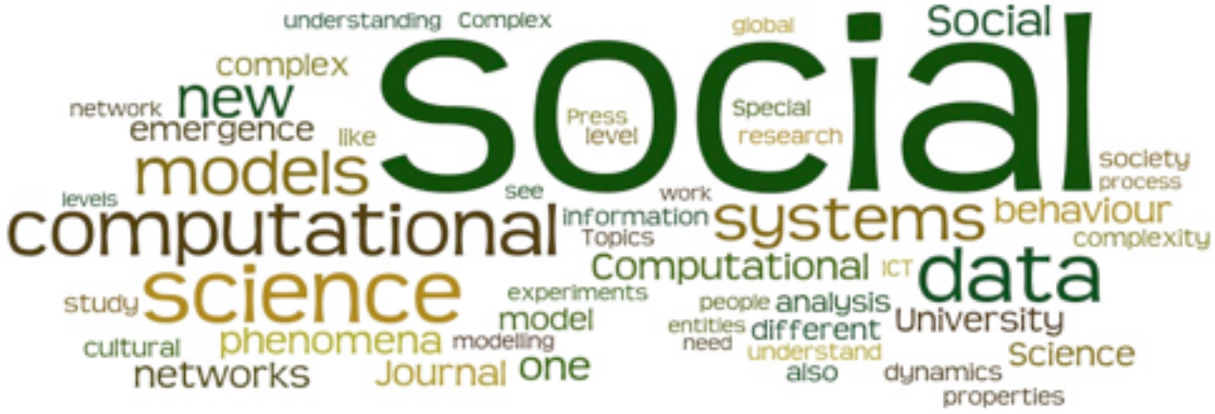
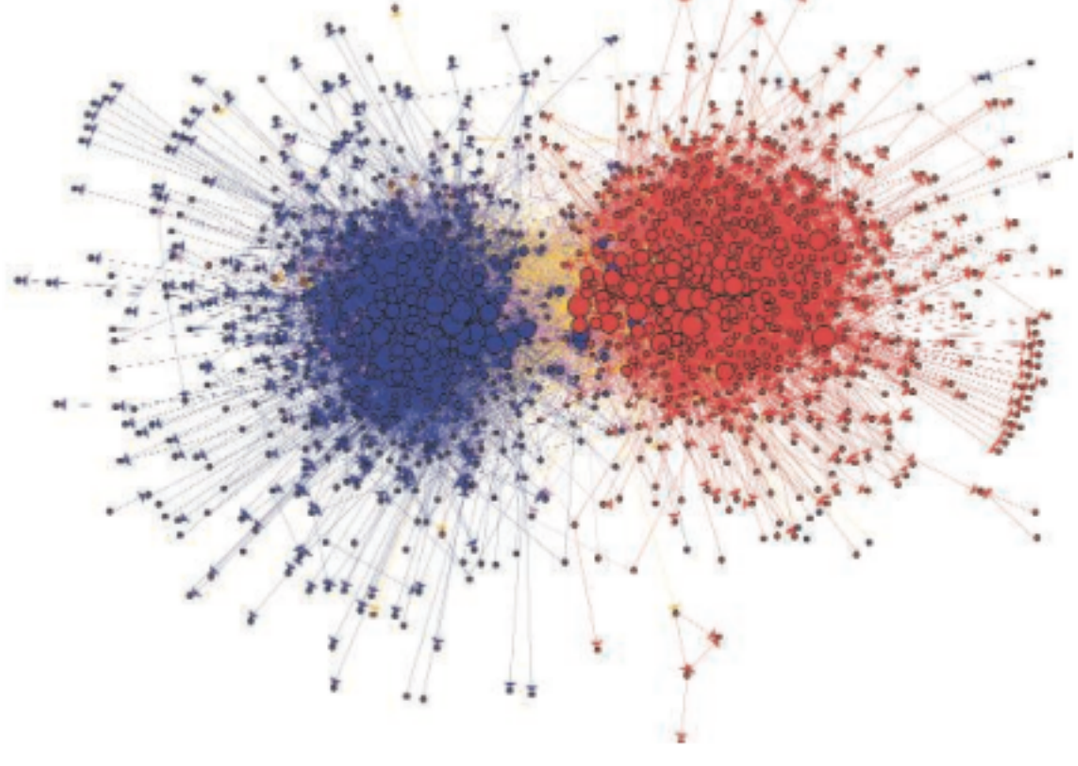
CSSLab @ IMT LUCCA

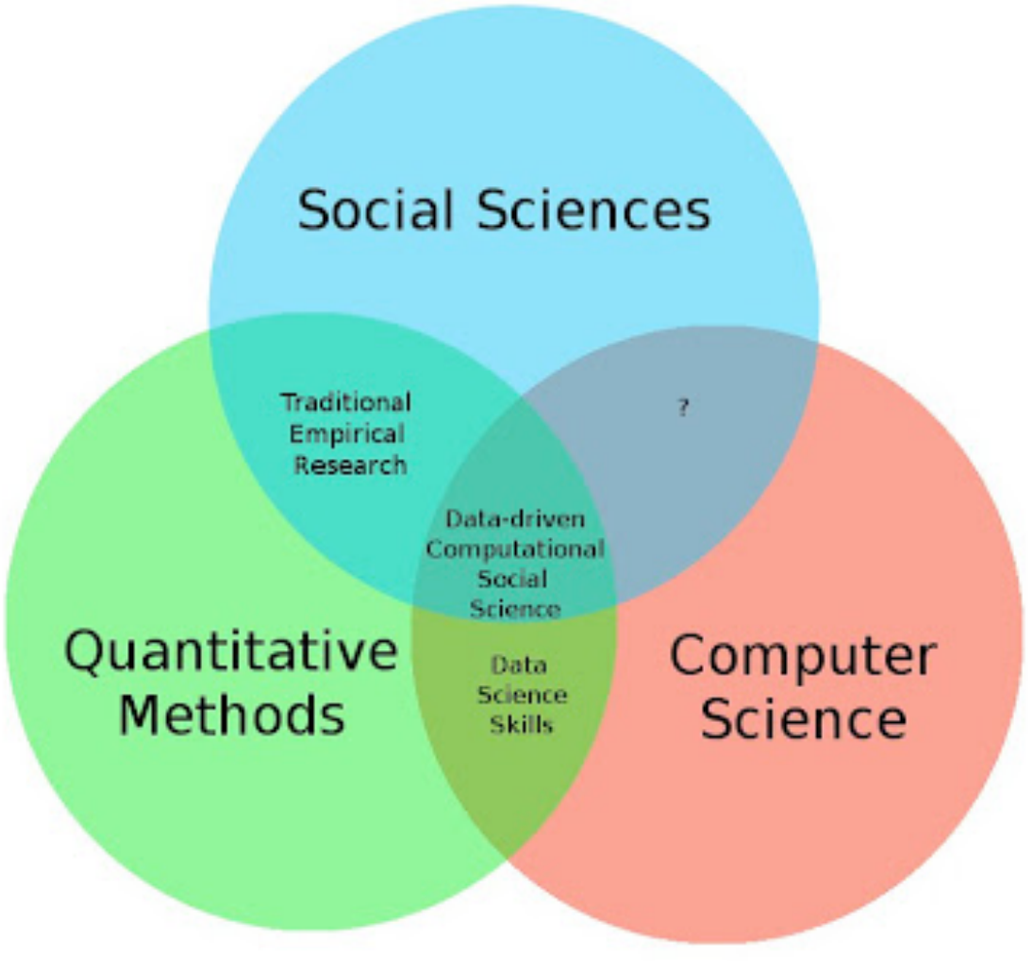
Walter Quattrociocchi
walter.quattrociocchi@imtlucca.it

Computational social science

refers to the academic sub-disciplines concerned with **computational approaches** to the **social sciences**.







Computational Social Science

Computational Social Science:
"The science that investigates social phenomena through the medium of computing and algorithmic data processing."
[adapted from CSSSA]

CSSSA: http://computationalsocialscience.org/

Science
AAAS

SOCIAL SCIENCE

Computational Social Science

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.

David Lazer,[1] Alex Pentland,[2] Lada Adamic,[3] Sinan Aral,[2,4] Albert-László Barabási,[5] Devon Brewer,[6] Nicholas Christakis,[7] 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]

We live life in the network. We check our e-mails regularly, make mobile phone calls from almost any location, swipe transit cards to use public transportation, and make purchases with credit cards. Our movements in public places may be captured by video cameras, and our medical records stored as digital files. We may post blog entries accessible to anyone, or maintain friendships through online social networks. Each of these transactions leaves digital traces that can be compiled into comprehensive pictures of both individual and group behavior, with the potential to transform our understanding of our lives, organizations, and societies.

ment agencies such as the U.S. National Security Agency. Computational social science could become the exclusive domain of private companies and government agencies. Alternatively, there might emerge a privileged set of academic researchers presiding over private data from which they produce papers that cannot be critiqued or replicated. Neither scenario will serve the long-term public interest of accumulating, verifying, and disseminating knowledge.

What value might a computational social science—based in an open academic environment—offer society, by enhancing understanding of individuals and collectives? What are the

- Harvard iQS,
- Stanford IRiSS,
- CMU CASOS,
- ESRC COSMOS
- Web Observatories
- …

Computational Social Science: Example

Stanley Milgram (1967)

- Social Scientist
- **Theory**: A small world
- 6 degrees of separation

Jure Leskovec (2008)

- Computer Scientist
- **(Found) Data**: 240 mio users
- 7 degrees of separation

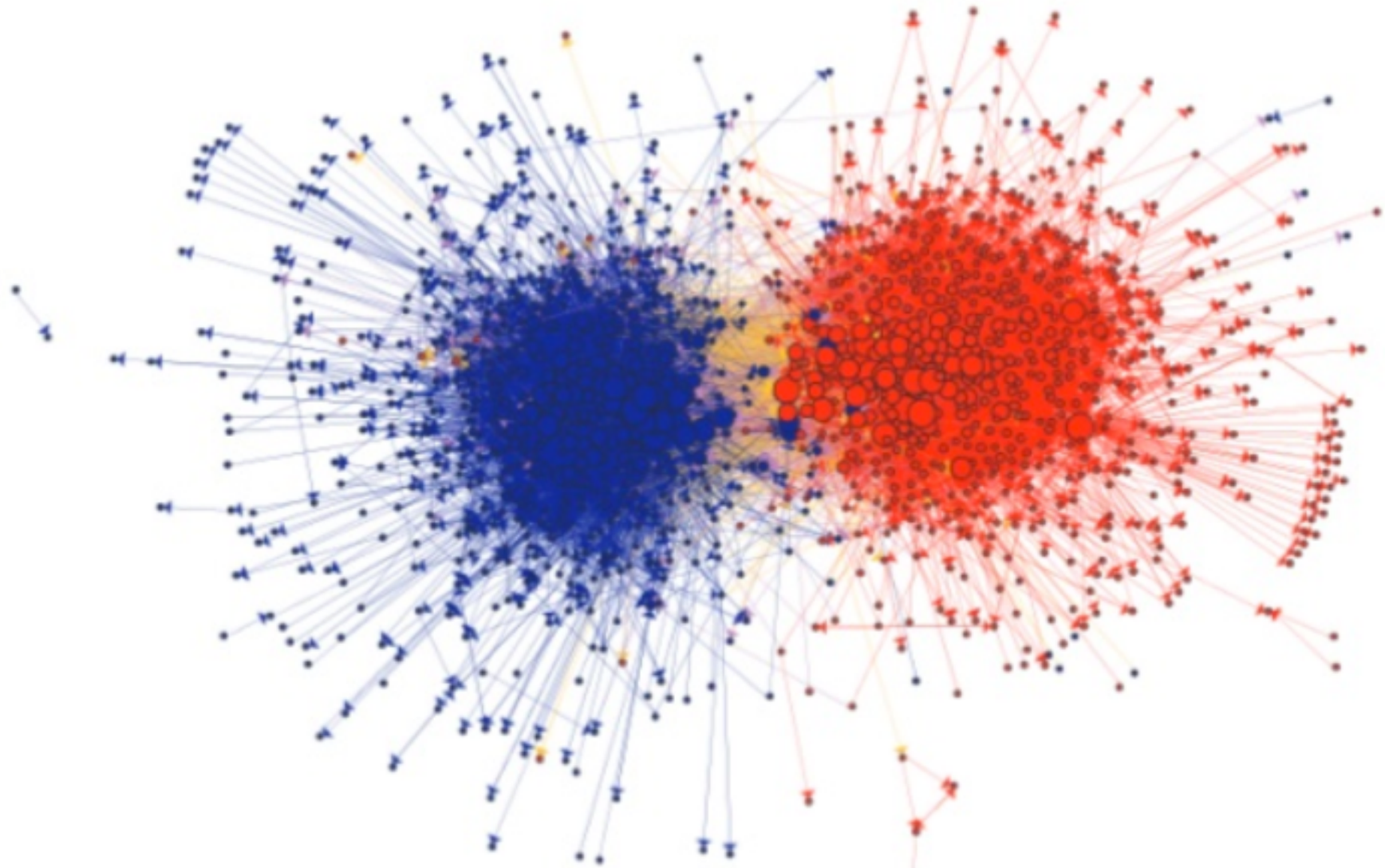
What disciplines are interested in microposts?

Table 1. Count of Twitter Studies by Discipline (2007-2012)

	2007	2008	2009	2010	2011	2012	Total
Business				3	9	3	15
Communications	1		5	10	21	15	52
Computer Science		1	8	45	41	50	145
Economics			3	5	5	2	15
Education		1	3	1	6	2	13
English				1	4	1	6
Environmental Sciences					1		1
Geographic Information Systems				2	4		6
Information Science	1		6	17	33	25	82
Law				2			2
Mathematics					3		3
Medicine				2	7	5	14
Physics					2		2
Political Science			1	5	4	3	13
Psychology			1		1	1	3
Sociology			1	2	4	2	9
Sport Sciences				1			1
Total	2	2	28	96	145	109	382

Zimmer, Michael / Proferes, Nicholas. (In press). "A topology of Twitter research: Disciplines, methods, and ethics". To appear in Aslib Journal of Information Management 3(66), special issue on Twitter data analysis.

Assessing online conversational practices of political parties



Lada Adamic and Natalie Glance. The political blogosphere and the 2004 US election: divided they blog. *Proceedings of the 3rd international workshop on Link discovery*. ACM, 2005.

Assessing online conversational practices of political parties on Twitter

During the German National Election 2013

(a) Following (H=0.83) (b) Retweeting (H=0.90) (c) Mentioning (H=0.79)

Haiko Lietz, Claudia Wagner, Arnim Bleier, and Markus Strohmaier, When Politicians Talk: Assessing Online Conversational Practices of Political Parties on Twitter, The International AAAI Conference on Weblogs and Social Media (ICWSM2014), Ann Arbor, MI, US, 2014.

OUR CHALLENGES:

- Quantitative understanding of Social Dynamics
- Prediction (and Control) of Social Phenomena

Social Contagion **Innovation Diffusion**

Opinion Dynamics **Human Mobility**

RESEARCH STREAMS:

- Theoretical and Data-Driven **Models** of Social Phenomena (Multi agent or Equation based)

- Detecting **Observables** from data

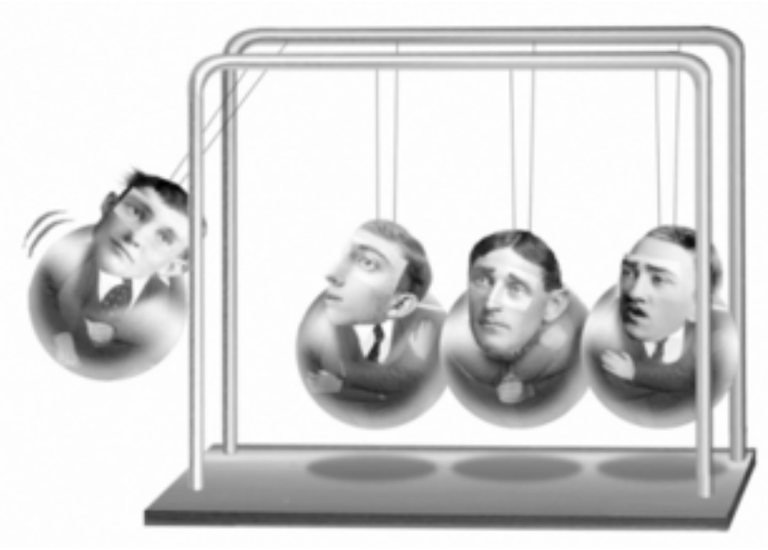
More realistic computational models (BIG PICTURE)

Epidemics:
Simple Contagion

Ideas and behaviors:
Complex Contagion

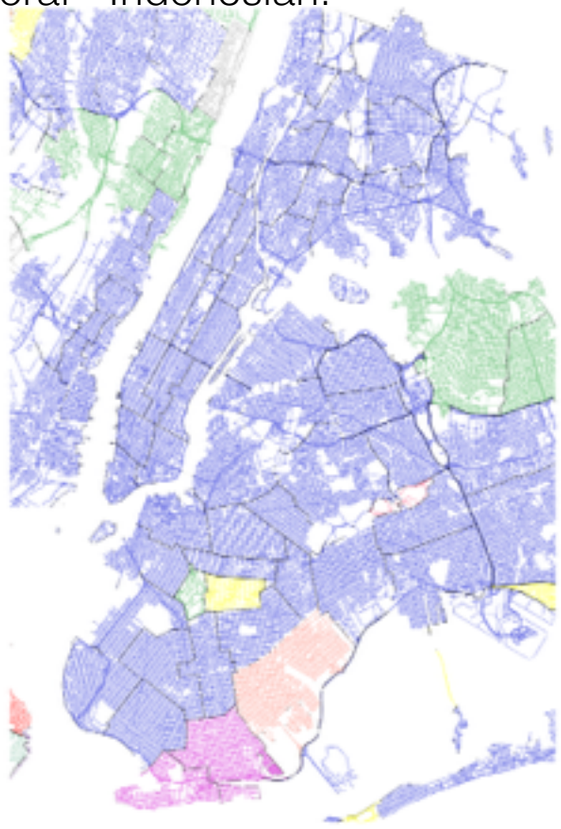
The idea of the social atom does not work

Intentionality
Trust
Information



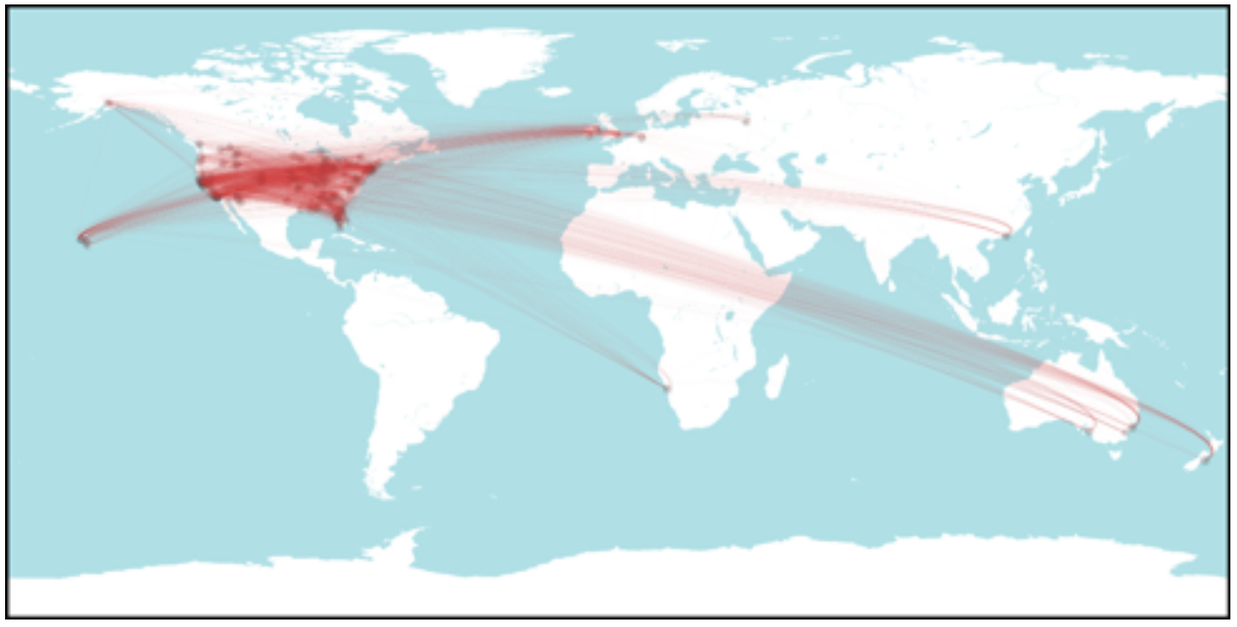
Exploiting Data: Tracking complex social dynamics

Language polarization in New York City, NY, USA.
The second language by district or municipality (in the case of New Jersey state) is shown. Blue - Spanish, Light Green - Korean, Fuchsia - Russian, Red - Portuguese, Yellow - Japanese, Pink - Dutch, Grey - Danish, Coral - Indonesian.



The Twitter of Babel: Mapping World Languages through Microblogging Platforms
Delia Mocanu,Andrea Baronchelli, Nicola Perra, Bruno Gonçalves, Qian Zhang, Alessandro Vespignani
PlosOne 2013

Users-Users network of @Occupy Movements around the World



Geographical and Social Communities within the Occupy Movement on Facebook
Qian Zhang, Delia Mocanu, Ana Pastore, Luca Rossi Walter Quattrociocchi and Alessandro Vespignani

in preparation

HOW TO DO MORE THAN THIS?

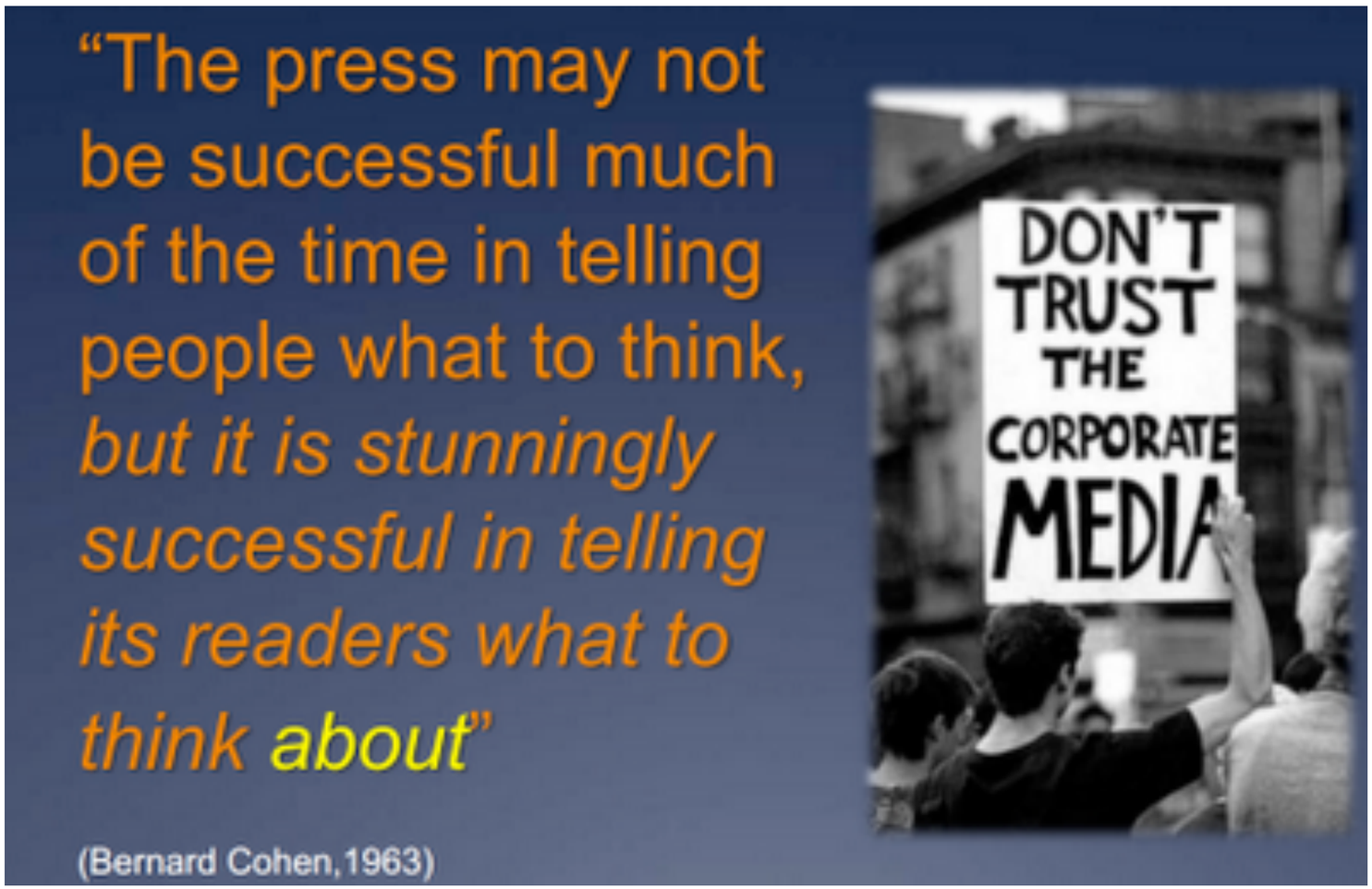
Challenges

Huberty (2014)

- N != all
  we have both N<all and N>all
- All (today) != All (tomorrow)
  user populations change
- Online behavior != Offline behavior
  multi-faceted identities
- Behavior of all (today) != Behavior of all (tomorrow)
  behavior changes and evolves

Challenge: <u>The cognitive aspects behind opinions</u>

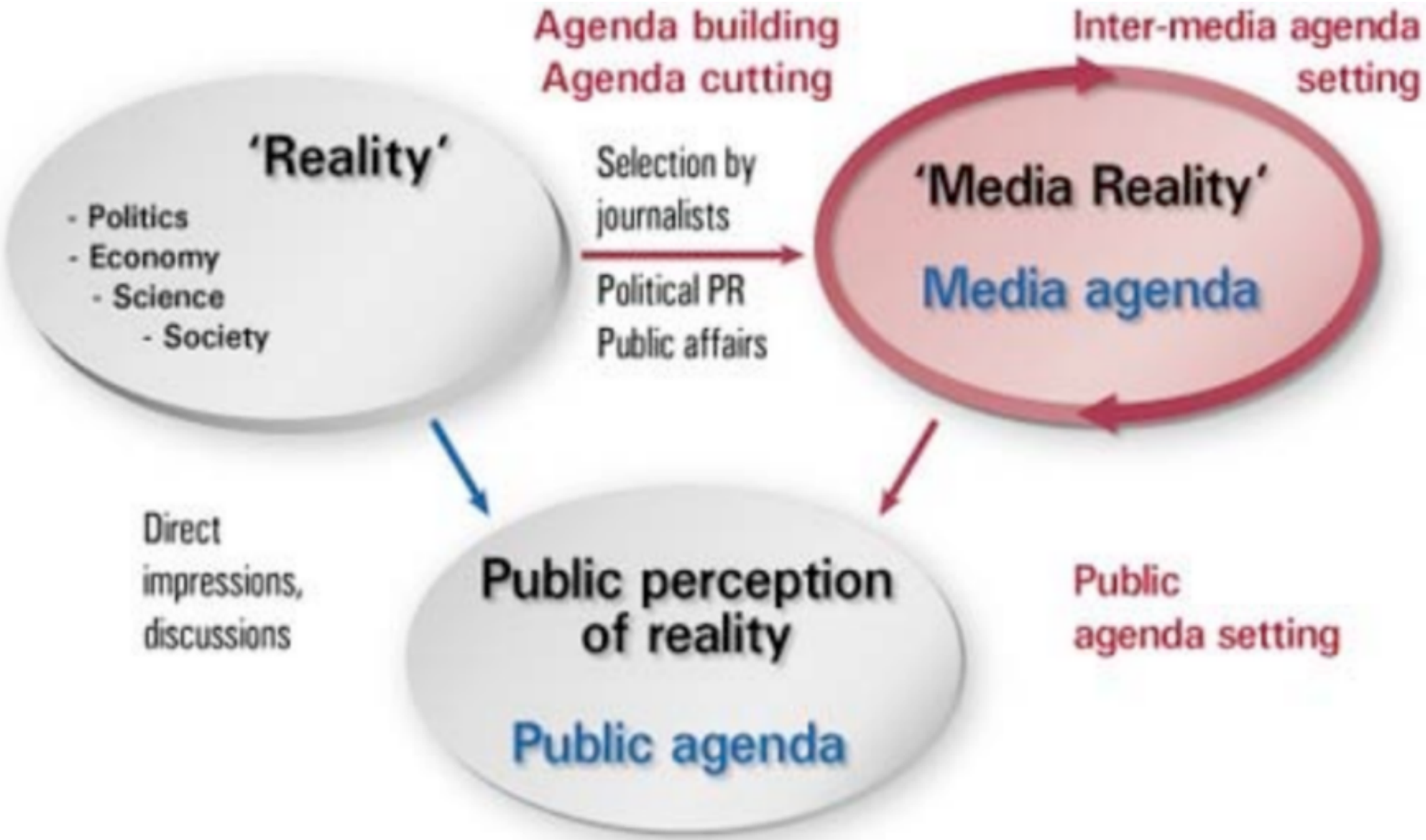
The interlink between information and opinions



Agenda Setting is the process of the mass media presenting certain issues frequently and prominently with the result that large segments of the public perceive those issues as more important than others.

<u>The more coverage an issue receives, the more important it is to people</u>

What is Agenda Setting?



INTERACTION BETWEEN INFORMATION SOURCES

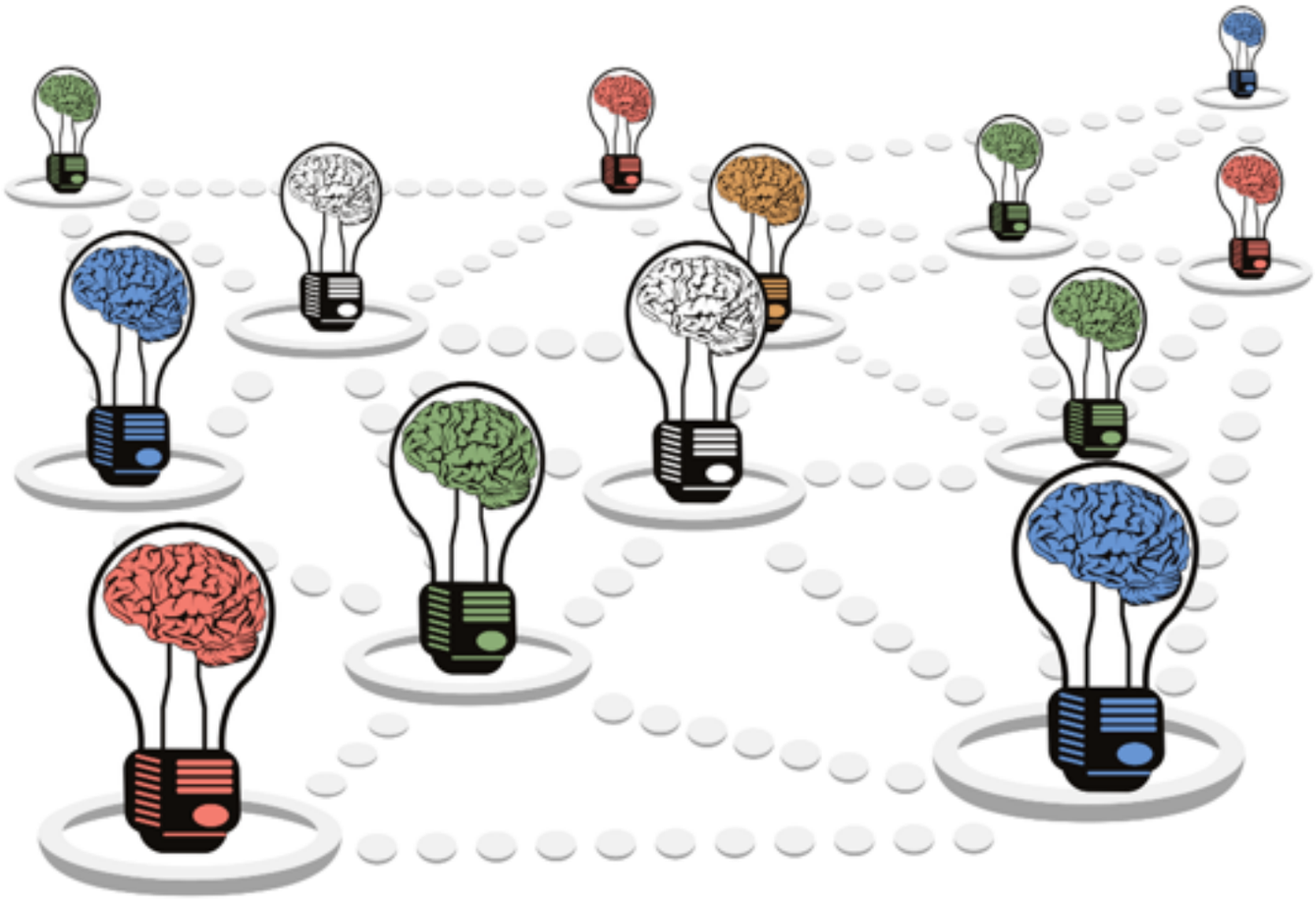
OLD MEDIA

- Follow the “Ritual of Objectivity”
- Publication patterns are driven by most followed sources (imitation) (Marlow 2005)

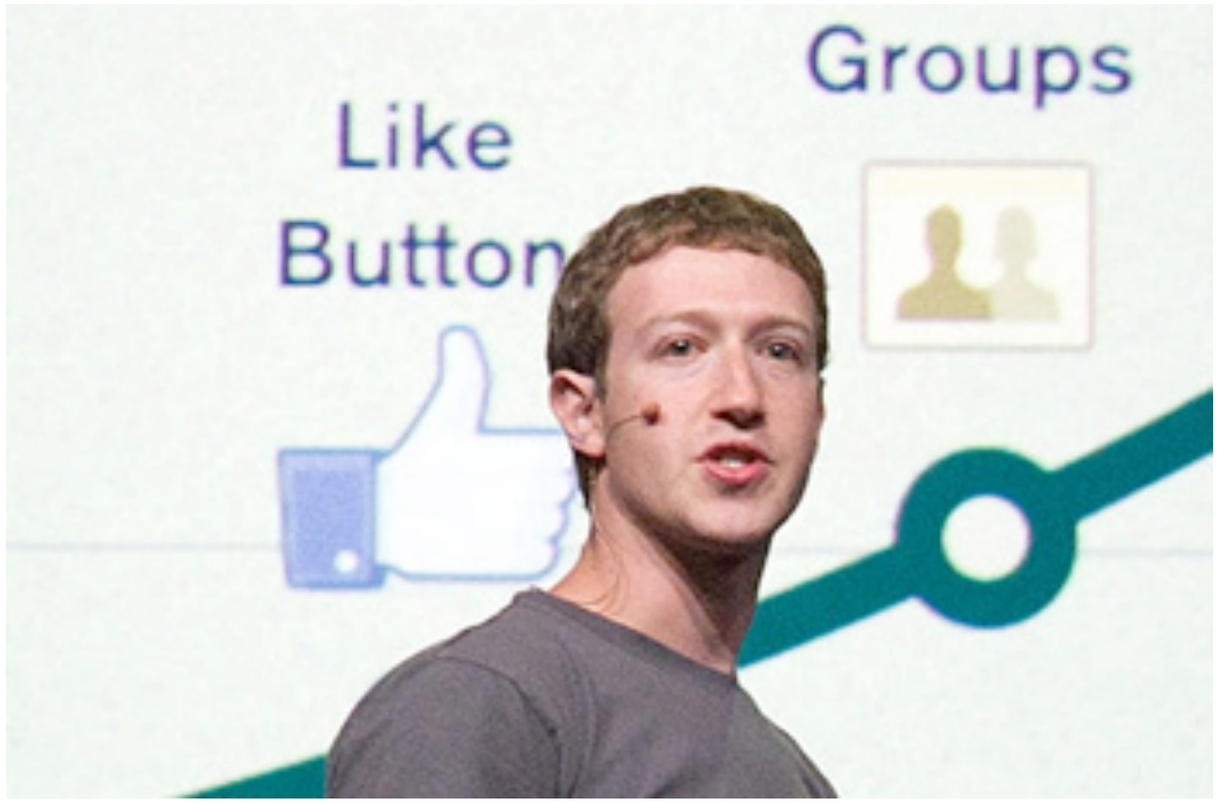
NEW MEDIA

- Information production is the work of interconnected actors spanning over organizations, professional identity and geographical location
- Imitation between mainstream agency on the internet (leader follower) in a shorter time span

COLLECTIVE INTELLIGENCE



- IS THE “MODERN SYNTHESIS BETWEEN PSYCHOLOGY AND THE NEW TECHNOLOGY”
- IT EMERGES IN THE FORM OF COLLECTIVE DEBATES ON BLOGS AND ON SOCIAL NETWORKS (TWITTER, FACEBOOK)
- COULD IMPROVE E-PARTICIPATION TO PUBLIC RELEVANT ISSUES



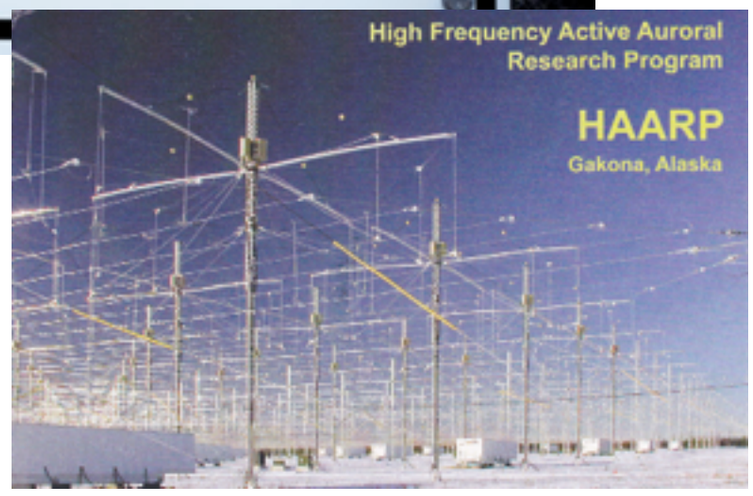
Mark Zuckerberg:

“We're not thinking about ourselves as a community — we're not trying to build a community — we're not trying to make new connections. [...]
What we're trying to do is just make it really efficient for people to communicate, get information and share information.
We always try to emphasize the utility component.”
Jul. 2007

COLLECTIVE INTELLIGENCE OR JUST ENTHUSIASTIC RHETORIC?

People are increasing their participation on public debates and internet is used as a medium for organization and discussion of political activities.

What about the quality of information **online**?



Conspiracy theories running rampant: How misinformation spreads on Facebook

Researchers studied how people interacted with "trolls" posting false information -- the results are terrifying

SARAH GRAY

Share 387 93 g+1 13 20

TOPICS: CONSPIRACY THEORIES, FACEBOOK, POLITICS, TECHNOLOGY, MISINFORMATION, INNOVATION NEWS, TECHNOLOGY NEWS, POLITICS NEWS

From the steady roll of theories on what happened to Malaysian Arlines Flight 370, to Sarah Palin's "death panels" panic, to Donald Trump's birther theories, misinformation spreads like wildfire in the age of Facebook.

In 2013, professor Walter Quattrociocchi of Northeastern University along with his team studied how more than 1 million Facebook users engaged with political information during the Italian election. During that election a post appeared titled: "Italian Senate voted and accepted (257 in favor and 165 abstentions) a law proposed by Senator Cirenga to provide policy makers with €134 billion Euros to find jobs in the event of electoral defeat."

WHAT ABOUT THE EFFECT OF ONLINE RUMORS?

The case of Italian political activism organizing on the web

TROLL:

Facebook pages used as a parodistic imitation of both alternative information sources and online political activism. Their activities range from controversial comments and posting satirical content mimicking alternative news sources, to the fabrication of purely fictitious statements, heavily unrealistic and sarcastic.

Sandro Pertini never said "*when the government does not do what people want must be fired with stones and sledgehammers.*" He has been President of the Republic (1978-1985).

"Italian Senate voted and accepted (257 in favor and 165 abstentions) a law proposed by Senator Cirenga aimed at funding with 134 billion Euros the policy-makers to find a job in case of defeat in the political competition."

LinkSicilia
Giornale OnLine

ULTIM'ORA e Crocetta giocano a ping pong... Caro presidente Crocetta, la Sicilia è senza soldi non per l'impugnativa

I Forconi: "Il Senato ha approvato una legge per i parlamentari in crisi". Chi non verrà rieletto, oltre alla buonuscita, si beccherà altri soldi. Sarà vero?

09 ago 2013 Scritto da Giulio Ambrosetti 3 Commenta

Like 227 Tweet 2 Share 3

Noi non ci crediamo. Intanto su facebook c'è anche il risultato della votazione...

La notizia la leggiamo sulla rete, per la precisione su **facebook**. L'ha postata il Movimento dei Forconi. **Stando a quanto si legge in questo post, ieri il Senato della Repubblica avrebbe approvato una legge (257 voti favorevoli, 165 astenuti) che istituisce il "Fondo per i parlamentari in crisi".**

A noi sembra uno scherzo. Soprattutto dopo che le più alte magistrature hanno fatto sapere che offendere la Repubblica e le istituzioni in generale – compresa quella parlamentari – è un reato, non possiamo credere che un ramo del Parlamento italiano abbia approvato una legge del genere.

Nel post dei Forconi si legge che il provvedimento stanzia 134 miliardi di euro (a qui secondo noi c'è un errore: si tratterà, quasi certamente, di 134 milioni di euro, perché 134 miliardi di euro sarebbe assurdo) per sostenere i parlamentari che, finita la legislatura, dovessero avere difficoltà a trovare lavoro.

Sostieni Lin

Donazi

to search type and hit ent

Si discut

Il popolo piange
Paolo Genco: "Pr
disparità di tratta
economico procu

EMIGRATO su Ar
Genco: "Pronta d

mary su Quelle lu
che vanno verso

Alexander su Mis
rimborsi corrispo
ai consiglieri com
sbagliati?

Giovanni Sorrenti
inizia valzer di de
Luigi Pirandello.
pilotati"

Tutte le ru

...e adesso
spogliati...

Caffe' Amaro

cirenga
Termine di ricerca

+ Aggiungi termine

Condividi

Interesse nel tempo

Intestazioni notizie Previsione

Incorpora

Interesse regionale

Regione | Città

Italia	100

Visualizza la variazione nel tempo

Incorpora

Incorpora

Ricerche correlate

Argomenti	Più cercati	In crescita

Volume di ricerca insufficiente per visualizzare i risultati.

Query	Più cercati	In crescita
senatore cirenga	100	
legge cirenga	60	
parlamentari in crisi	35	
legge senatore cirenga	25	

Movimentocinquestelletoseitu Torino ha condiviso la foto di Semplicemente me.
47 minuti fa

> CONDIVIDETE questa incredibile vicenda accaduta negli Stati Uniti d'America. I media non ne parlano, anzi dileggiano chi si azzarda a trattare l'argomento. QUESTE COSE NON LE SENTIRETE ALLA TV, CONDIVIDI PER INFORMARE. Quasimodo

Questa bella bambina si chiama Maggie Ducanan ed è nata a Boston in Connecticut. Sembrava una bambina sana e solare fino a quando non le sono state fatte le normali vaccinazioni post natali. Da quel momento la mamma della piccola Maggie ha notato che quando passava con sua figlia dalle porte antifurto dei centri commerciali la piccola Maggie faceva scattare l'allarme. La madre insospettita ha fatto una radiografia alla piccola Maggie e le è stato scoperto ed in seguito rimosso un microchip sub-cutaneo. E pensare che in Italia ancora crediamo che non esistano!!! CONDIVIDI questa storia!!!

Condividi 1

A 3 persone piace questo elemento.

Ogni condivisione o mi piace aumenta la conosicenza.
VOLEVATE L'ANALISI DELLE SCIE CHIMICHE? Ancora a dirci che sono scie di condenza. Noi abbiamo chiesto ai nostri amici dubbiosi di verificare i composti dell'aria ed ecco cosa ne è uscito. La gente deve sapere. Questo muro di omertà è diventato assurdo. Dobbiamo dubitare e amare nel dubbio.

Continuante a seguirci e a farci seguire su Simply Humans

Mi piace · Commenta · Condividi · 12 ore fa

Giuseppe Povia Strano che non siano ancora arrivati commenti in bimbomikiese tipo:
HHAHAHHAHAHAH PVIA ZE1 PROPQIO UN ERET1CO A CRDERE ANKORA A QSRE CZZATE HAHAHAHAHAHAHAH
LISLANDA KOSAA??? HAHAHAHAHAAHAHAH PNSA A CANTARE DI PICCION111 XDDD1!!!!
9 ore fa · Mi piace · 2

Beyond pure observation of data by triggering viral phenomena.

DATA ANALYSIS

The relationship between online political activism and online information sources

2.4 Million Facebook users

REAL

TROLL (1.6M sharing)

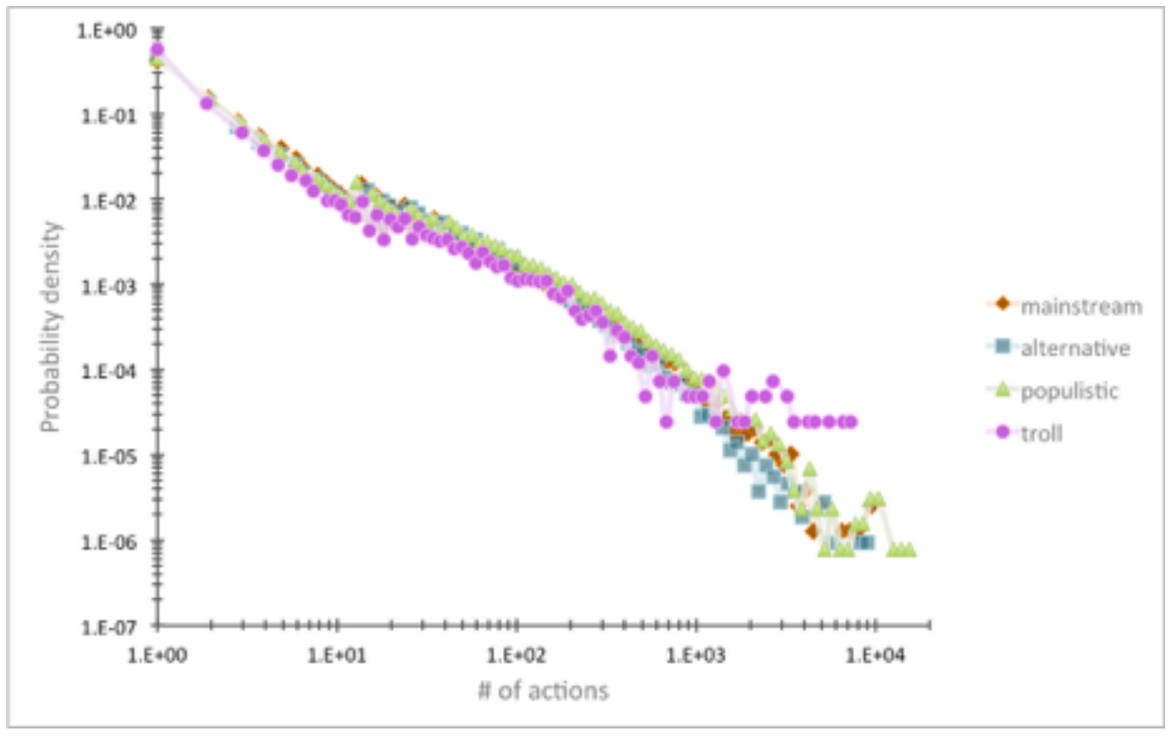
Collective attention in the age of (mis)information

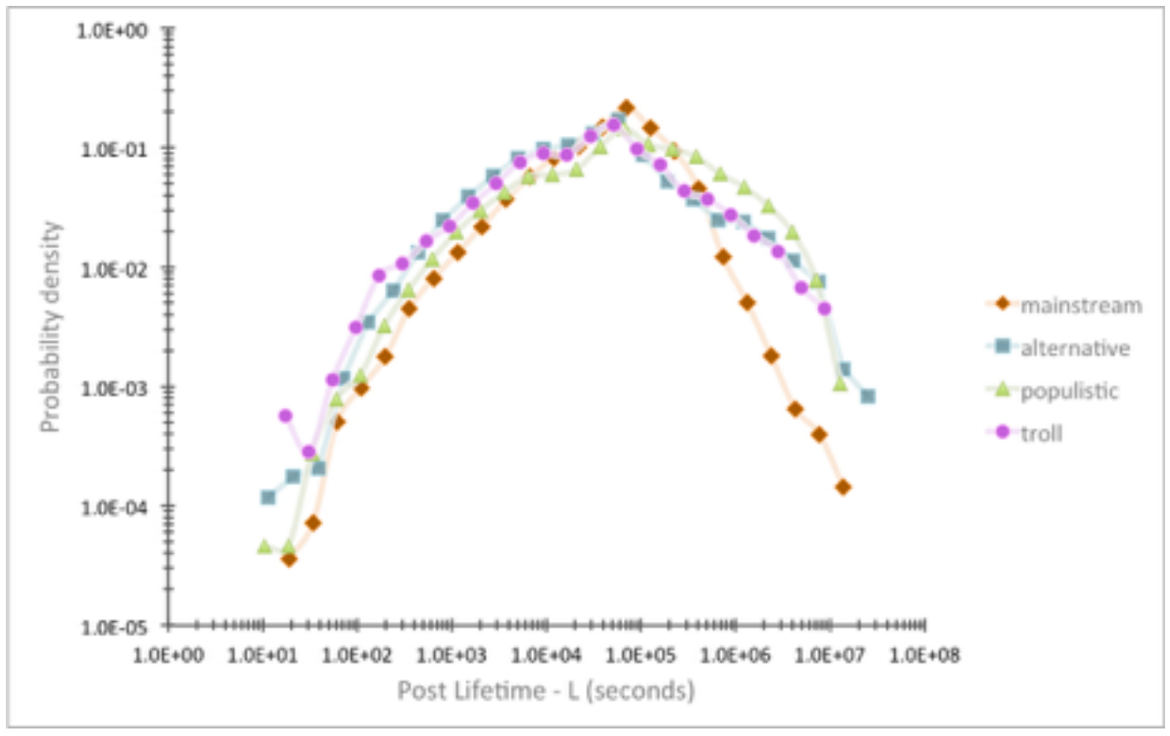
We study (over a period of 3 months) several Facebook communities centered around discussions (both directly and indirectly) related to the political scene

Sept 1st, 2012 - Feb 28th, 2013

	Total	Mainstream News	Alternative News	Political Activism
Distinct users	2,368,555	786,952	1,072,873	1,287,481
Pages	50	8	26	16
Posts	193,255	51,500	92,566	49,189
Likes	23,077,647	4,334,852	7,990,225	10,752,570
Comments	4,395,363	1,719,409	935,527	1,740,427
Likes of Comment	4,731,447	1,710,241	1,146,275	1,874,931

Table 1. Breakdown of Facebook dataset. **Mainstream News**: all the national newspapers present on Facebook. **Alternative News**: pages which disseminate controversial information, most often lacking supporting evidence and sometimes contradictory of the official news.. **Political Activism**: gathering users to publicly convey discontent against the current political and socio-economic situation.

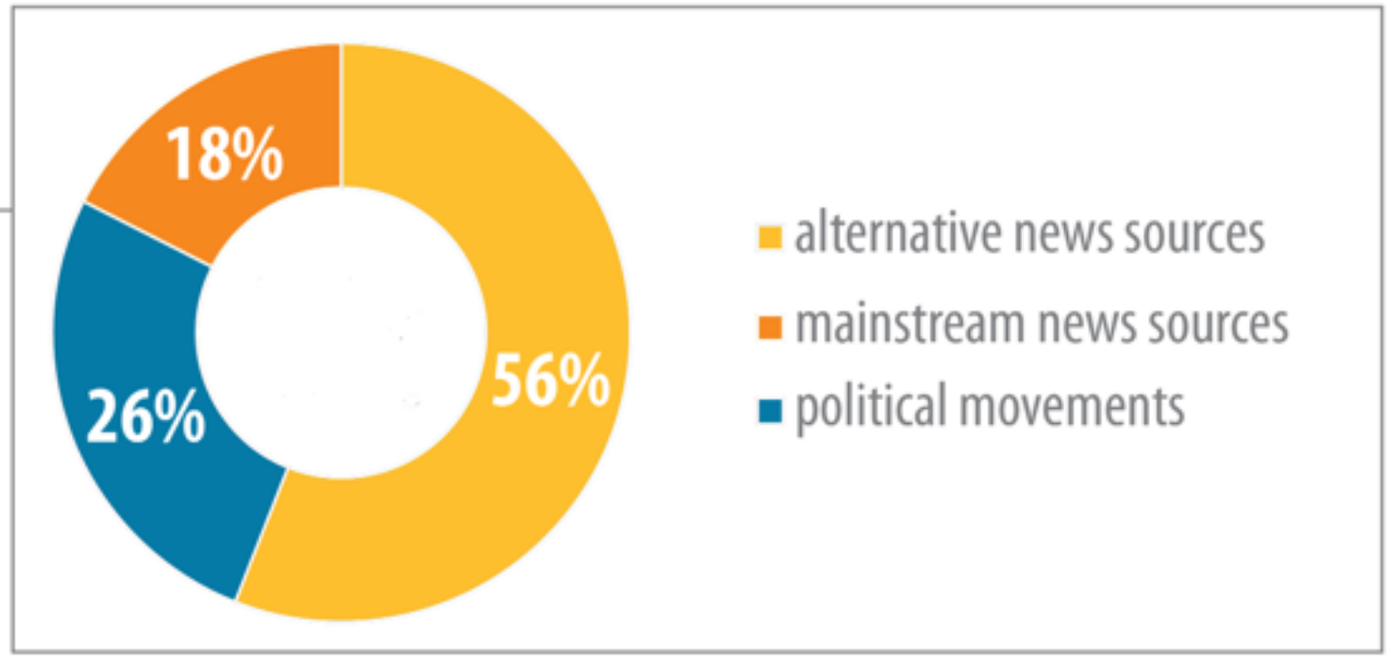




Response to Intentional False Information

2788 posts from troll page
7430 unique users
1279 users *labeled*

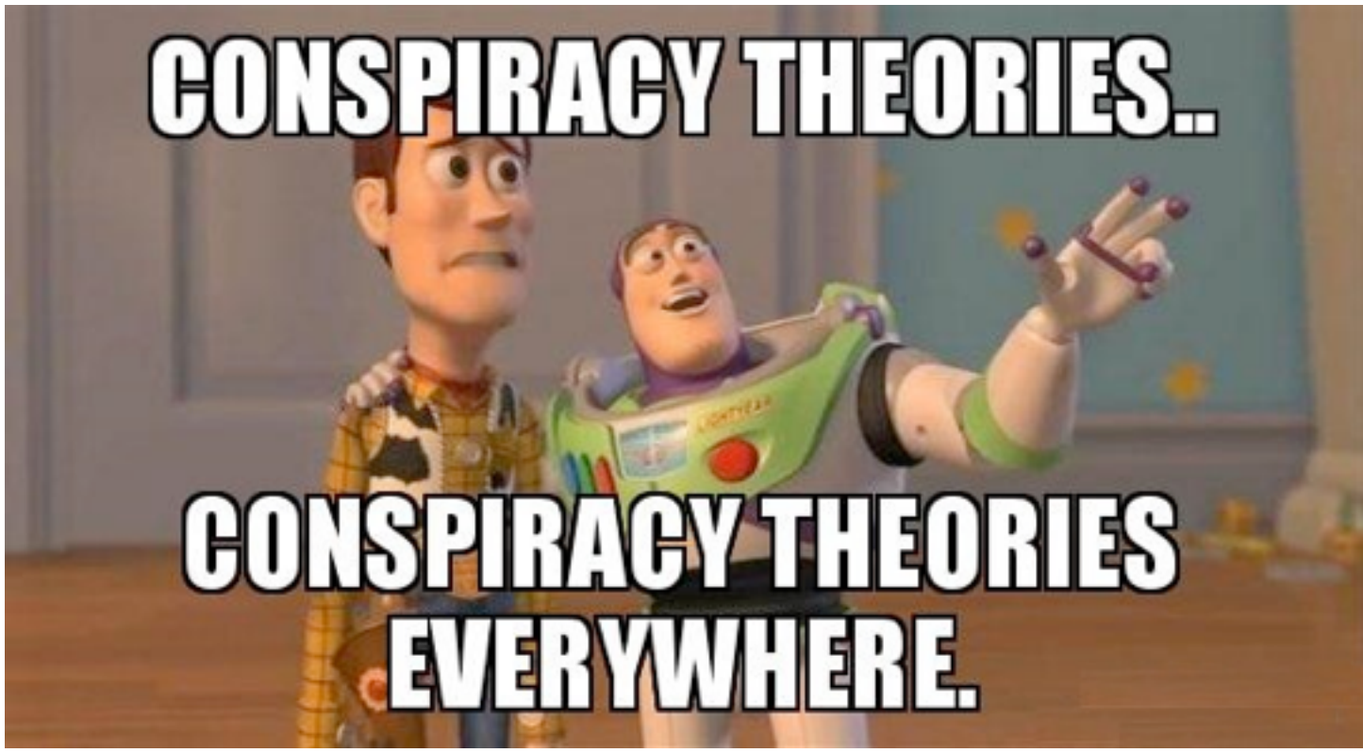
39098 actions
18212 likes
11337 comments
9549 likes of comment



PARADOX: USERS INTENTIONALLY AVOIDING THE “MAIN STREAM MEDIA MANIPULATION” ARE THE MORE PRONE TO BE MANIPULATED

THE DISINTERMEDIATION PROBLEM

THE WORLD ECONOMIC FORUM HAS POINTED OUT
MASSIVE DIGITAL MISINFORMATION
AS ONE OF THE MAIN RISKS FOR OUR SOCIETY

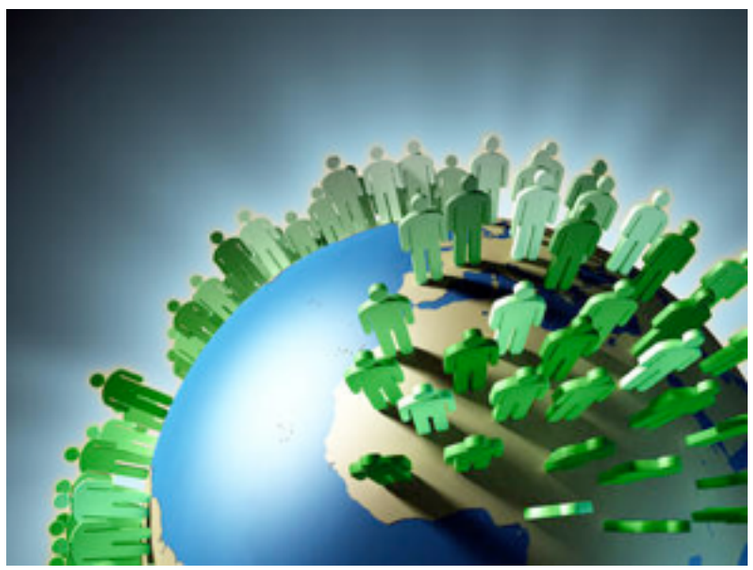


UNSUBSTANTIATED CLAIMS REVERBERATE AS LONG
AS MORE VERIFIED INFORMATION AND
IT IS REALLY EASY FOR THEM TO PASS
THE CREDULITY BARRIER

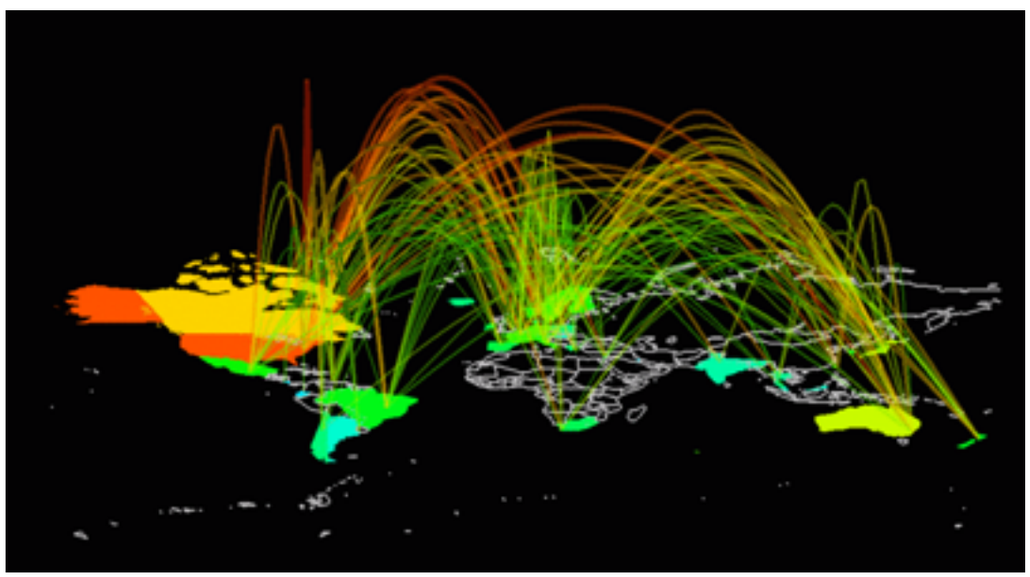
WHY?

COMPLEXITY OF THE WORLD AND SOCIETY IS INCREASING

GLOBALIZATION

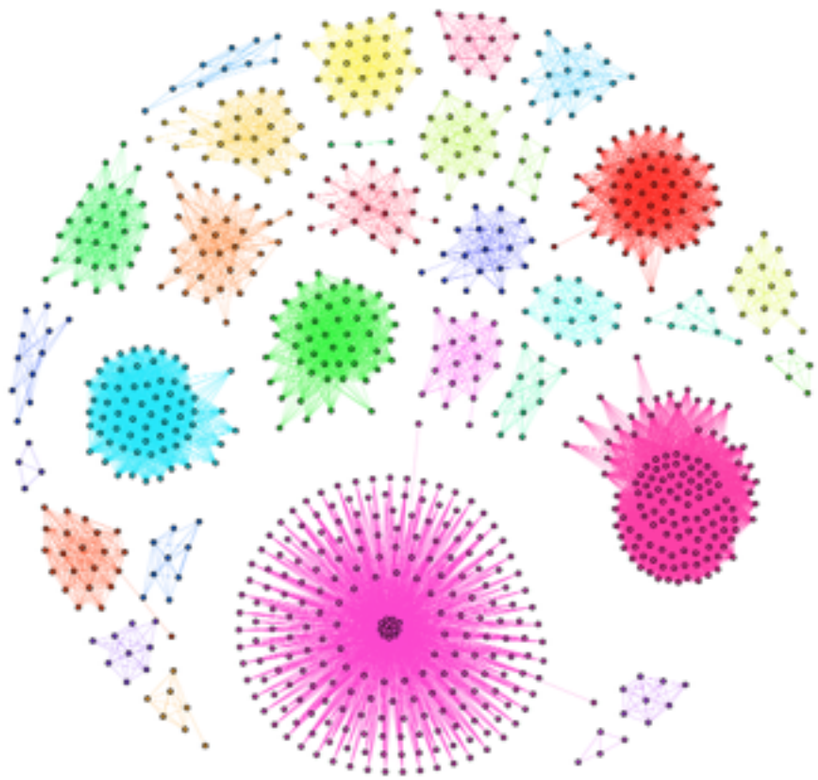


INTERCONNECTIVITY



TECHNOLOGICAL AND SCIENTIFIC PROGRESS

FREE CIRCULATION OF CONTENTS AND COMMUNITIES

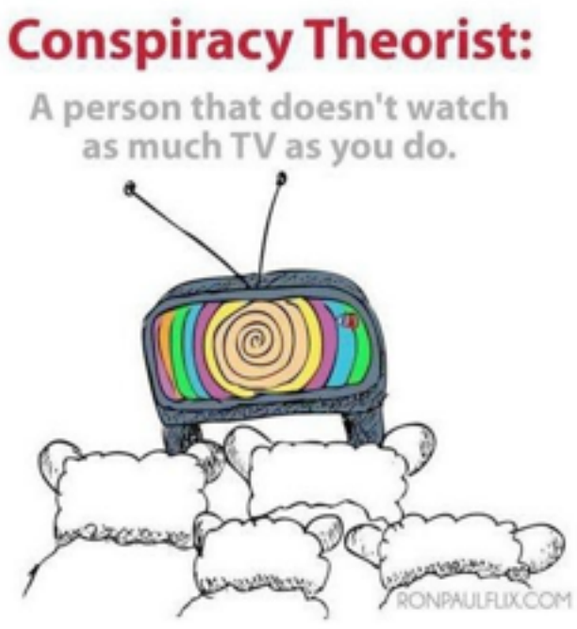


On the internet several and differentiate cultures coexist:

- On Facebook, they are mainly organized around **PAGES SPECIALIZED ON SPECIFIC TOPICS**, e.g. against or pro vaccination, vegan, politics, science, conspiracy, etc.

- On these pages users interact between them by commenting the posts and thus **SHARING, DEBATING** and **REINFORCING** their opinions.

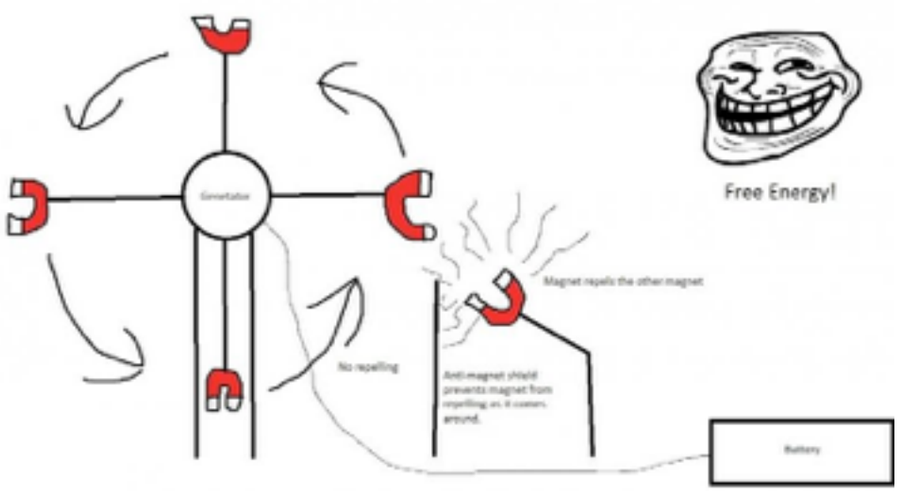
CONSPIRACY THINKING



THERE IS A SECRET PLOT AIMED AT MANIPULATING SOCIETY. SOME SECRET ORGANIZATION MOTIVATED BY ECONOMIC INTEREST IS:

- HIDING THE EXISTENCE OF SECRET TECHNOLOGIES
- MANIPULATING THE WEATHER
- TRYING TO UNIFY THE WORLD UNDER A UNIQUE GOVERNMENT
- MANIPULATING THE PUBLIC OPINION THROUGH MAIN STREAM MEDIA

- MYTHOLOGICAL NARRATIVES TO THE COMPLEXITY OF THE WORLD -

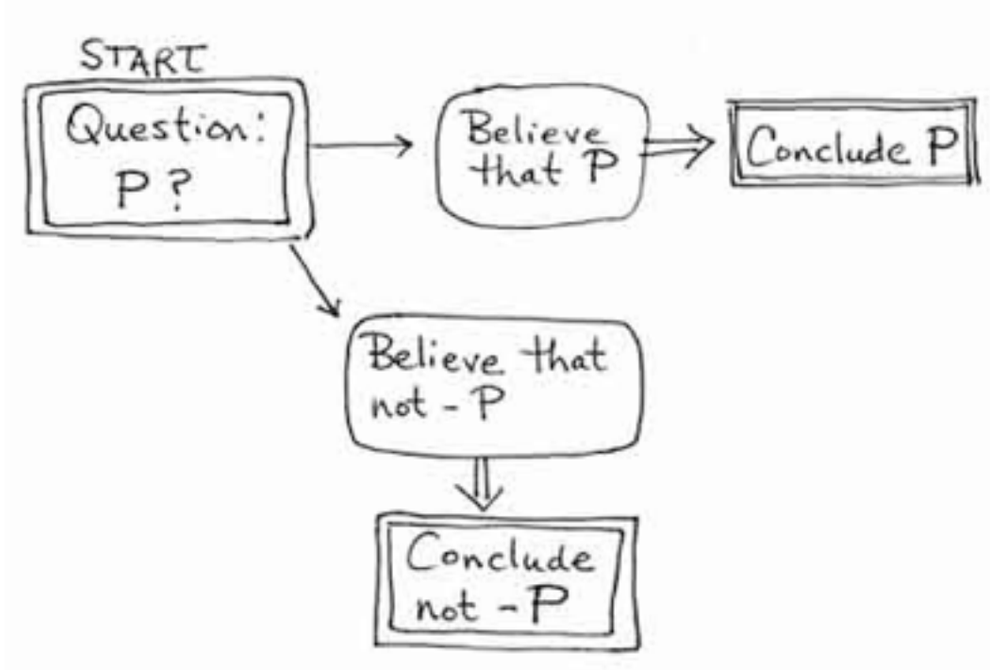


SCIENTIFIC THINKING

SCIENCE AND SCIENTIFIC THINKING COULD BE THE CORRECT NARRATIVE TO UNDERSTAND REALITY.

- TRYING DO DISENTANGLE COMPLEXITY
- SKEPTICISM
- LOOKING FOR EVIDENCE

- **DIFFICULT TO DISSEMINATE BECAUSE OF ITS COMPLEXITY**



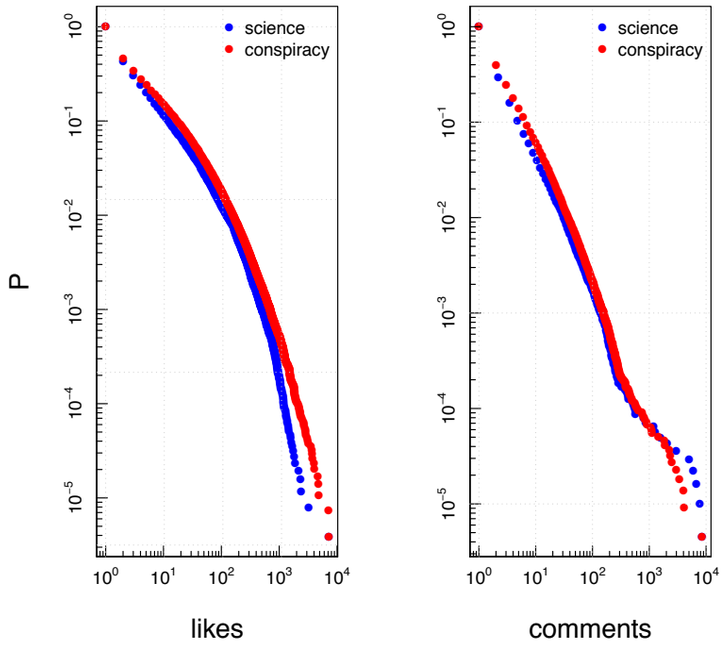
Science Vs Conspiracy: Collective Narratives in the Age of Misinformation

FACEBOOK PAGES IN THE ITALIAN CONTEXT FROM TWO CATEGORIES, **SCIENCE** AND **CONSPIRACY**, OVER A 4 YEARS TIMESPAN (2010 - 2014)

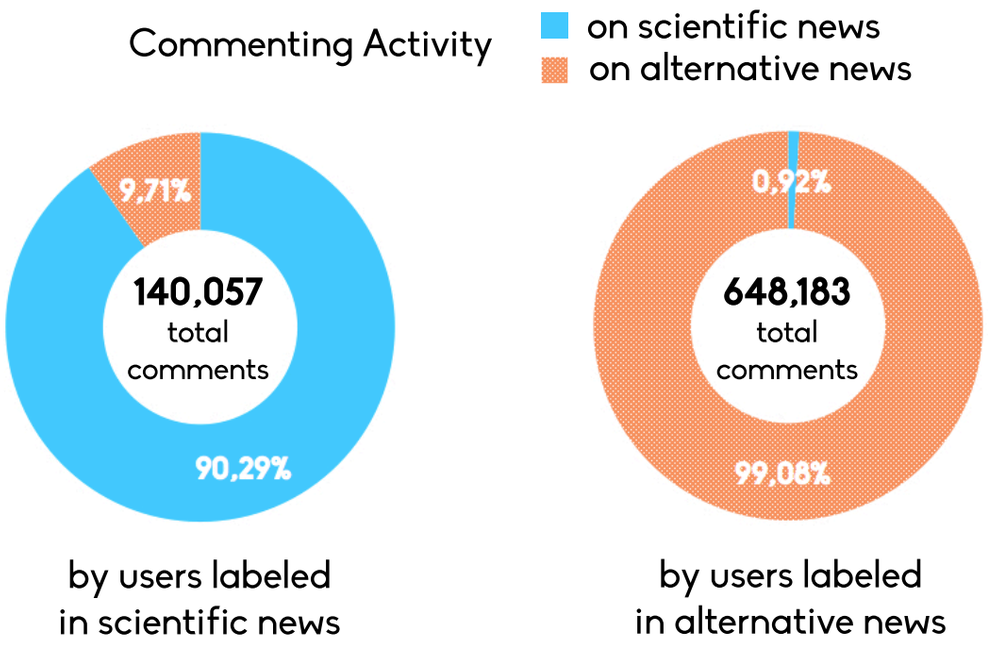
	Total	Science News	Conspiracy News
Pages	73	34	39
Posts	271,296	62,705	208,591
Likes	9,164,781	2,505,399	6,659,382
Comments	1,017,509	180,918	836,591
Likers	1,196,404	332,357	864,047
Commenters	279,972	53,438	226,534

Polarized Users activities in their community

	COMMENTS		LIKES	
SCIENCE	44,965	84,14%	255,225	76,79%
CONSPIRACY	218,27	96,35%	790,899	91,53%



Global consumption patterns for polarized users.
Here we continue our analysis on **95%-threshold** polarized users, i.e. users that liked at least 95% of posts in a specific category. The figure shows Cumulative Distribution Function for likes (panel a) and comments (panel b) of polarized users. Also the analysis on polarized users show some similarities in consumption patterns.



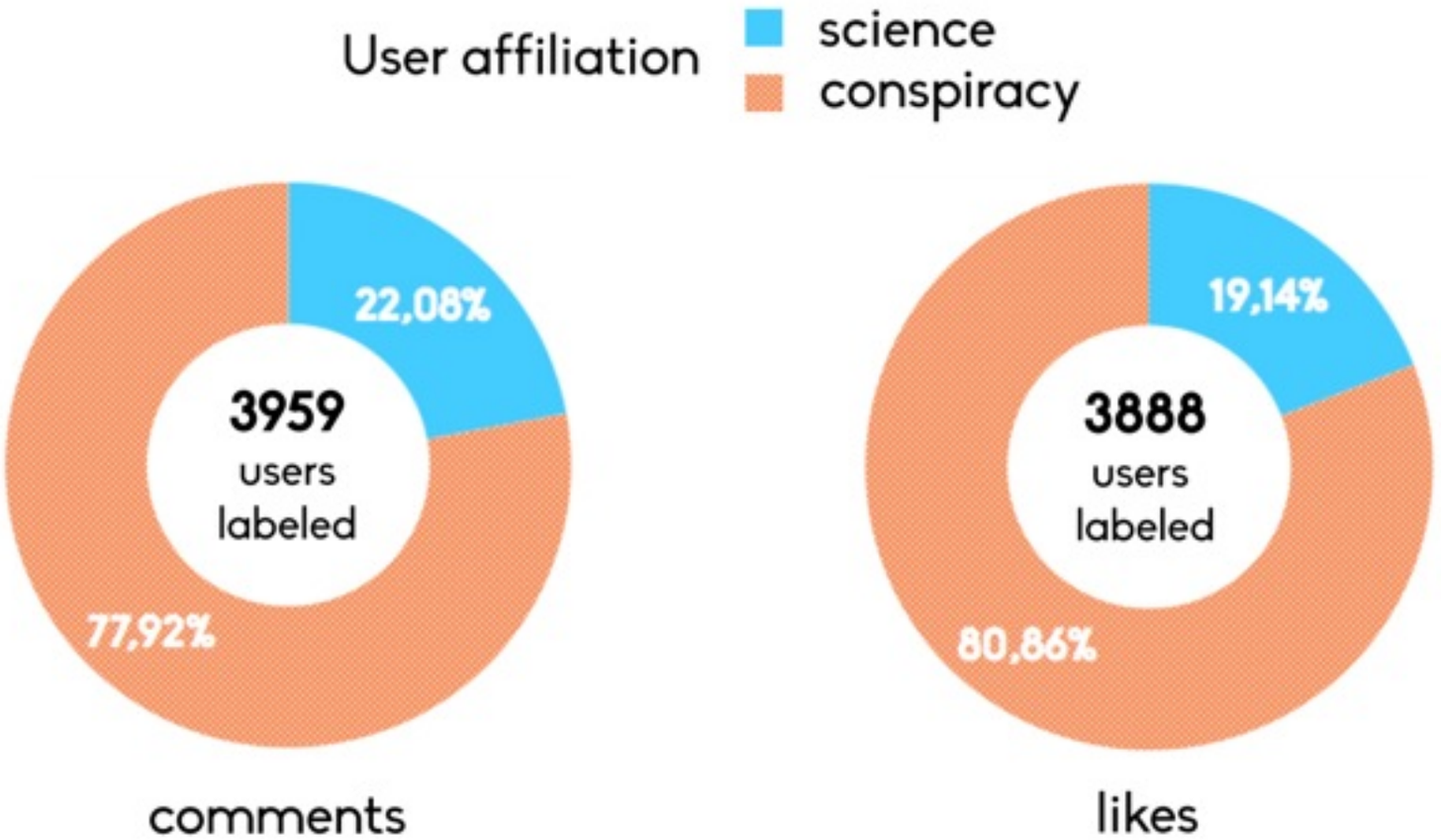
Commenting activity of polarized users.
Out of **7551** posts (**1991** in Science and **5790** in Conspiracy), we look at the activity of **95%-threshold** polarized users on the opposite category in terms of comments.

Users polarized on **Science** makes the **9.71%** of their total commenting activity on Conspiracy posts, whereas

users polarized in **Conspiracy** commented on Science posts only the **0.92%** of their total activity in terms of comments.

Hence, users polarized in Science pay attention to the activity on the opposite community and try to interact with users polarized in Conspiracy; conversely, users polarized in Conspiracy are focused on the inner dynamics of their community.

HOW DO THEY COMMENT ON **4000** INTENTIONAL FALSE CLAIMS?



Polarized users on false information.

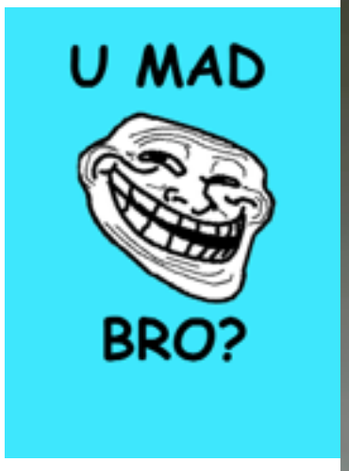
Percentage of likes and comments on intentional false memes posted by a satirical page from polarized users of the two categories.

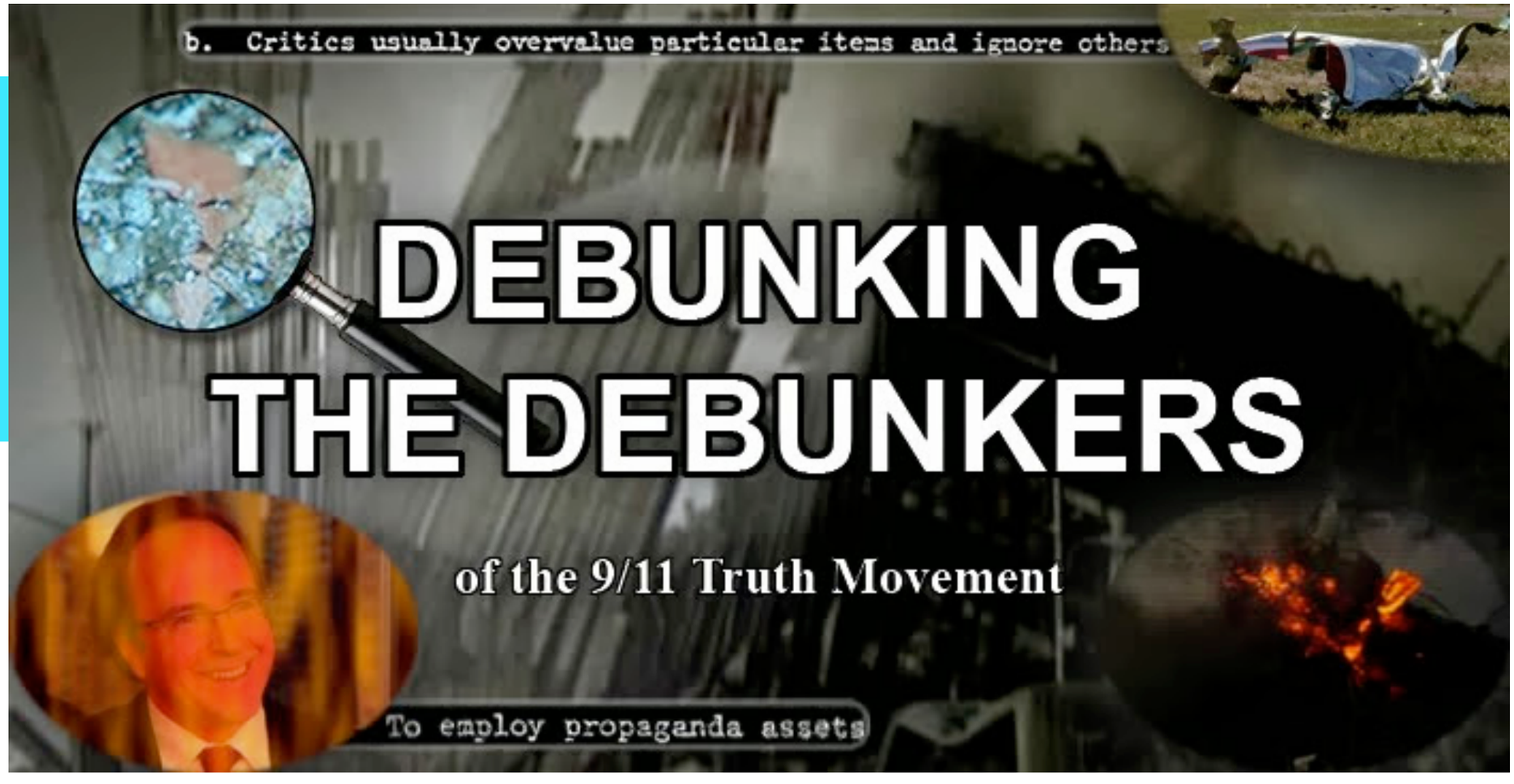
We notice the presence of debates among the two groups on the comments, while conspiracist confirm their attitude on liking.

RESISTANCE TO CORRECTIONS

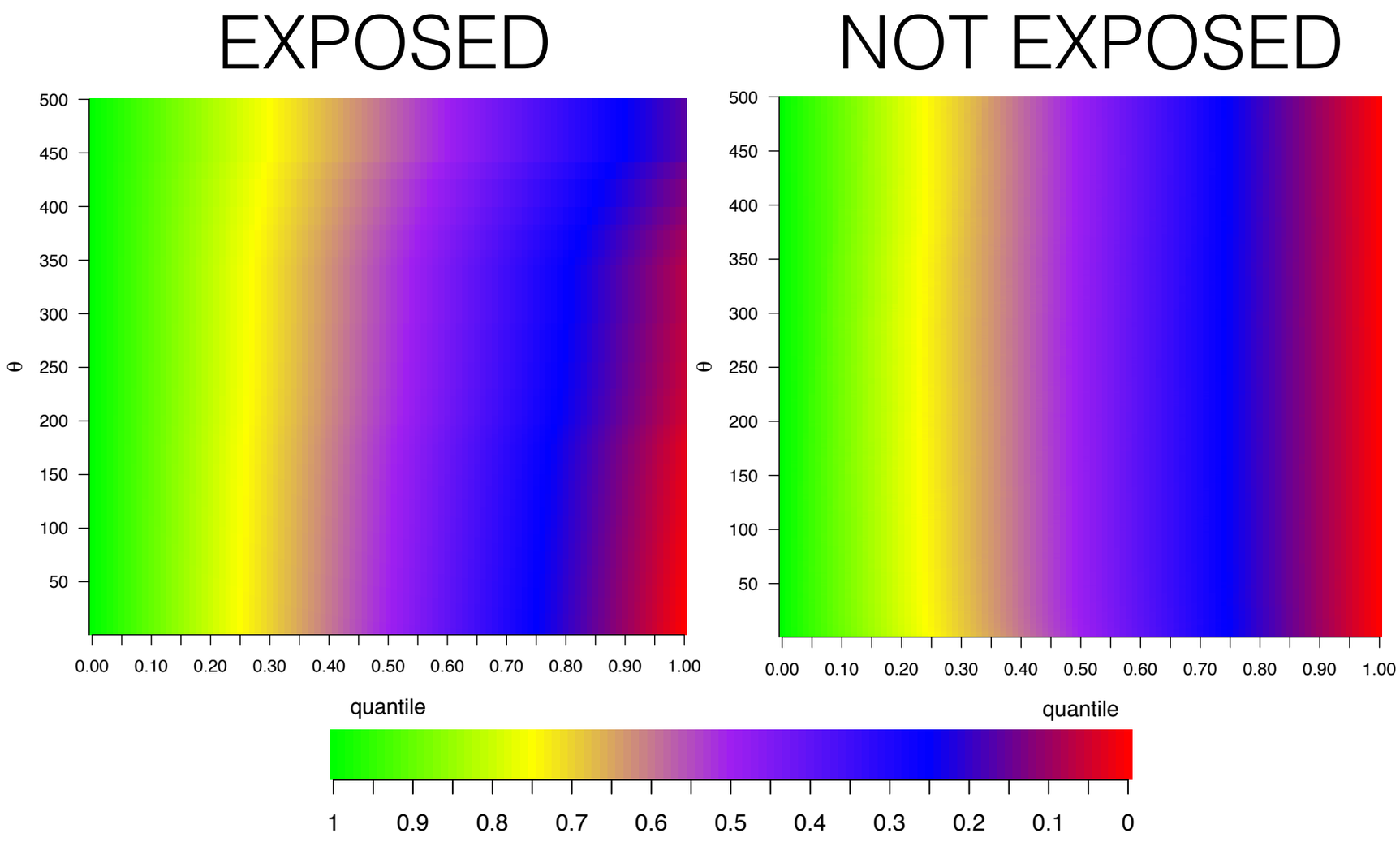
Debunkers: users or pages posting information aiming at correcting false claims circulating on the web.

Are Debunkers effective on persuading conspiracists?





Social Determinants of Content Selection in the age of misinformation

SURVIVAL PROBABILITY: the probability to interact with a post of a given category after a specified time T

Polarized users of **conspiracy-like** information and **debunking** posts