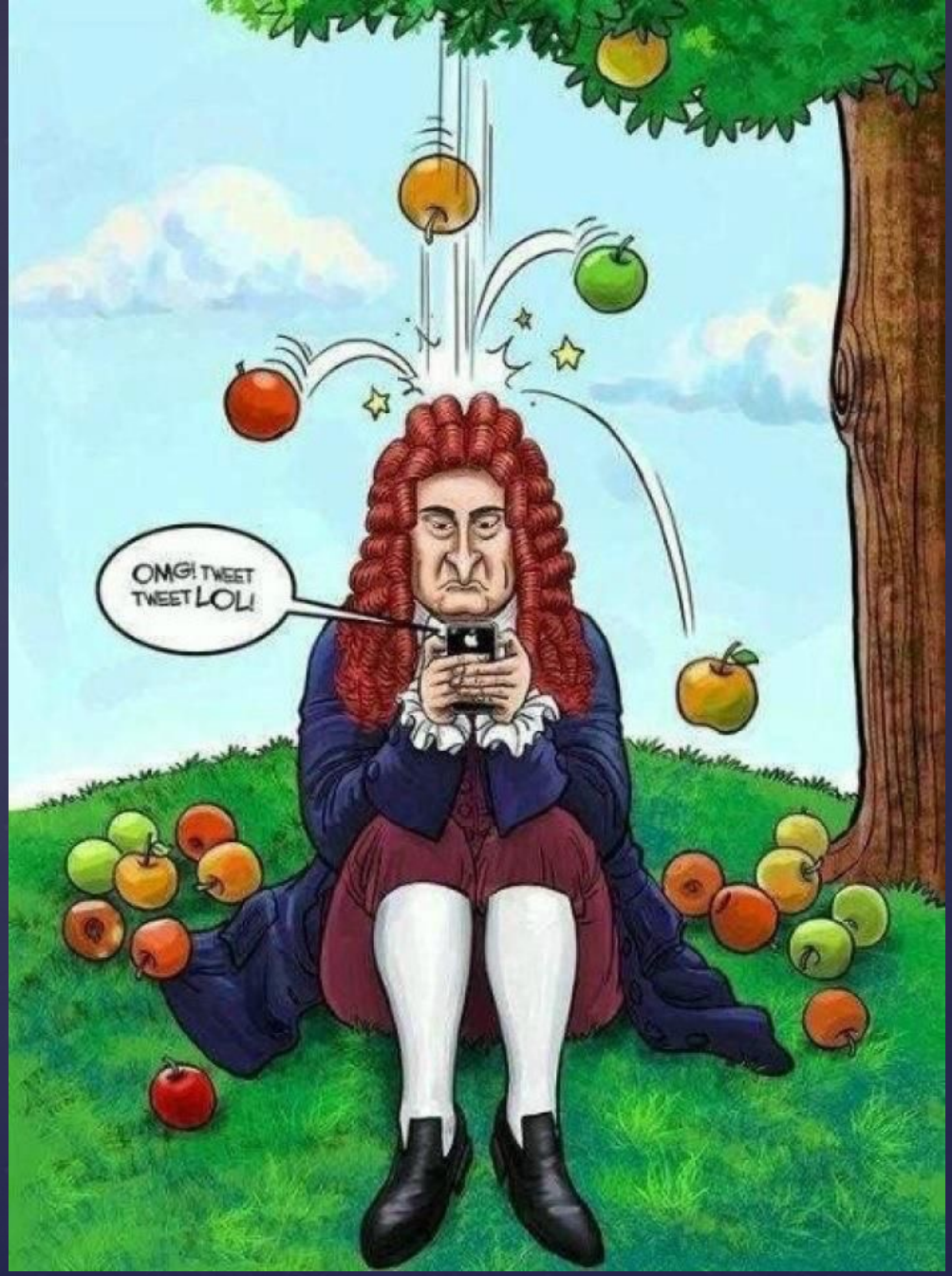
Can crowdsourcing rescue the social marketplace of ideas?

Taha Yasseri
School of Sociology and Geary Institute for Public Policy
University College Dublin

Buenos Aires, 8 Feb 2023

@TahaYasseri



Sanhorn Chen (Pinterest)

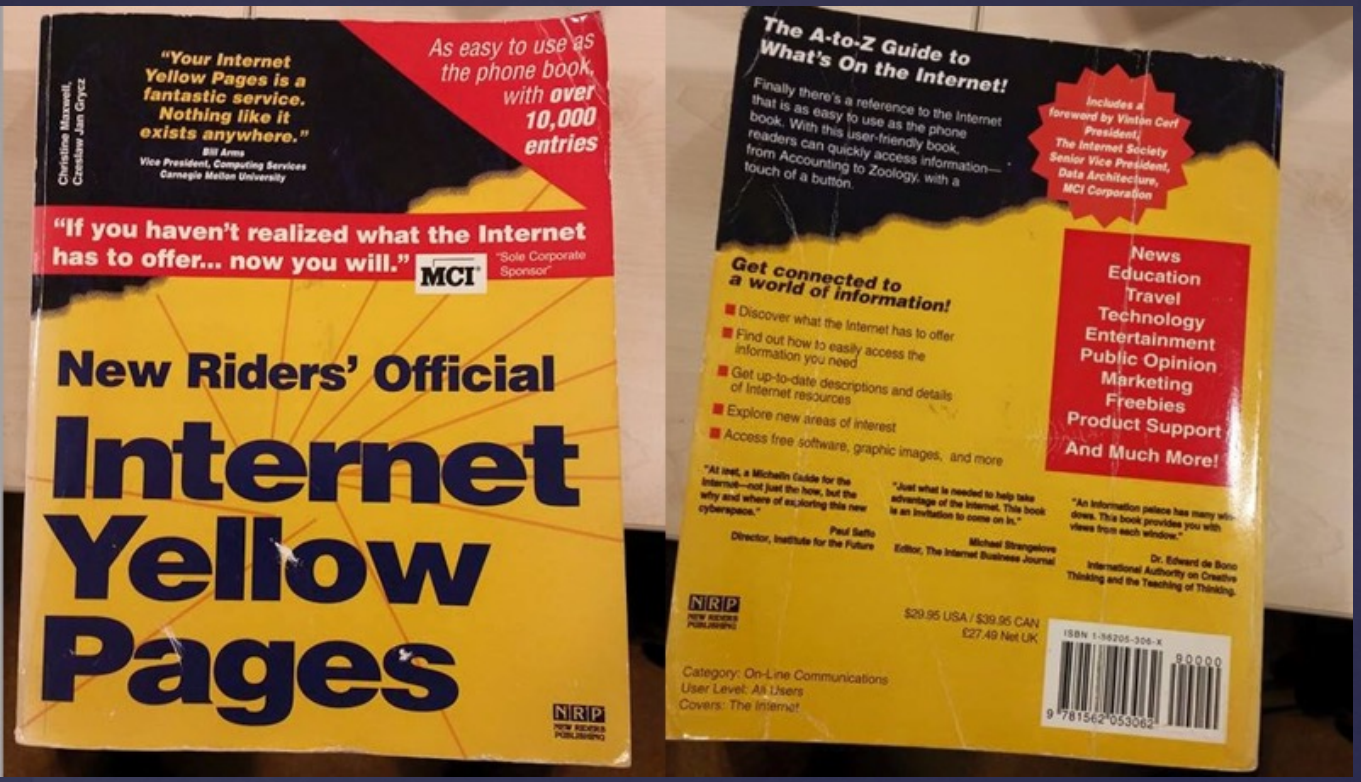
@TahaYasseri

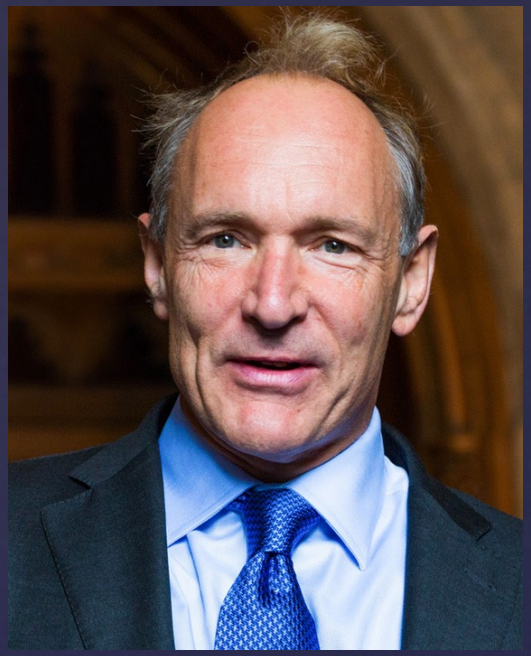
NGC 4414 (Source: The Hubble Heritage Team)

Brücke Osteuropa

CERN (1954 –)

@TahaYasseri





Tim Berners-Lee

WWW (1989)

Dot-com bubble (1990's)

Web2.0 (1999)

2001–

Zeroth Law of Wikipedia:

In theory it should never work, it only works in practice!

@TahaYasseri

Arab Spring (2010-2012)

2004

2006

Social Media

2015

2016

2016

The Guardian

2021

@TahaYasseri

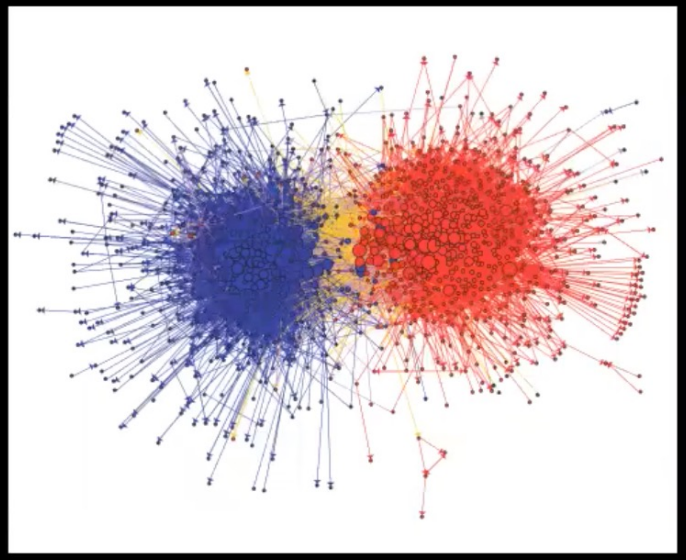
Now

- Polarization
- Hate-speech
- Misinformation

Social Media

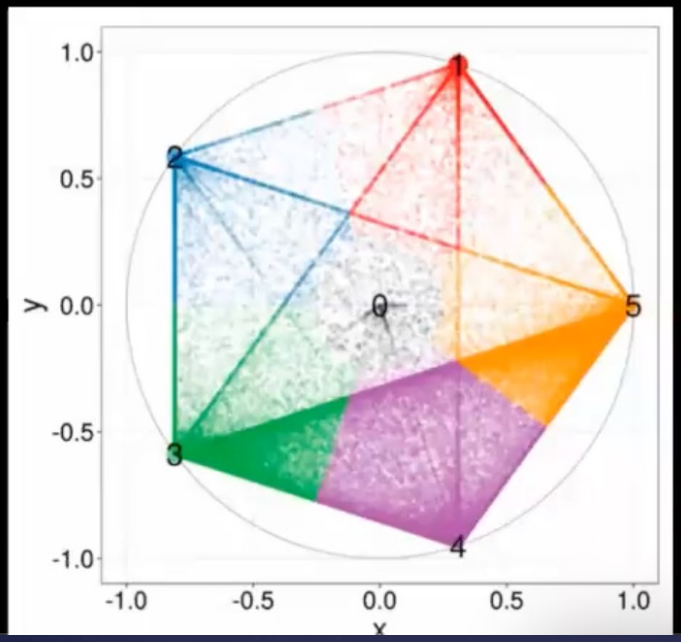
How did a new technology make people dumber?

@TahaYasseri



Blogs

Adamic & Glance, 2005

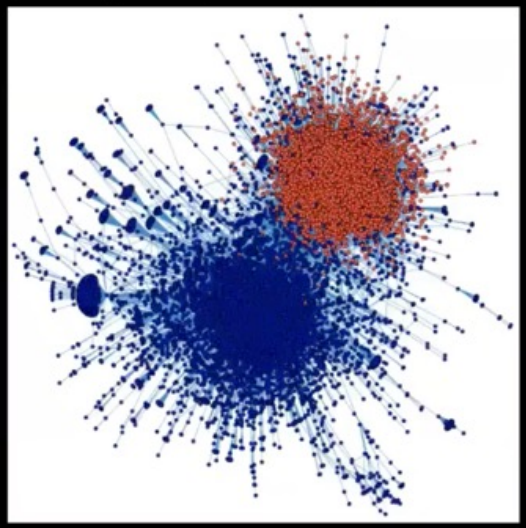


Facebook

Schmidt, 2017

Twitter

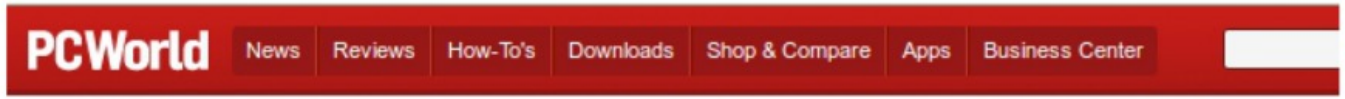
Conover, Gonçalves, Flammini, Menczer, 2012



Polarization

Why Does Wikipedia Work and Social Media Don't?

The Paradox of Openness in the Digital Age



Wikipedia Wars: 10 Biggest Edit Battles

We round up the most heated, most bitterly contested, and most pointless confrontations over facts in Wikipedia's 10-year history. What makes people leap into these frays will amaze and amuse you.

By David Daw, PCWorld Jul 26, 2011 3:00 am

SF WEEKLY

News

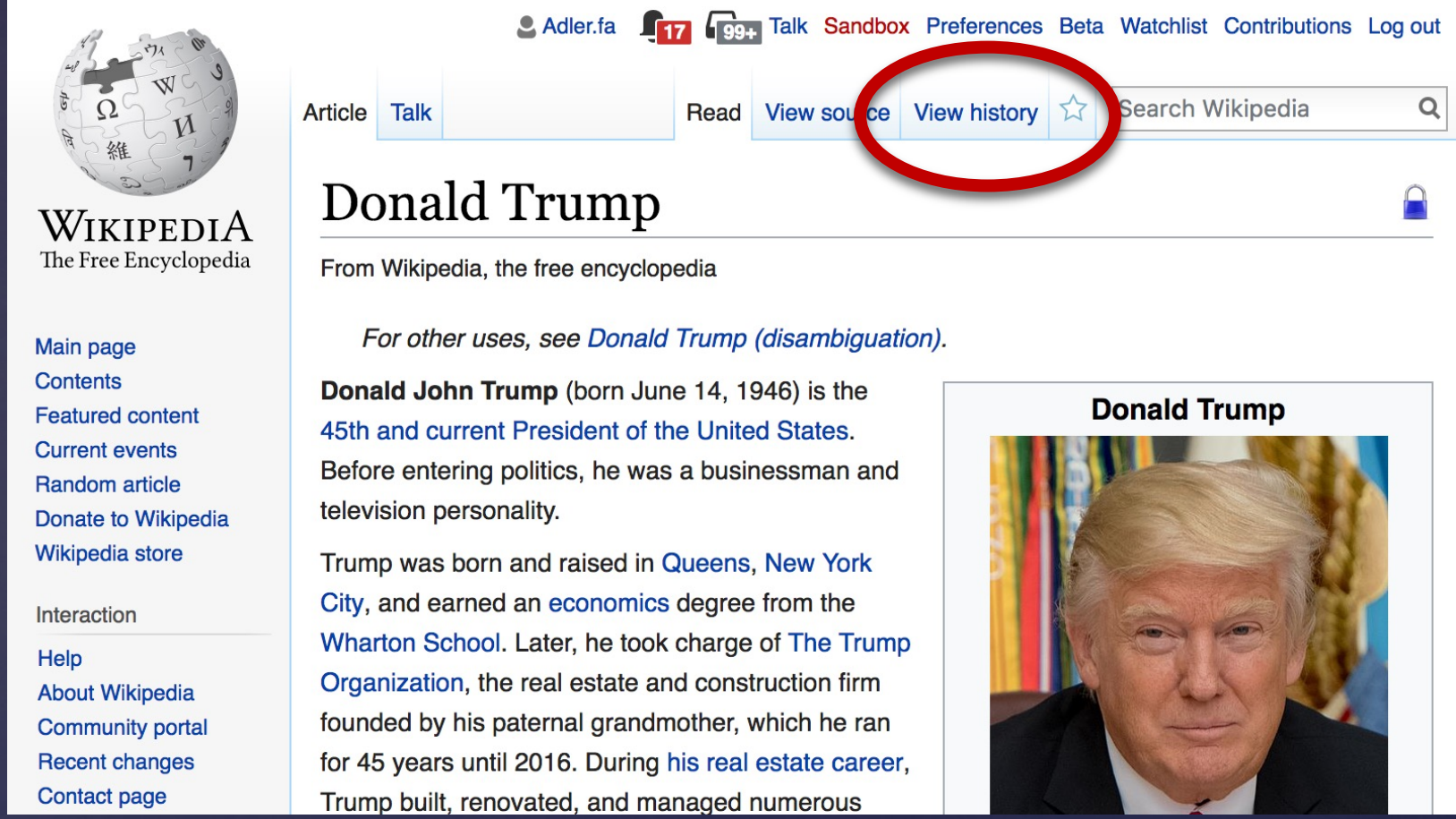
Wikipedia Idiots: The Edit Wars of San Francisco

The Telegraph

Michael Jackson's death sparks Wikipedia editing war

Wikipedia Edit Wars

@TahaYasseri



02:03, 21 June 2017 General Ization (talk | contribs) m . . (302,073 bytes) (-260) . . (*Reverted edits by PerfectlyIrrational (talk) to last version by Moxy*) (thank)

01:53, 21 June 2017 PerfectlyIrrational (talk | contribs) . . (302,333 bytes) (+260) . . (*More pictures*) (thank)

23:56, 20 June 2017 Moxy (talk | contribs) m . . (302,073 bytes) (-184) . . (*Reverted edits by PerfectlyIrrational (talk) to last version by Power~enwiki*) (thank)

23:51, 20 June 2017 PerfectlyIrrational (talk | contribs) . . (302,257 bytes) (+184) . . (*Undid revision 786684669 by PerfectlyIrrational (talk)*) (thank)

23:51, 20 June 2017 PerfectlyIrrational (talk | contribs) . . (302,073 bytes) (-184) . . (thank)

23:51, 20 June 2017 PerfectlyIrrational (talk | contribs) . . (302,257 bytes) (+184) . . (thank)

Wikipedia Reverts

@TahaYasseri

Network of reverts

- January 2001 – October 2011
- 13 languages
- 4.7 million reverts

Data

Value production in a collaborative environment
T Yasseri & J Kertész
Journal of Statistical Physics **151** (3-4), 414-439 (2013)

M top-10 lists

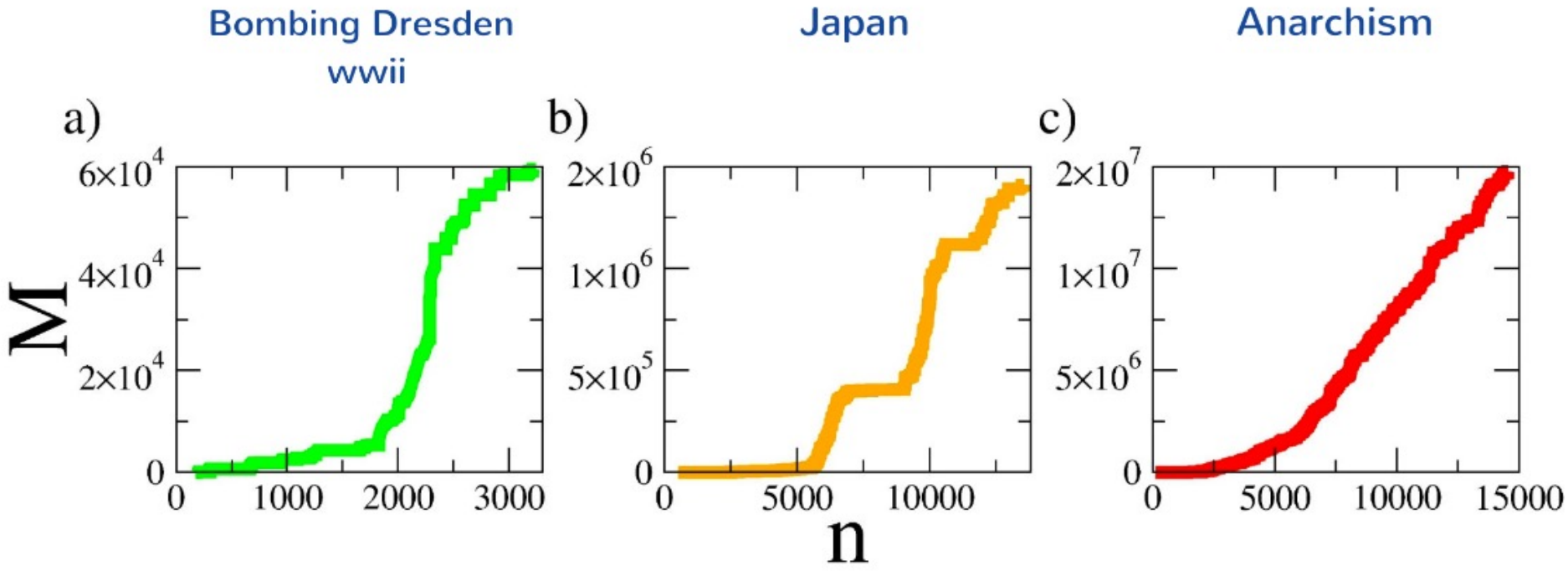
English	German	French	Spanish	Czech
George W. Bush	Croatia	Ségolène Royal	Chile	Homosexuality
Anarchism	Scientology	Unidentified flying object	Club América	Psychotronics
Muhammad	9/11 conspiracy theories	Jehovah's Witnesses	Opus Dei	Telepathy
LWWEe	*Fraternities*	Jesus	Athletic Bilbao	Communism
Global warming	Homeopathy	Sigmund Freud	Andrés Manuel López Obrador	Homophobia
Circumcision	Adolf Hitler	September 11 attacks	Newell's Old Boys	Jesus
United States	Jesus	Muhammad al-Durrah incident	FC Barcelona	Moravia
Jesus	Hugo Chávez	Islamophobia	Homeopathy	Sexual orientation change efforts
Race and intelligence	Minimum wage	God in Christianity	Augusto Pinochet	Ross Hedvíček
Christianity	Rudolf Steiner	Nuclear power debate	Alianza Lima	Israel

M: Measure of Controversy of the Article

T. Yasseri, A. Spoerri, M. Graham, & J. Kertész.
"The most controversial topics in Wikipedia: A multilingual and geographical analysis." (2013)

@TahaYasseri

M: Controversy measure



Dynamics of Conflicts in Wikipedia
Yasseri T, Sumi R, Rung A, Kornai A, Kertész J (2012), PLoS ONE 7(6): e38869

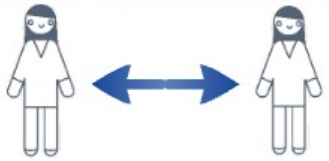
Deffuant Model

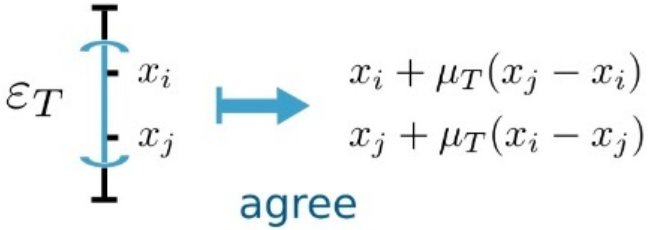
Mixing beliefs among interacting agents
Guillaume Deffuant et al, Advs. Complex Syst. 03, 87 (2000)

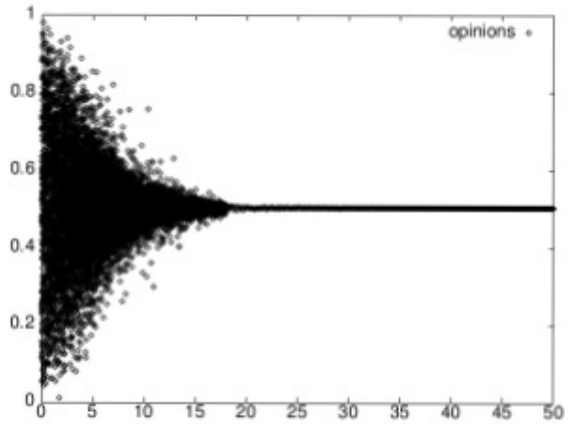
N editors

opinion $x_i \in [0, 1]$

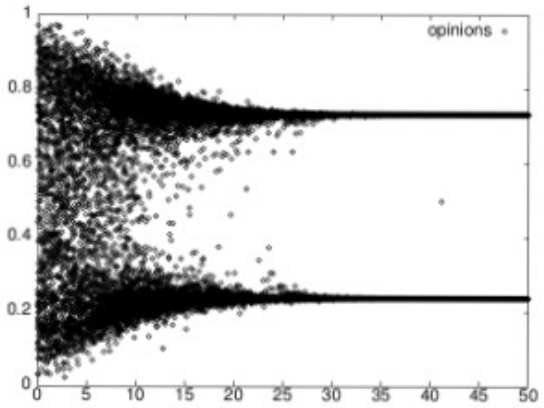
Bounded Confidence Model







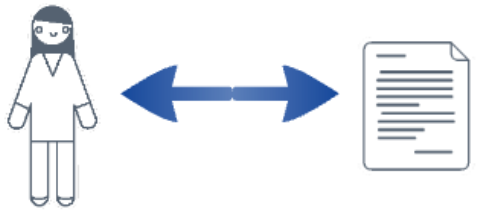
Large Tolerance



Small Tolerance

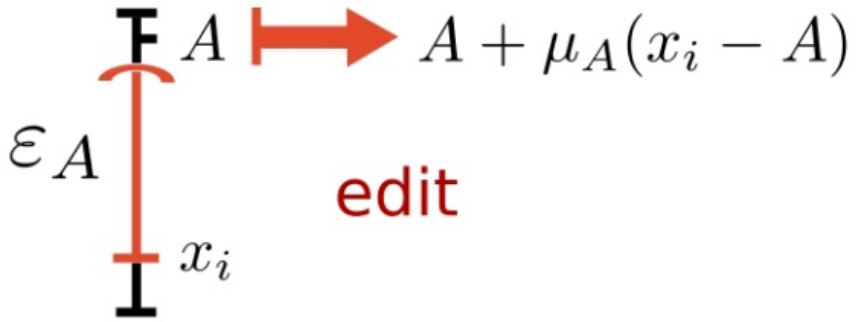
Additional Ingredient

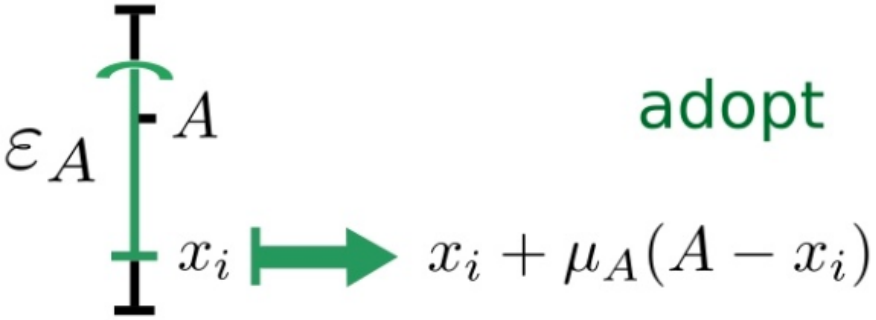
Editor-article interaction



1 article

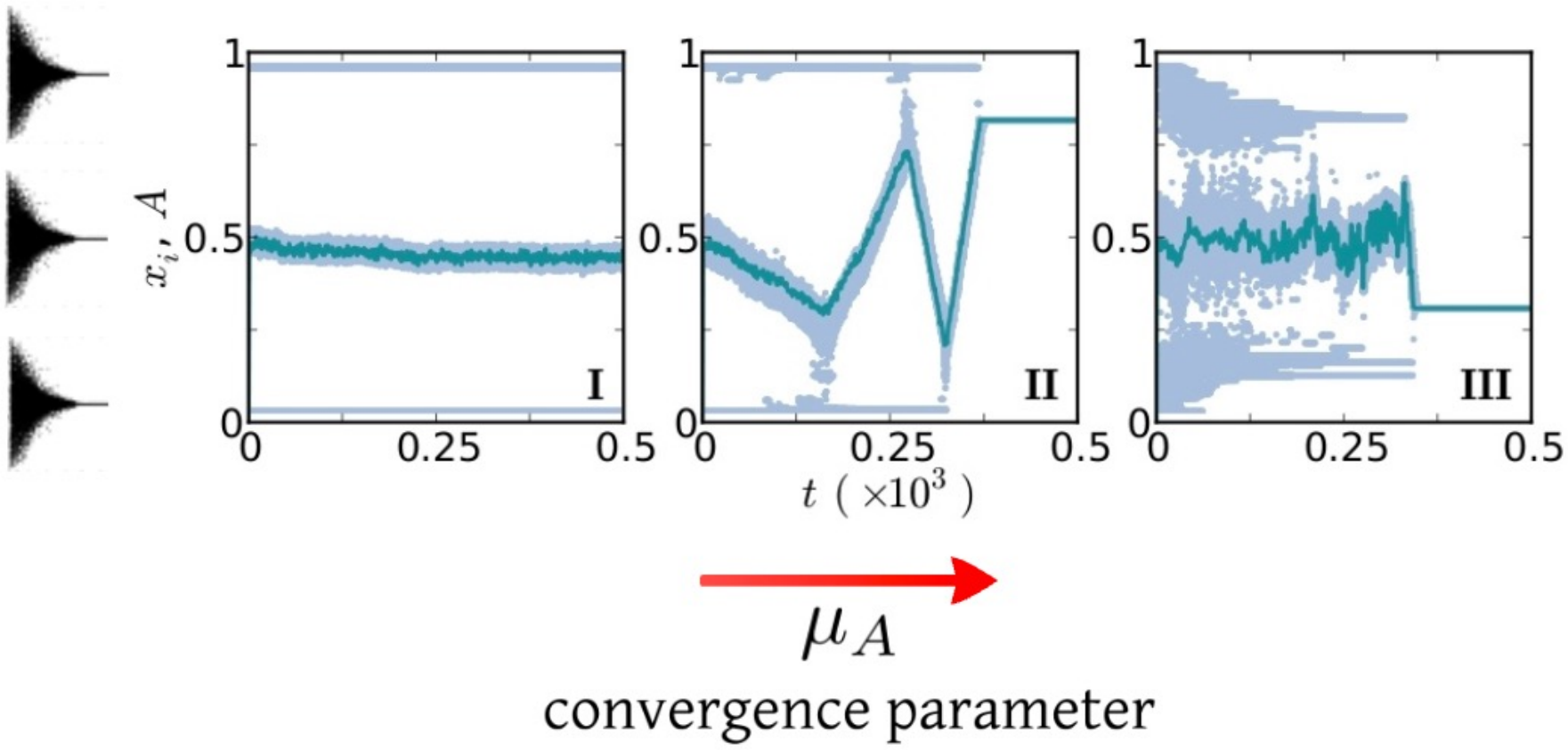
state $A \in [0, 1]$



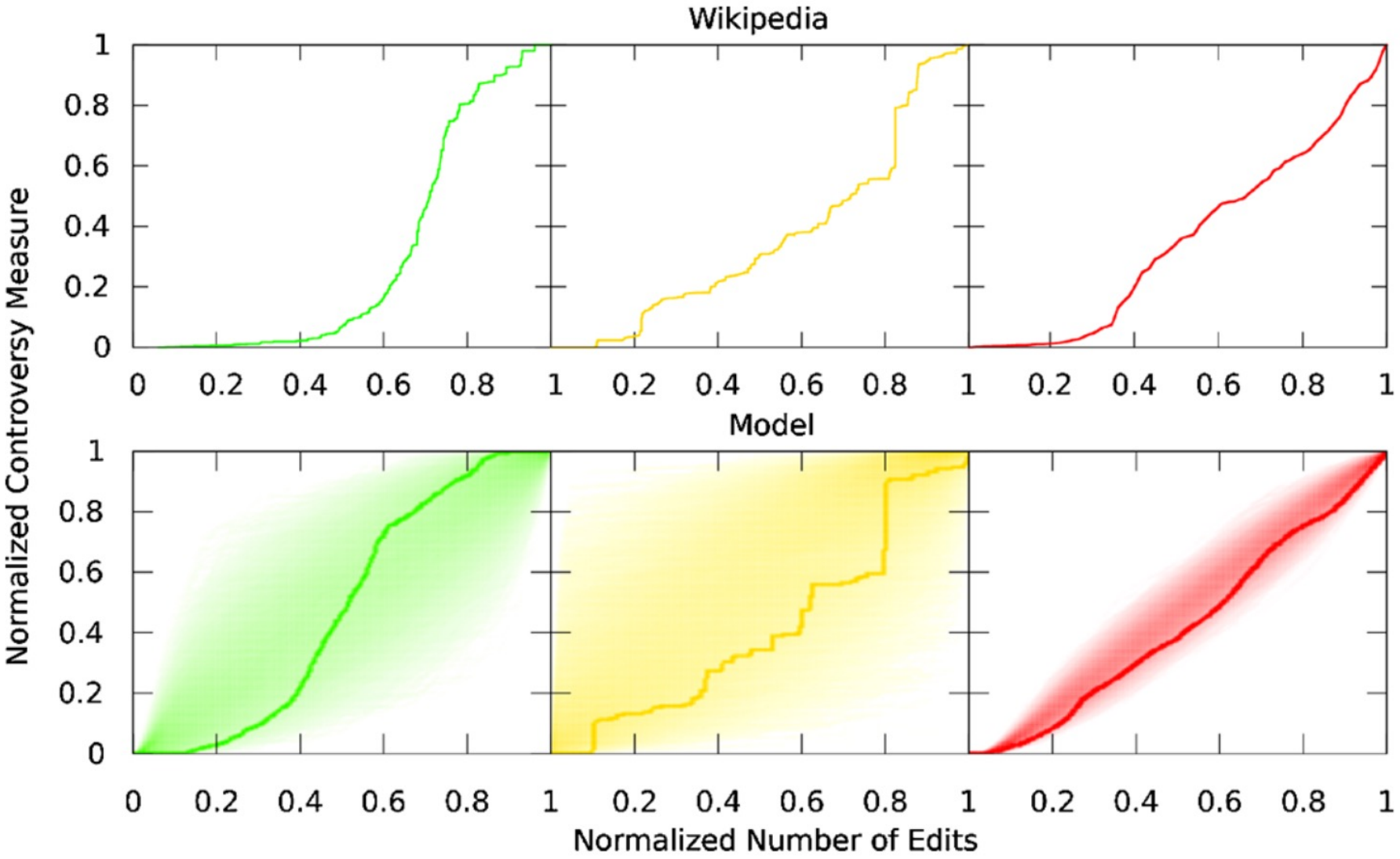


Opinions, Conflicts, and Consensus: Modeling Social Dynamics in a Collaborative Environment

Török J., Iñiguez G., Yasseri T., San Miguel M., Kaski K., and Kertész J. (2013) Phys. Rev. Lett. 110, 088701


1
0.5
0
x_i, A
0
0.25
0.5
I
II
III
t ($\times 10^3$)
μ_A
convergence parameter

Calibration



Main Messages

1. Collaboration encourages consensus

PLOS ONE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Understanding and coping with extremism in an online collaborative environment: A data-driven modeling

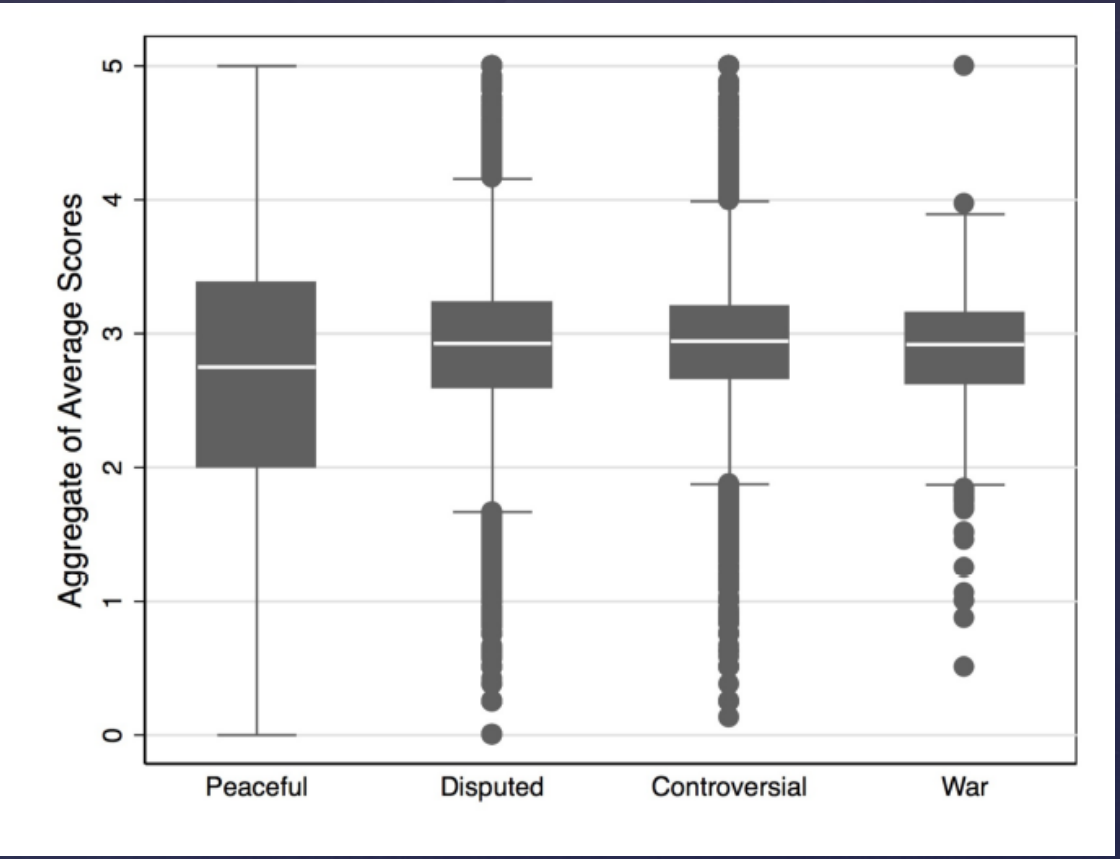
Csilla Rudas, Olivér Surányi, Taha Yasseri, János Török

Published: March 21, 2017 • https://doi.org/10.1371/journal.pone.0173561

2. Banning editors with extreme opinions postpones the consensus

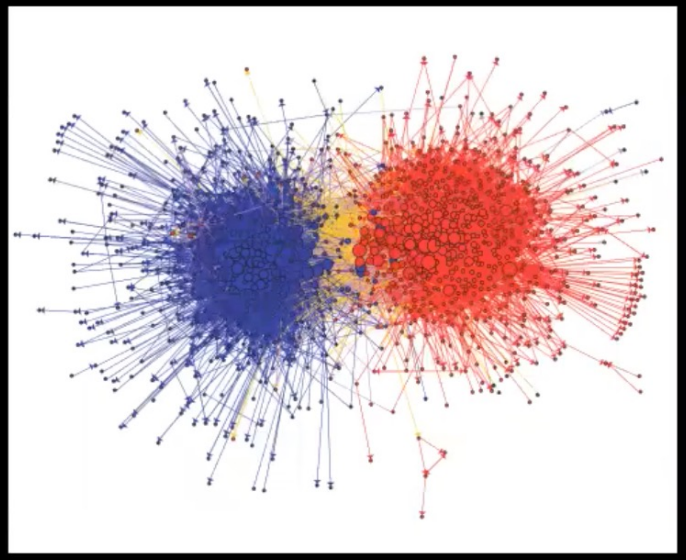
There is an upside to anger—or, at least, moderate discord….

Evan Rachel Wood



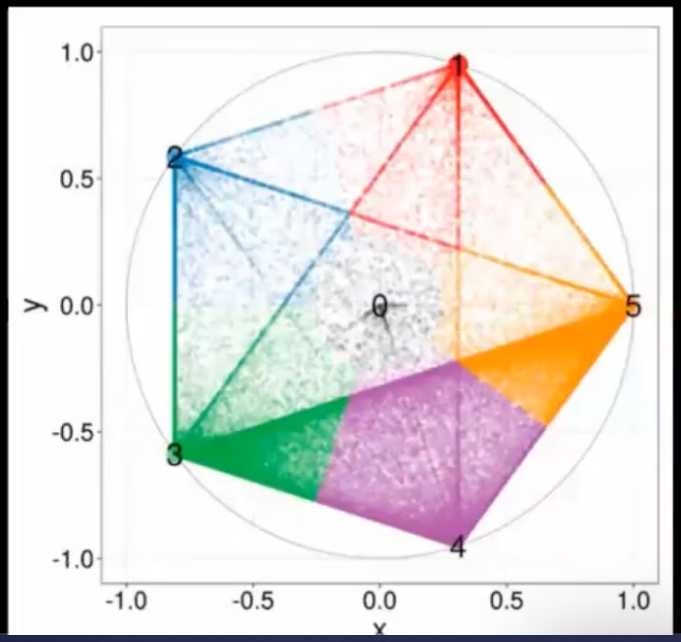
Fiadeiro & Yasseri (Forthcoming)

3. Certain amount of conflict increases the quality



Blogs

Adamic & Glance, 2005

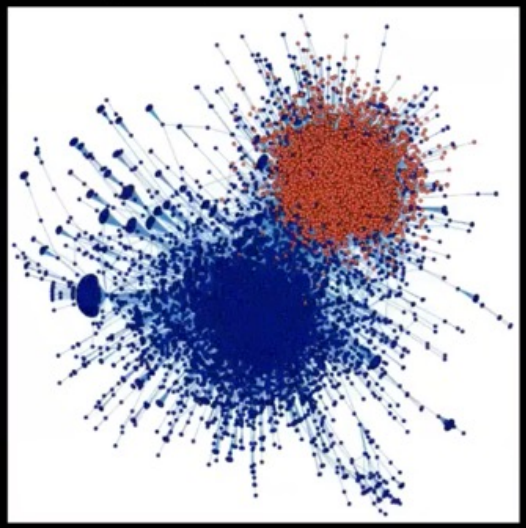


Facebook

Schmidt, 2017

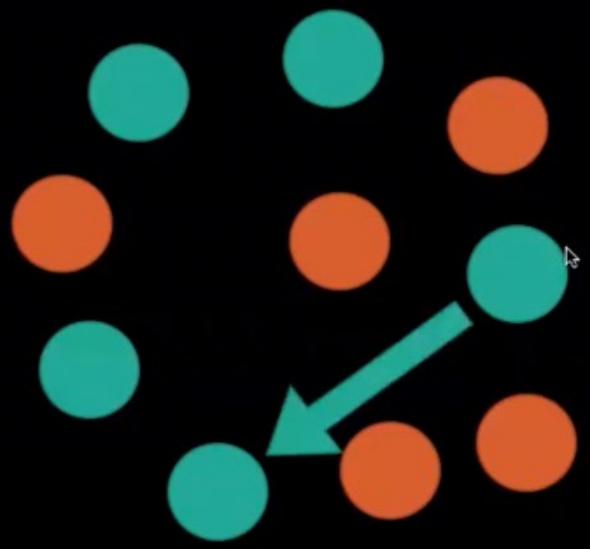
Twitter

Conover, Gonçalves, Flammini, Menczer, 2012



Polarization

A model with homophily



Blex, C., & Yasseri, T. (2020). Positive algorithmic bias cannot stop fragmentation in homophilic networks. *The Journal of Mathematical Sociology*, 1-18.

$$p_t = 1 - q_t = f\left(\frac{\mathbb{E}\left[E_{d,t-1}\right]}{E_{t-1}}\right) \tag{1}$$

$$\mathbb{E}\left[E_{d,i,t}\right] = q_{i,t} + \sum_{j=0}^{t} k_{d,t-j} q_{i,t-1-j} \tag{2}$$

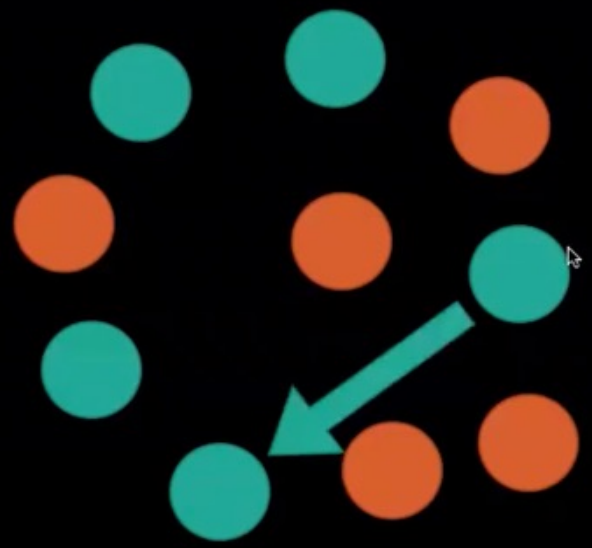
where $f' > 0$ and $f'' < 0$

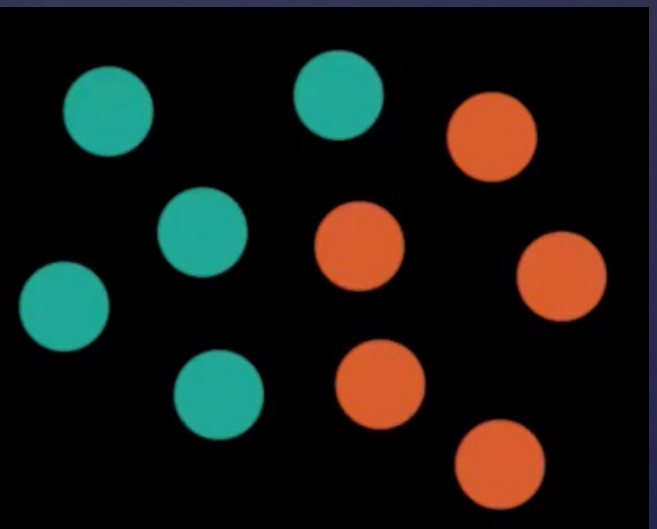
The derivative of the probability of making similar connections with respect to the probability of making dissimilar connections is given by

$$\frac{\partial}{\partial q_{i,t-1}} p_{i,t} = f'\left(\frac{\mathbb{E}[E_{d,t-1}]}{E_{t-1}}\right) t^{-1} > 0 \tag{3}$$

This implies that as $t \to \infty$, $\frac{\partial}{\partial q_{i,t-1}} p_{i,t} \to 0$. And thus for all strictly monotonically increasing concave functions f on $[0,1]$, $p_{i,t} \to 1$, implying $q_{i,t} \to 0$ and $\mathbb{E}\left[E_{d,i,t}\right] \to 0$.

A model with homophily





Blex, C., & Yasseri, T. (2020). Positive algorithmic bias cannot stop fragmentation in homophilic networks. *The Journal of Mathematical Sociology*, 1-18.

$$p_t = 1 - q_t = f\left(\frac{\mathbb{E}\left[E_{d,t-1}\right]}{E_{t-1}}\right) \tag{1}$$

$$\mathbb{E}\left[E_{d,i,t}\right] = q_{i,t} + \sum_{j=0}^{t} k_{d,t-j} q_{i,t-1-j} \tag{2}$$

where $f' > 0$ and $f'' < 0$

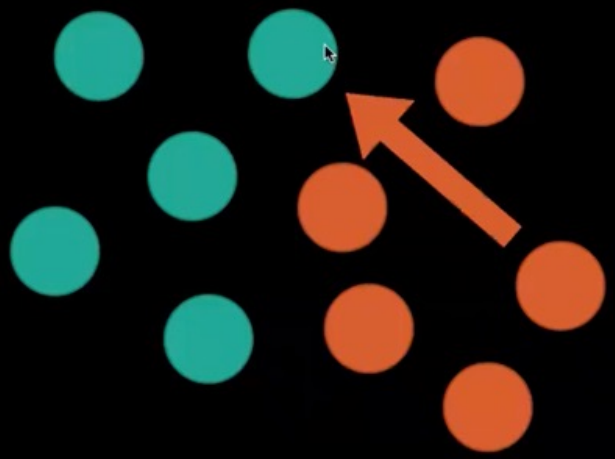
The derivative of the probability of making similar connections with respect to the probability of making dissimilar connections is given by

$$\frac{\partial}{\partial q_{i,t-1}} p_{i,t} = f'\left(\frac{\mathbb{E}[E_{d,t-1}]}{E_{t-1}}\right) t^{-1} > 0 \tag{3}$$

This implies that as $t \to \infty$, $\frac{\partial}{\partial q_{i,t-1}} p_{i,t} \to 0$. And thus for all strictly monotonically increasing concave functions f on $[0,1]$, $p_{i,t} \to 1$, implying $q_{i,t} \to 0$ and $\mathbb{E}\left[E_{d,i,t}\right] \to 0$.

A model with homophily and **weak ties**

Granovetter (1983)
Bakshy, et al., (2012)



$$\mathbb{E}[E_{d,i,t}] = q_{i,t} + q_{i,t}p_{k,t} + p_{i,t}q_{k,t} + k_d E_{d,i,t-1}$$

After substituting the expression for $E_{d,k,t}$, iterating the equation backwards and considering it in continuous time:

$$\mathbb{E}[E_{d,i,t}] = q_{i,t} + q_{i,t}p_{k,t} + p_{i,t}q_{k,t} + \sum_{j=0}^{t} k_{d,t-j}[q_{i,t-j-1} + q_{i,t}p_{k,t-j-1} + p_{i,t-j-1}q_{k,t-j-1}]$$

Theorem 2 For any concave strictly monotonically increasing function $f : [0,1] \rightarrow [0,1]$, p_t converges to 1 and thus $\mathbb{E}[E_{d,t}]$ converges to zero, even in the presence of secondary ties.

4.2. Proof

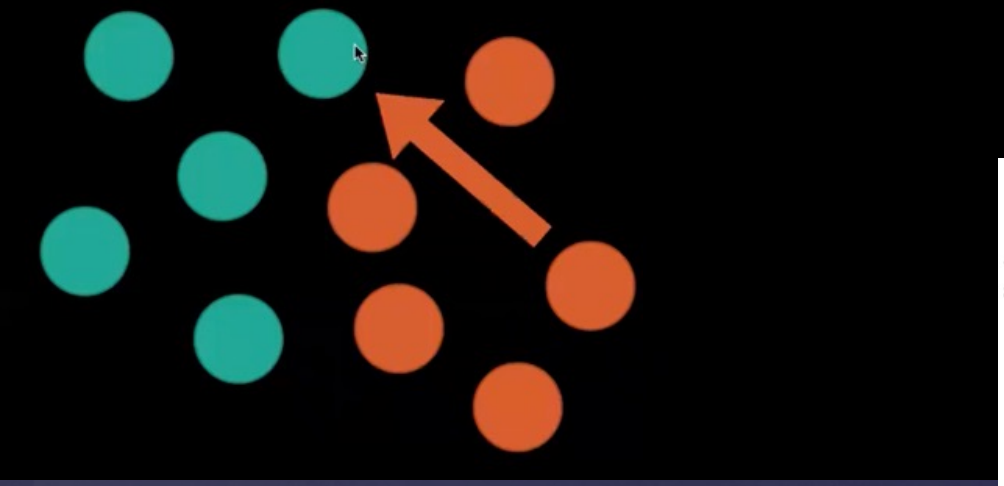
The derivative of Eq (3) has now changed to

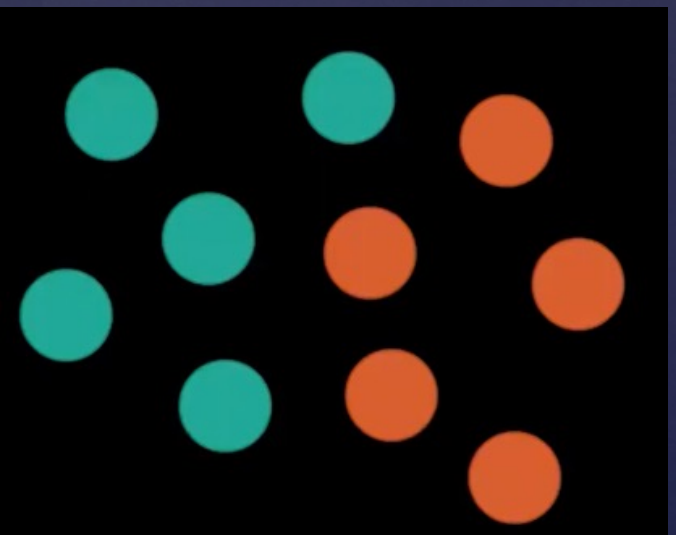
$$\frac{\partial}{\partial q_{i,t-1}} p_{i,t} = f'\left(\frac{\mathbb{E}[E_{d,t-1}]}{E_{t-1}}\right)\frac{1+p_{k,t-1}}{t} = f'\left(\frac{\mathbb{E}[E_{d,t-1}]}{E_{t-1}}\right)\frac{1+(1-q_{k,t-1})}{t} > 0$$

Given that $p_{k,t}$ is bounded in [0,1] for all t, this means that the derivative converges to zero. Note that the convergence rate is being slowed down compared to the result of the *Homophily Theorem*, but that this deceleration quickly declines over time as t increases and by symmetry $q_{k,t-1}$ decreases. Thus, for all strictly monotonically increasing functions f on [0,1], $p_{i,t} \rightarrow 1$, implying $q_{i,t} \rightarrow 0$ and $\mathbb{E}[E_{d,i,t}] \rightarrow 0$, even in the presence of weak ties.

A model with homophily and **weak ties**

Granovetter (1983)
Bakshy, et al., (2012)





$$\mathbb{E}[E_{d,i,t}] = q_{i,t} + q_{i,t}p_{k,t} + p_{i,t}q_{k,t} + k_d E_{d,i,t-1}$$

After substituting the expression for $E_{d,k,t}$, iterating the equation backwards and considering it in continuous time:

$$\mathbb{E}[E_{d,i,t}] = q_{i,t} + q_{i,t}p_{k,t} + p_{i,t}q_{k,t} + \sum_{j=0}^{t} k_{d,t-j}[q_{i,t-j-1} + q_{i,t}p_{k,t-j-1} + p_{i,t-j-1}q_{k,t-j-1}]$$

Theorem 2 For any concave strictly monotonically increasing function $f : [0,1] \rightarrow [0,1]$, p_t converges to 1 and thus $\mathbb{E}[E_{d,t}]$ converges to zero, even in the presence of secondary ties.

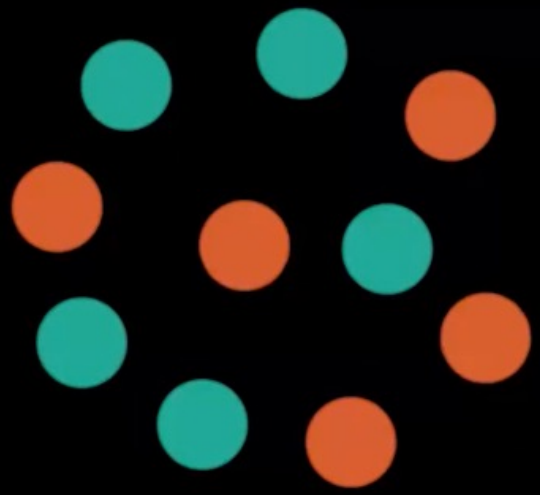
4.2. Proof

The derivative of Eq (3) has now changed to

$$\frac{\partial}{\partial q_{i,t-1}} p_{i,t} = f'\left(\frac{\mathbb{E}[E_{d,t-1}]}{E_{t-1}}\right)\frac{1+p_{k,t-1}}{t} = f'\left(\frac{\mathbb{E}[E_{d,t-1}]}{E_{t-1}}\right)\frac{1+(1-q_{k,t-1})}{t} > 0$$

Given that $p_{k,t}$ is bounded in [0,1] for all t, this means that the derivative converges to zero. Note that the convergence rate is being slowed down compared to the result of the *Homophily Theorem*, but that this deceleration quickly declines over time as t increases and by symmetry $q_{k,t-1}$ decreases. Thus, for all strictly monotonically increasing functions f on [0,1], $p_{i,t} \rightarrow 1$, implying $q_{i,t} \rightarrow 0$ and $\mathbb{E}[E_{d,i,t}] \rightarrow 0$, even in the presence of weak ties.

A positive **algorithmic bias?**



$$\mathbb{E}[E_{d,i,t}] = q_{i,t} + \sum_{j=0}^{t} k_{d,t-j} q_{i,t-1-j} + \sum_{j=0}^{t} k_{b,t-j} \sum_{l=1}^{t} \left(\phi_{i,t-l} - k_{s,t-l}\right)\left(t - l - 1 - q_{i,t-l-1}\right) + \left(\phi_{i,t} - k_{s,t}\right)\left(t - 1 - q_{i,t-1}\right)$$

,

with $k_d < k_b < k_s$ and k being assumed to be a Weibull process $k = 1 - \lambda^{-\gamma}\gamma t^{\gamma-1}$, where $\lambda \geq 1$ is a scale parameter and $\gamma < 1$.

Consider the equation above with all variables at their steady-state value, denoted here by an *. N.B. that k in all cases converges to one for large t.

$$\mathbb{E}[E_{d,i,t^*}] = q^* + t^* q^* + t^*(t^* - 1)(\phi^* - 1)(t^* - 2 - q^*) + (\phi^* - 1)(t^* - 1 - q^*)$$

Since ϕ is bounded on [0,1] for all t, the optimal and in this case, the minimal level of bias is the boundary solution $\phi^* = 1$. This implies that every similar connection is changed to a dissimilar connection. Consequently,

$$\mathbb{E}[E_{d,i,t^*}] = (1 + t^*)q^*$$

And

$$\frac{\mathbb{E}[E_{d,i,t^*}]}{\mathbb{E}[E_{s,i,t^*}]} = \frac{(1 + t^*)q^*}{(1 + t^*)p^*} = \frac{q^*}{p^*}$$

For the probability of making similar connections in the steady-state this implies

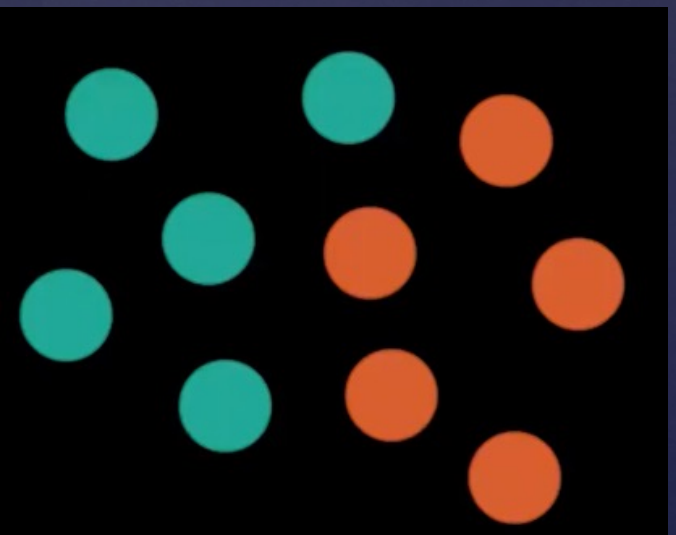
$$p^* = f\left(\frac{\mathbb{E}[E_{d,i,t-1^*}]}{t^*}\right),$$

which in the limit converges to

$$p^* = f(1) \text{ for } q_{t^*-1} > 0$$

The latter implies $p^* = 1$ and therefore $q^* = 0$ and $\mathbb{E}[E_{d,i,t^*}] = 0$. Thus, for any $\phi^* < 1$, $\mathbb{E}[E_{d,i,t^*}]$ diverges and in fact becomes negative, while for $\phi^* = 0$,

A positive **algorithmic bias?**



$$\mathbb{E}\left[E_{d,i,t}\right] = q_{i,t} + \sum_{j=0}^{t} k_{d,t-j} q_{i,t-1-j} + \sum_{j=0}^{t} k_{b,t-j} \sum_{l=1}^{t} \left(\phi_{i,t-l} - k_{s,t-l}\right)\left(t - l - 1 - q_{i,t-l-1}\right) + \left(\phi_{i,t} - k_{s,t}\right)\left(t - 1 - q_{i,t-1}\right)$$

,

with $k_d < k_b < k_s$ and k being assumed to be a Weibull process $k = 1 - \lambda^{-\gamma}\gamma t^{\gamma-1}$, where $\lambda \geq 1$ is a scale parameter and $\gamma < 1$.

Consider the equation above with all variables at their steady-state value, denoted here by an *. N.B. that k in all cases converges to one for large t.

$$\mathbb{E}\left[E_{d,i,t^*}\right] = q^* + t^* q^* + t^*(t^* - 1)(\phi^* - 1)(t^* - 2 - q^*) + (\phi^* - 1)(t^* - 1 - q^*)$$

Since ϕ is bounded on [0,1] for all t, the optimal and in this case, the minimal level of bias is the boundary solution $\phi^* = 1$. This implies that every similar connection is changed to a dissimilar connection. Consequently,

$$\mathbb{E}\left[E_{d,i,t^*}\right] = (1 + t^*)q^*$$

And

$$\frac{\mathbb{E}\left[E_{d,i,t^*}\right]}{\mathbb{E}\left[E_{s,i,t^*}\right]} = \frac{(1 + t^*)q^*}{(1 + t^*)p^*} = \frac{q^*}{p^*}$$

For the probability of making similar connections in the steady-state this implies

$$p^* = f\left(\frac{\mathbb{E}\left[E_{d,i,t-1^*}\right]}{t^*}\right),$$

which in the limit converges to

$$p^* = f(1) \text{ for } q_{t^*-1} > 0$$

The latter implies $p^* = 1$ and therefore $q^* = 0$ and $\mathbb{E}\left[E_{d,i,t^*}\right] = 0$. Thus, for any $\phi^* < 1$, $\mathbb{E}\left[E_{d,i,t^*}\right]$ diverges and in fact becomes negative, while for $\phi^* = 0$,

4. Costless rewiring and homophily are responsible for polarization and they are hard to fight

4. Costless rewiring and homophily are responsible for polarization and they are hard to fight

Social influence and unfollowing accelerate the emergence of echo chambers

Kazutoshi Sasahara ✉, Wen Chen, Hao Peng, Giovanni Luca Ciampaglia, Alessandro Flammini & Filippo Menczer

Journal of Computational Social Science (2020) | Cite this article

Solution?

Ullstein Bild

Pre-Internet life?

arxiv.org/pdf/2104.13754.pdf

Can crowdsourcing rescue the social marketplace of ideas?

Taha Yasseri[1,2,3,4]*, Filippo Menczer[5,6]

Introducing Birdwatch, a community-based approach to misinformation

By Keith Coleman

Monday, 25 January 2021

A

Amazon News @amazonnews · Mar 24

Replying to @repmarkpocan

1/2 You don't really believe the peeing in bottles thing, do you? If that were true, nobody would work for us. The truth is that we have over a million incredible employees around the world who are proud of what they do, and have great wages and health care from day one.

14.3K 19.4K 4.7K

B

Currently rated helpful

Informative · Cites high-quality sources

Potentially misleading Mar 25

Amazon has a documented history of labor violations, including pushing employees to work so much they do not have time to use the restroom.
https://www.theguardian.com/technology/2020/feb/05/amazon-workers-protest-unsafe-grueling-conditions-warehouse

https://www.newsweek.com/amazon-drivers-warehouse-conditions-workers-complains-jeff-bezos-bernie-1118849

https://www.npr.org/2020/07/31/897836765/amazon-workers-respond-to-jeff-bezos-testimony-before-congress

C

Currently not rated helpful

Sources not included or unreliable · Misses key points

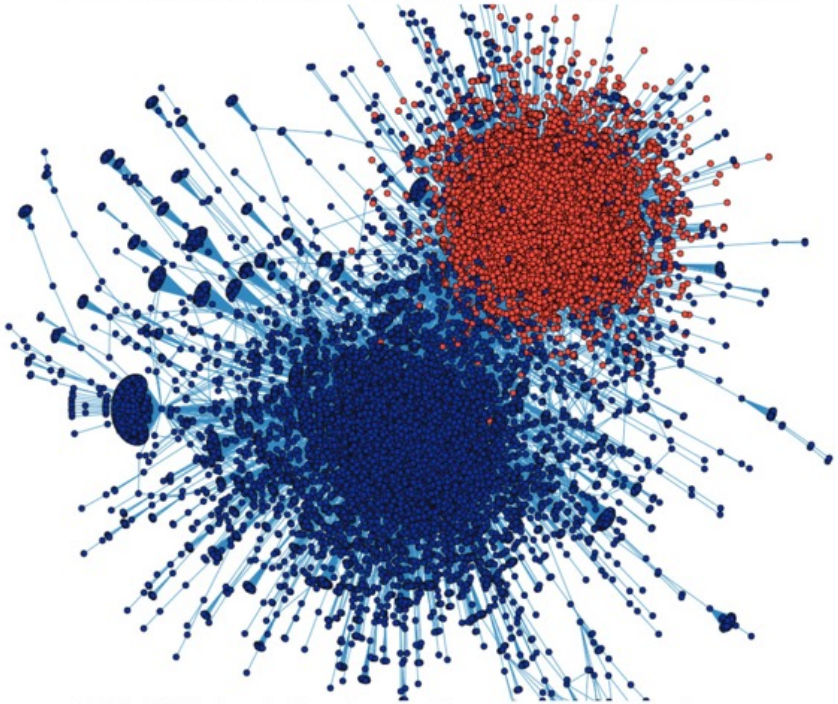
Potentially misleading Mar 25

Amazon workers are treated awfully. Thank you.

Data:

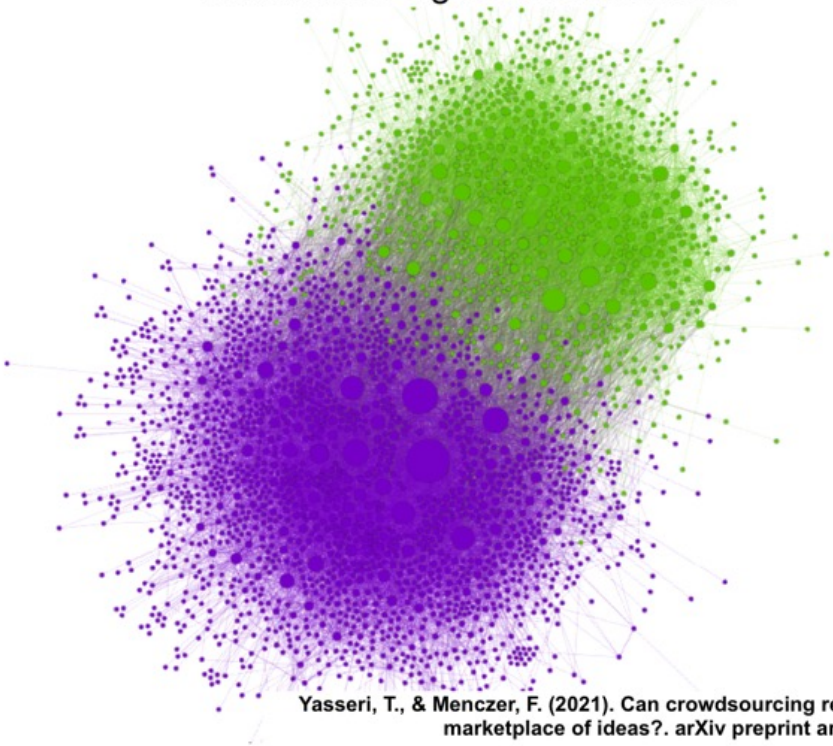
2,323 Tweets
2,756 Birdwachers
5,798 notes
36,786 votes

Polarization in the Twitter's Retweet network



Conover, M. D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011, July). Political polarization on twitter. In Fifth international AAAI conference on weblogs and social media.

Polarization among the Birdwatch contributors' agreement network



Yasseri, T., & Menczer, F. (2021). Can crowdsourcing rescue the social marketplace of ideas?. arXiv preprint arXiv:2104.13754.

Hang on!

Crowd-sourcing, Collaboration, Validation?

Experiment!

Collaborative "birdwatching"

Direct Engagement
Co-Ownership

Design matters
Connections matter
Community matters

Thank you!

Robert Sumi, András Rung, András Kornai, János Török, János Kertész, Gerardo Iñiguez, Kimmo Kaski, Maxi San Miguel, Anselm Spoerri, Mark Graham, Milena Tsvetkova, Ruth Garcia, Csilla Rudas, Olivér Surányi, Joao Fiadeiro, and Luciano Floridi, Fil Menczer, Chris Blex.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Program under grant agreement No 645043.

@TahaYasseri

School of Sociology, University College Dublin

tahayasseri.com