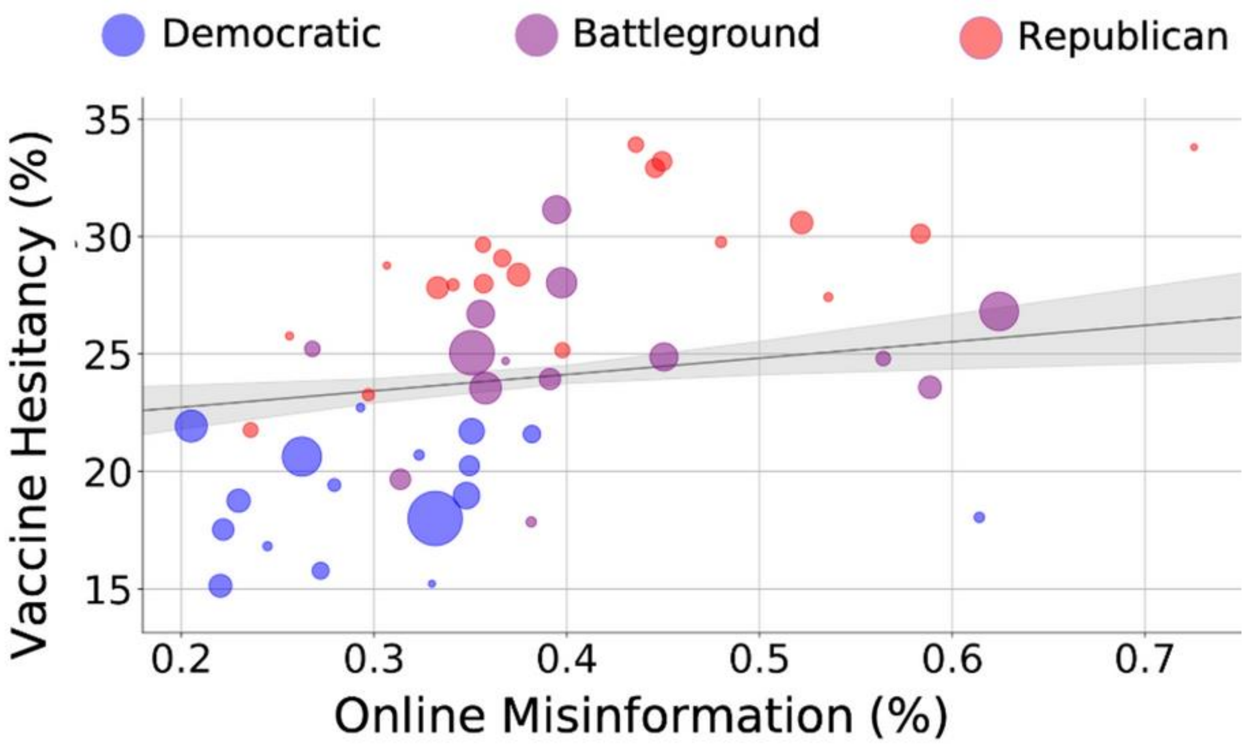
Online misinformation can drive vaccine hesitancy and refusal

Significant ~~associations~~ correlations between misinformation and hesitancy in the US

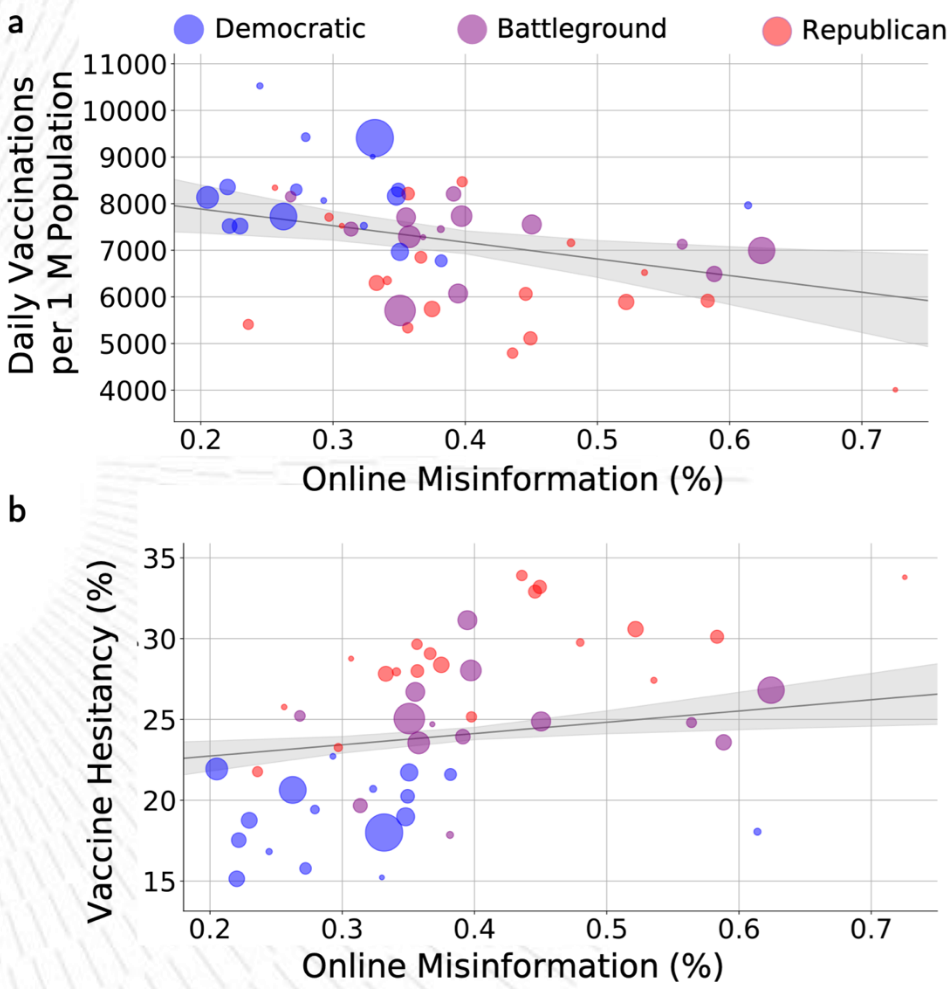
scientific reports

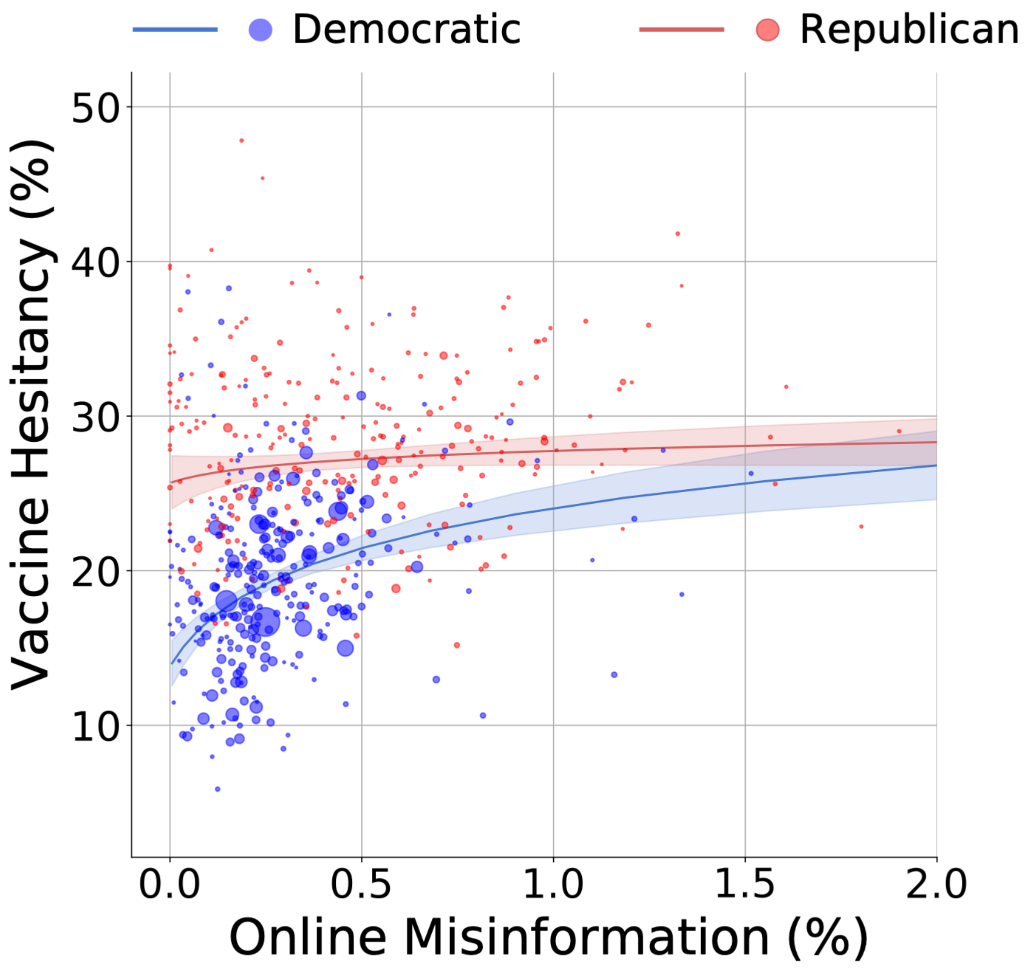
Online misinformation is linked to early COVID-19 vaccination hesitancy and refusal

Francesco Pierri, Brea L. Perry, Matthew R. DeVerna, Kai-Cheng Yang, Alessandro Flammini, Filippo Menczer & John Bryden

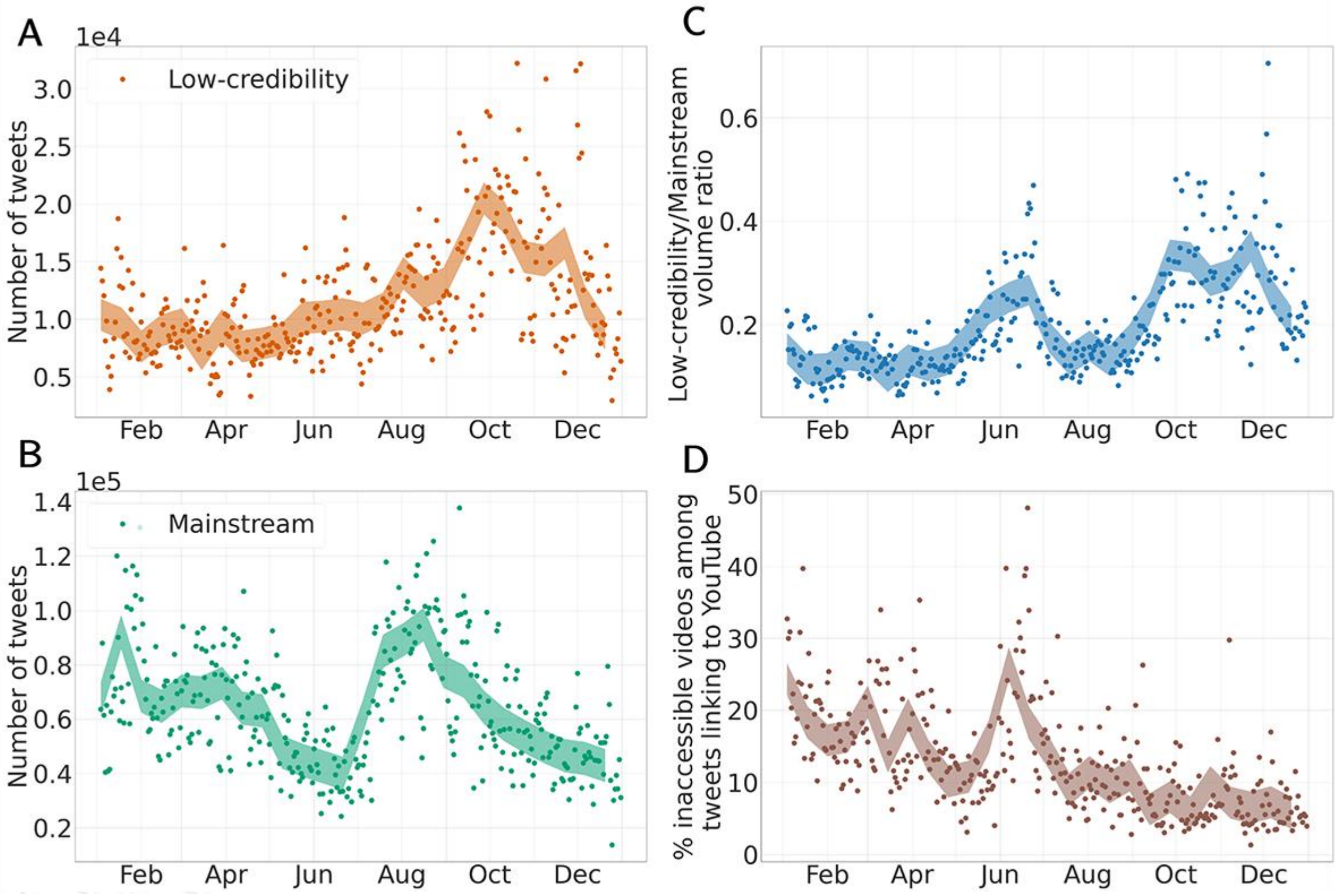


VACCINE HESITANCY AND ONLINE MISINFORMATION





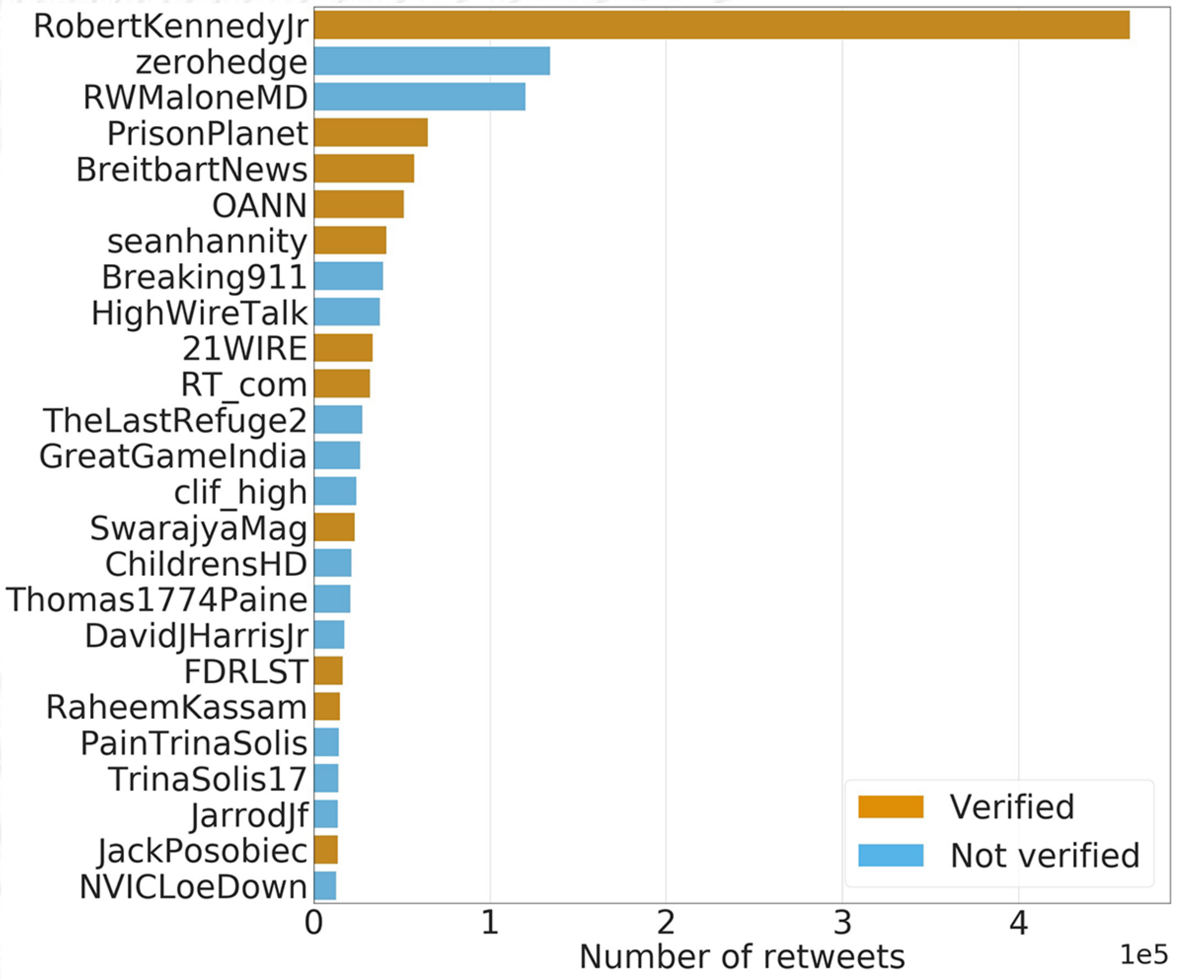
ONE YEAR OF VACCINE MISINFORMATION

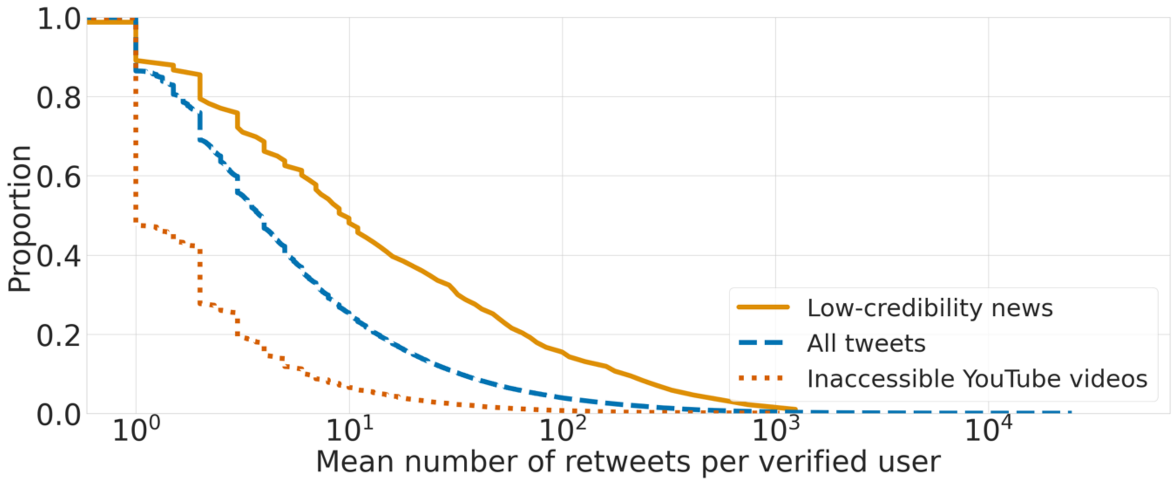


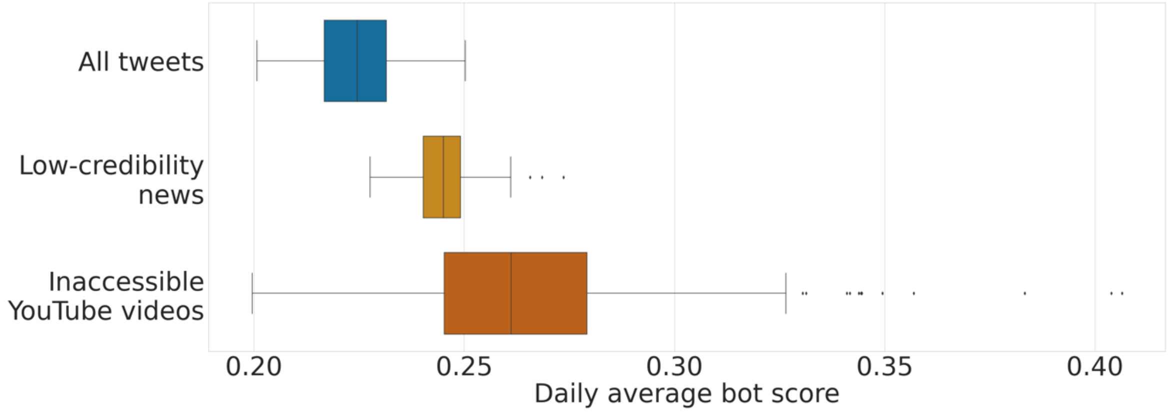
Francesco Pierri[1,2], PhD; Matthew R DeVerna[2], MA; Kai-Cheng Yang[2], MS; David Axelrod[2], MA; John Bryden[2], PhD; Filippo Menczer[2], PhD

JMIR Publications
Advancing Digital Health & Open Science
Journal of Medical Internet Research

ONE YEAR OF VACCINE MISINFORMATION







HOW CAN WE REDUCE MISINFORMATION?

Reduce financial incentives

Shadowbanning

Down-ranking

Soft labels

Deplatforming

Source: PlaygroundAI

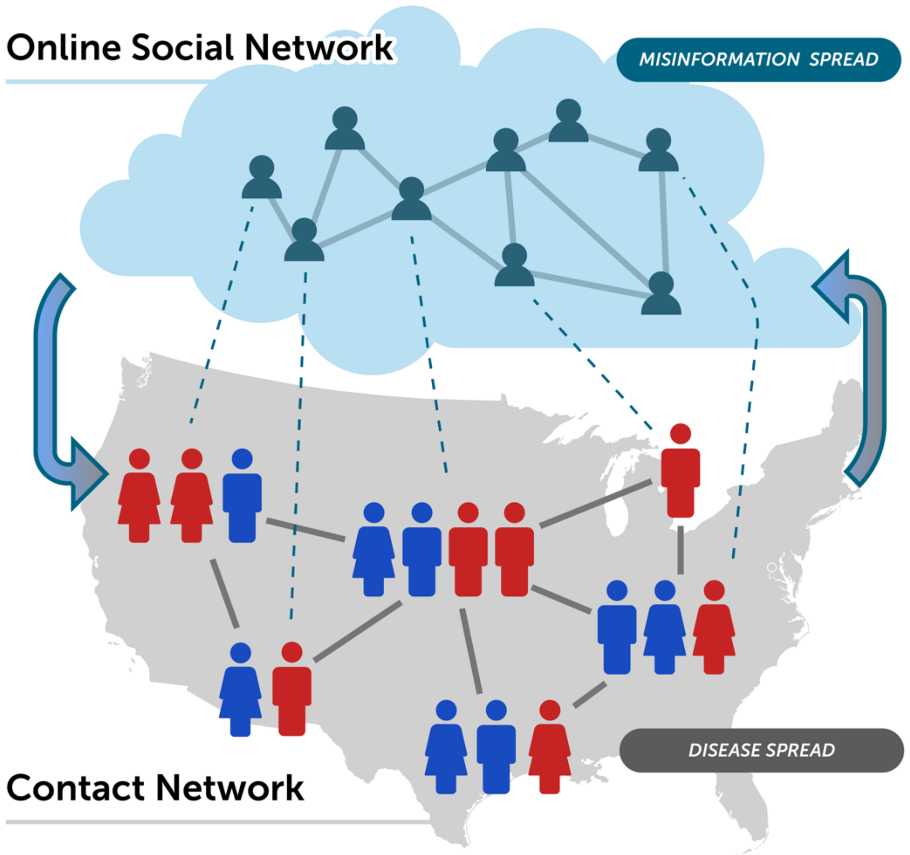
MODELING THE INTERPLAY BETWEEN MISINFO AND VIRUS

<u>How do misinformed populations influence the spread of disease?</u>

SMIR model (Susceptible-Misinformed-Infected-Recovered)

Mean field and ABM simulations

Twitter, mobile phone traces, voting record data



MEAN FIELD RESULTS

All individuals have equal chance of interacting

Infection prob. (β) for **M**isinformed is higher than **O**rdinary

Scaling factor $\lambda = \beta_M / \beta_O$

Greater λ, earlier and higher peak

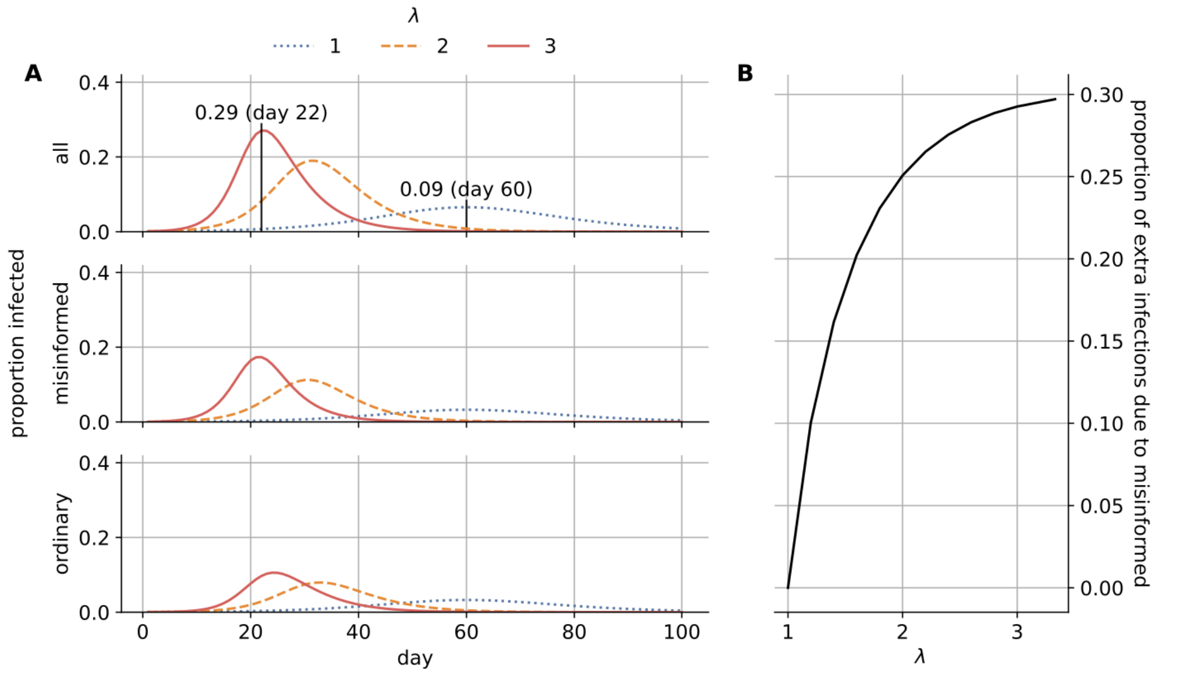


Figure 2: Increasing $\lambda = \beta_M/\beta_O$ accelerates and amplifies the infection. We use $\beta_O = 0.3$, $\gamma = 0.2$, and $\mu = 0.5$. (**A**) Proportion of the population infected on each day. (**B**) Extra proportion of the total population that is infected as a function of λ.

EFFECTS OF HOMOPHILY

α parameter to control interactions between **M**isinformed and **O**rdinary users (α = 1, full homophily)

Less infectious disease, higher harm for **M** nodes

Higher homophily, more shield to the **O** population

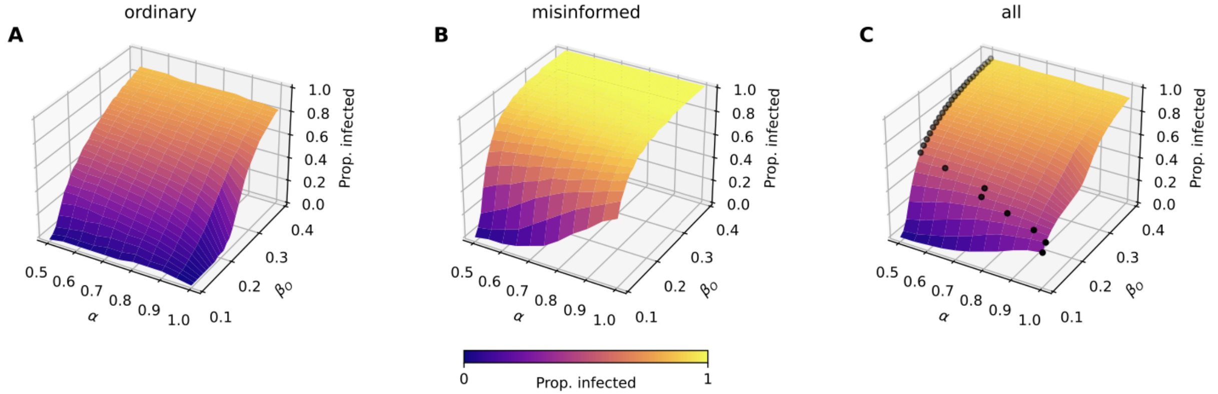


Figure 3: Homophily in the contact network worsens the infection among misinformed individuals, especially for lower transmission rates. The combined effects of transmission and homophily parameters, β_O and α, are examined with the mean-field approximation when $\lambda = 3$, $\gamma = 0.2$, and $\mu = 0.5$. We plot the proportions of infected individuals in (**A**) the ordinary population, (**B**) the misinformed population, and (**C**) the overall population. The maximum proportion of the overall population infected for each β_O is marked with a black dot. When the transmission rate is sufficiently high, homophily benefits the entire population but harms the misinformed group.

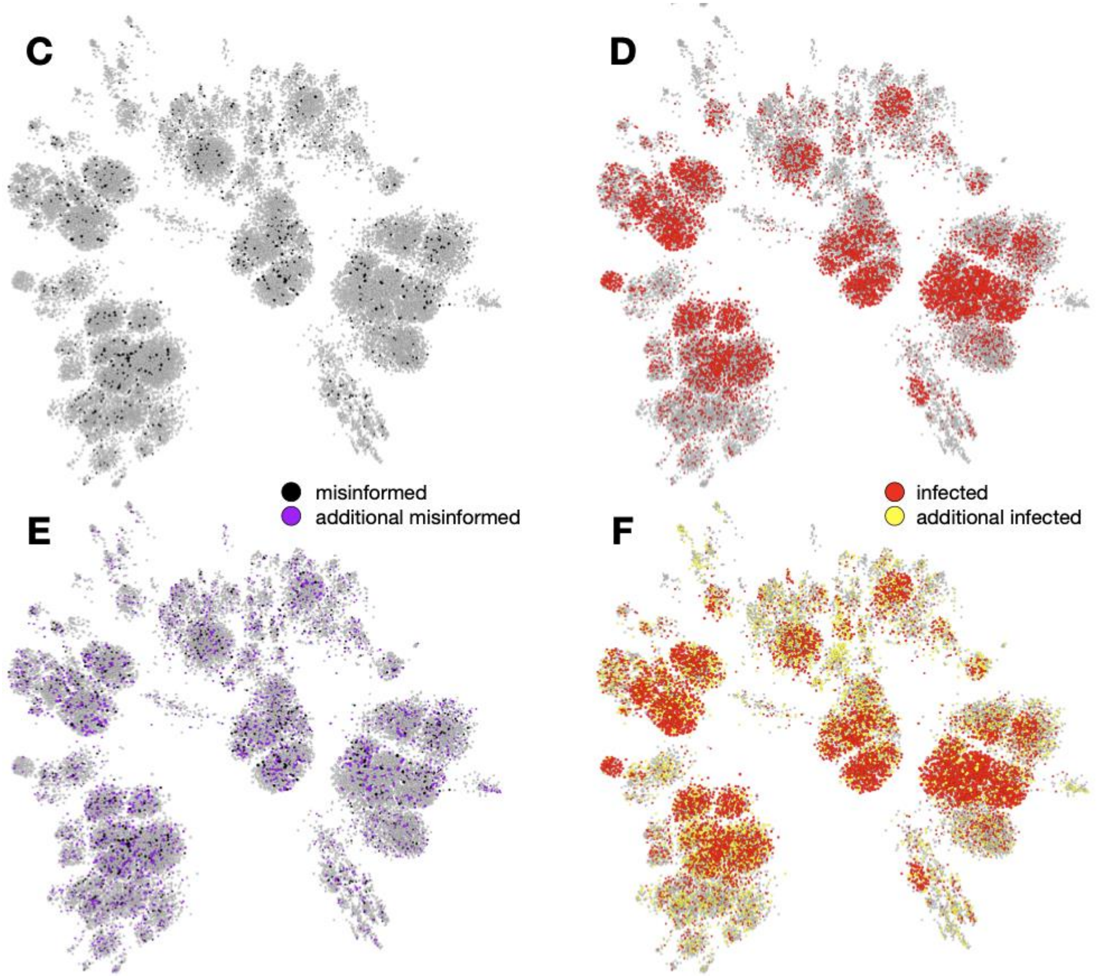
DATASETS AND NETWORK CONSTRUCTION

Covaxxy (2021): 25M COVID-19 vaccine tweets shared by over 2M geolocated US accounts

Political leaning inferred from web domains shared

Newsguard ratings

Safegraph mobility data (2019)



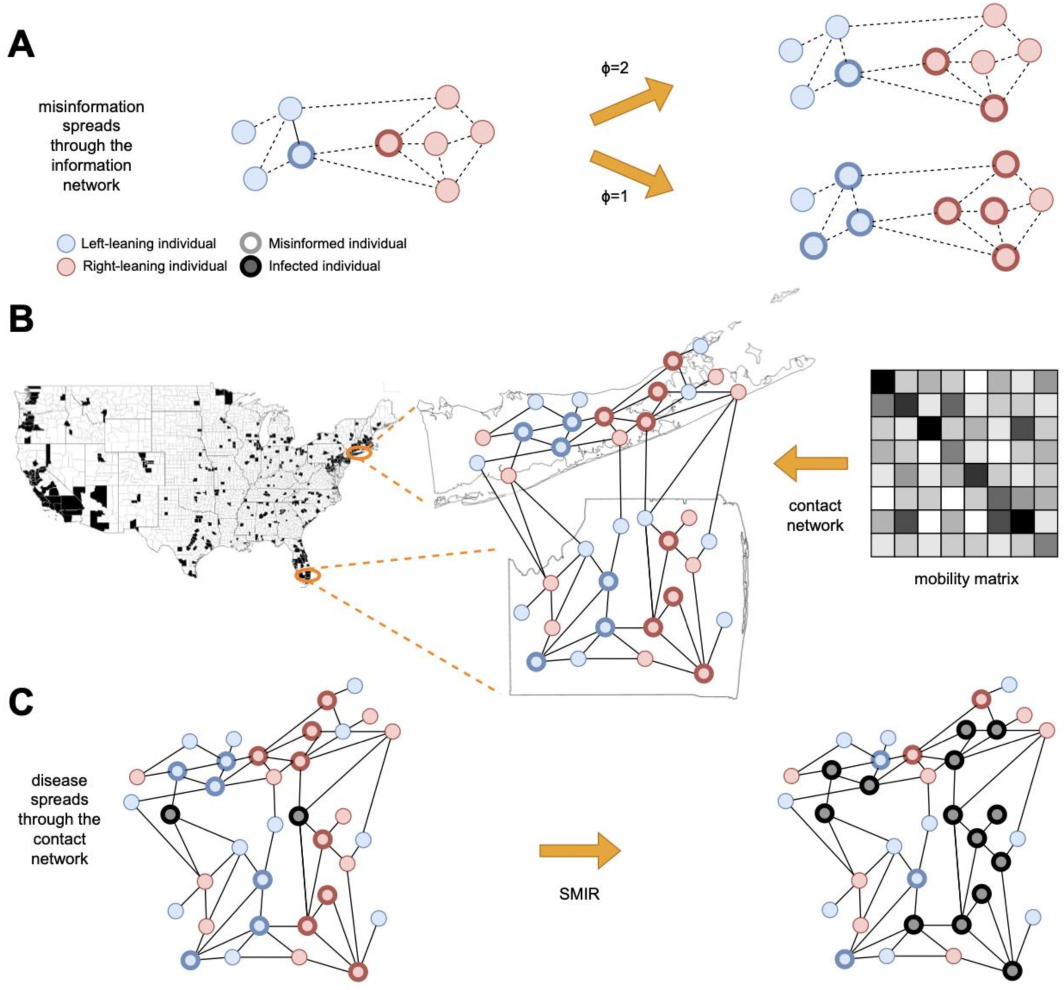
AGENT-BASED MODEL

Information network → Twitter

Contact network → Safegraph

Identify misinformed Twitter users

Sample geolocated users in each county based on Dem/Rep proportions



IMPACT OF MISINFORMATION

Propagate "misinformation" to Twitter neighbours with different thresholds (ϕ)

Extreme infectivity values:

$$p_M = 1, p_O = 0.01$$

Simulate disease dynamics

Worst case: +14% population infected

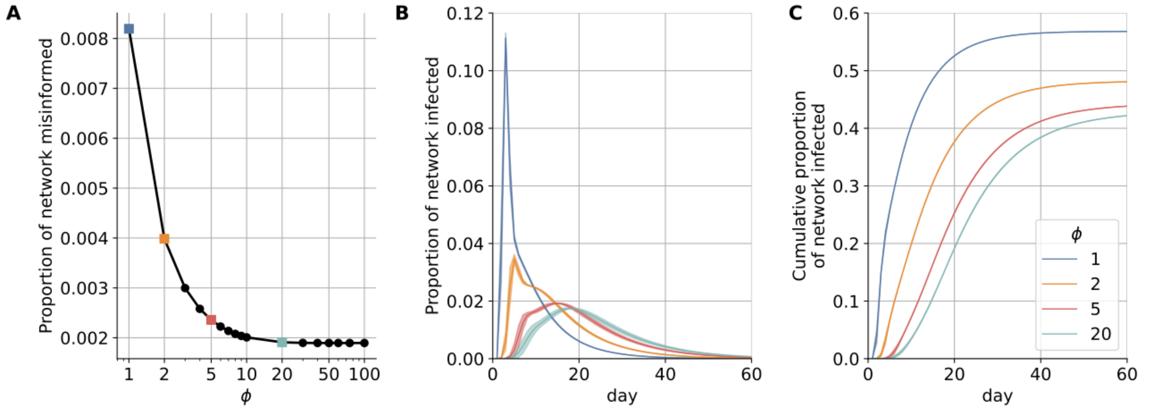


Figure 5: More misinformed individuals lead to a larger portion of the network becoming infected. Decreasing the resilience ϕ (**A**) increases the size of the misinformed subpopulation, leading to (**B**) faster infection spreading and (**C**) a greater cumulative number of infections. In panels (**B**, **C**), lines and corresponding shaded regions represent the mean and standard deviation across simulations, respectively.

LIMITATIONS

Beliefs and behaviours vary over time

Twitter is NOT representative of the general population

Simplifying dichotomy between Misinformed and Ordinary individuals

How to properly link misinformation to prob. infection?

Source: PlaygroundAI

ONGOING WORK

Urban epidemic model based on demographics

Electoral precinct data and political affiliation

Socio-economic features

How to link misinformation to infectiousness?

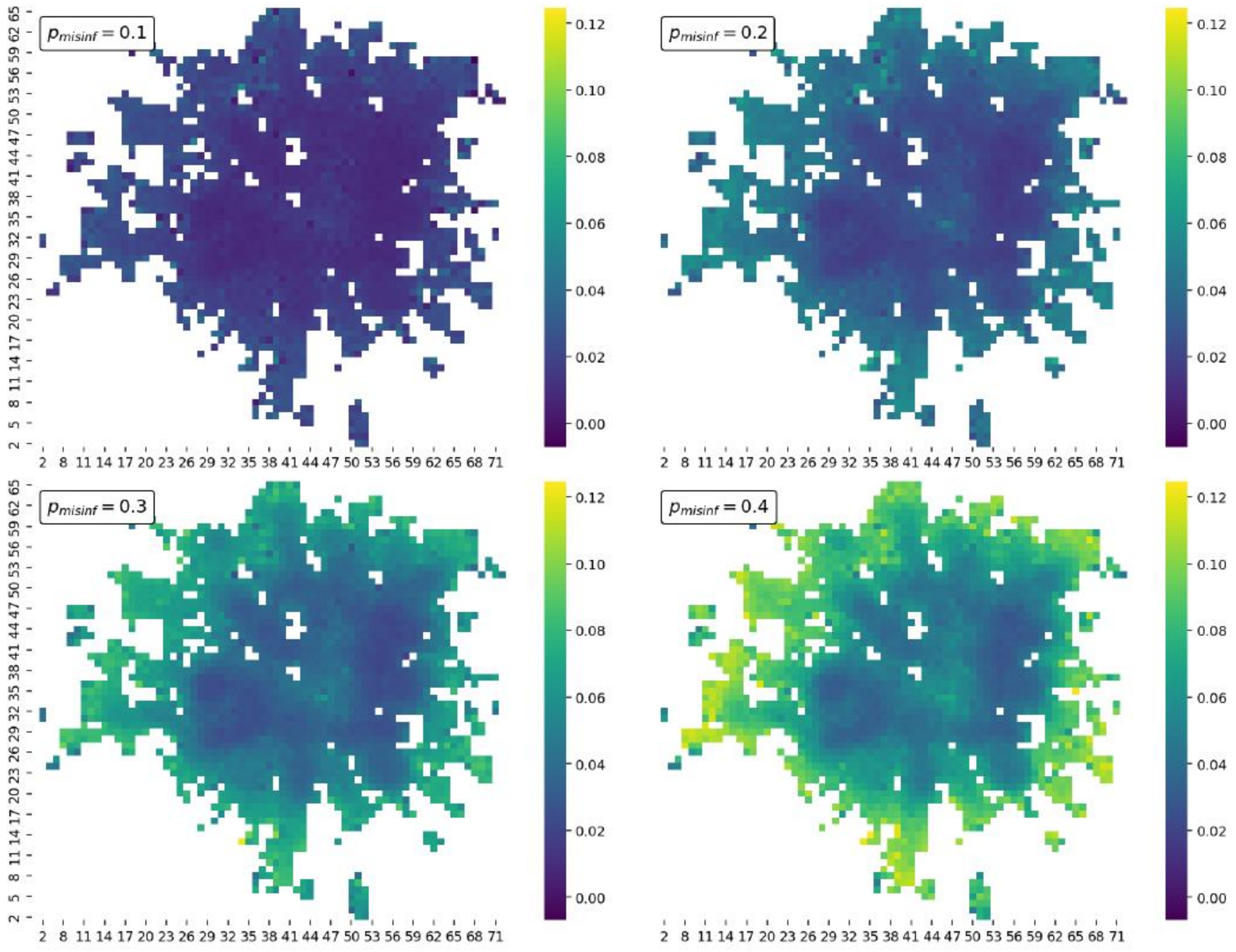


Figure 2: Increase in AR_i in Milano (250m)

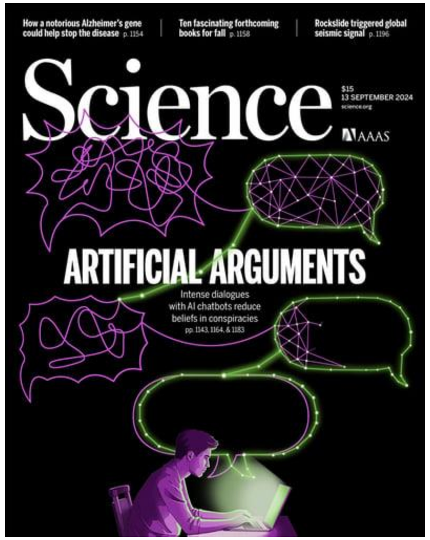
AI VERSUS CONSPIRACY THEORIES

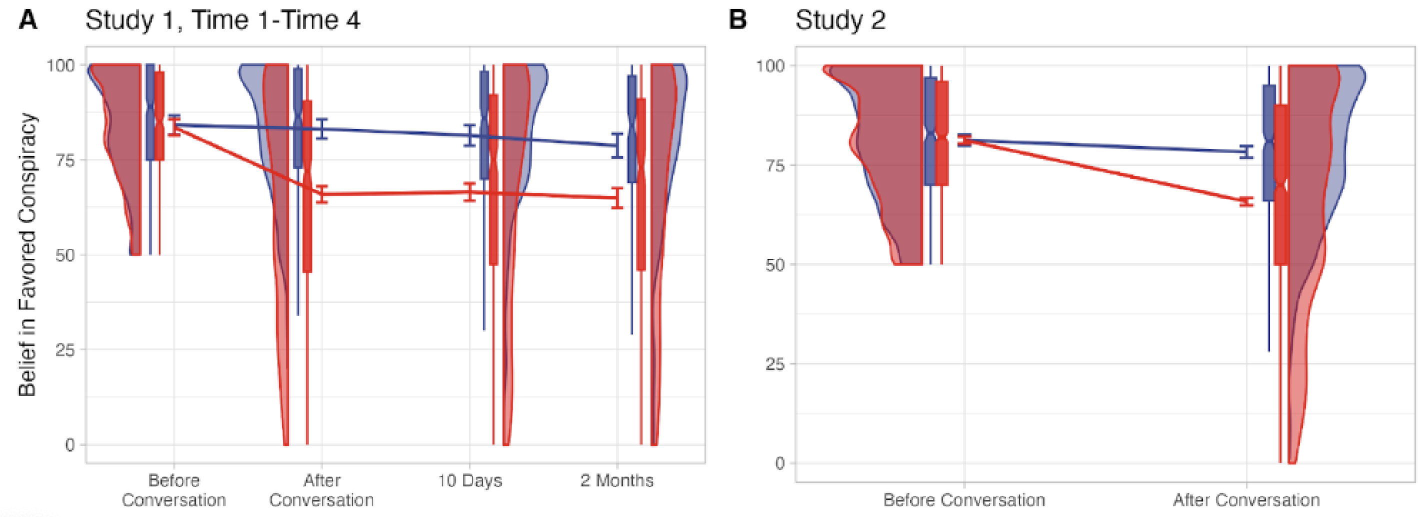
N=2190 CT believers

3-round dialogues with GPT4

Banal topic versus reduce conspiracy beliefs

20% reduction, consistent over time and spillovers





Source: Costello et al. Science (2024)

THESIS PROPOSALS

- Leveraging Large Language Models to counter online misinformation and hate speech
- Simulating online social media platforms with LLMs
- Studying human behaviour in online multiplayer games (e.g. League of Legends, DOTA)
- Investigating conspiracy theories on TikTok and YouTube

THANK YOU FOR YOUR ATTENTION!

THOUGHTS?