



### **Network Science**

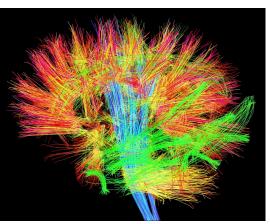
Guido Caldarelli

IMT Alti Studi LUCCA, Istituto Sistemi Complessi CNR ECLT, Venezia



### **Complexity Theory**









...the twenty-first century would be the "century of complexity". ....



#### Why should we care?

 $2.5 \ 10^{18}$  bytes per day = 2500 PB/day ~ 1 Million PB/year



#### IBM

IBM Offering Information

#### 10 Key Marketing Trends for 2017

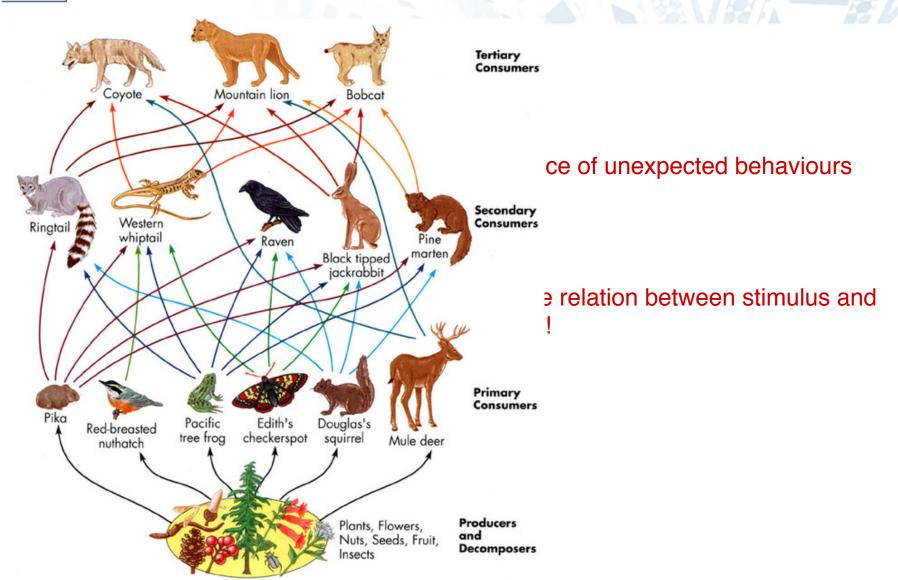
Every day we create 2.5 quintillion bytes of data. Find out which bytes matter most.

Learn more

Launch demo



#### **Complex Systems Properties**



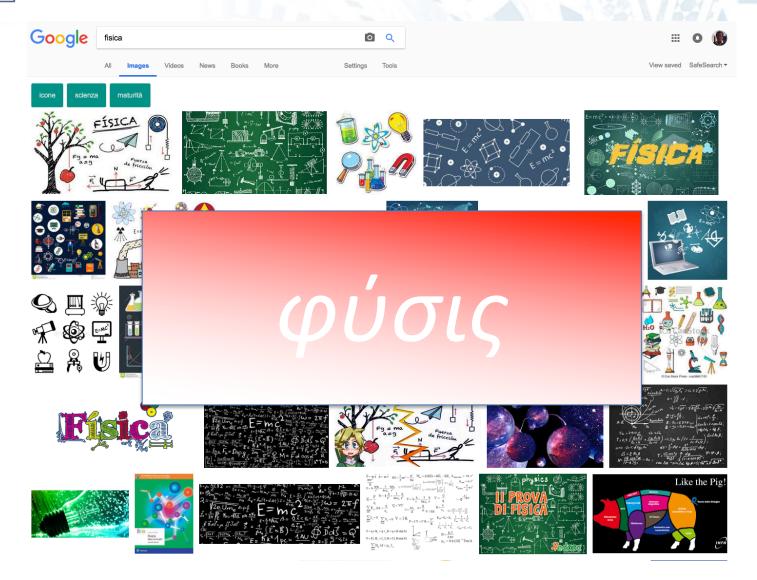






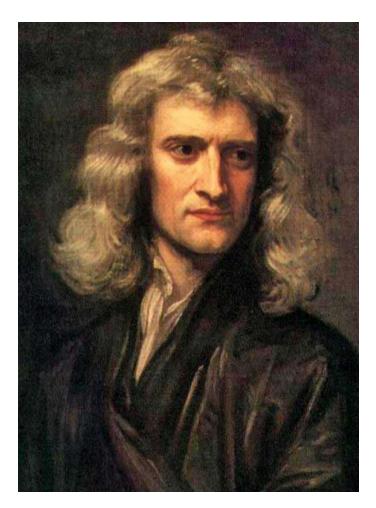


### **Complexity in Physics**











Physics (at least the aspect) is time invariant!!!!!



## Wembley July 13th 1985





#### Dr. Brian May (Queen)

Dr. Brian H. May (PhD Imperial College, 2007) Built himself his guitar (RED SPECIAL)

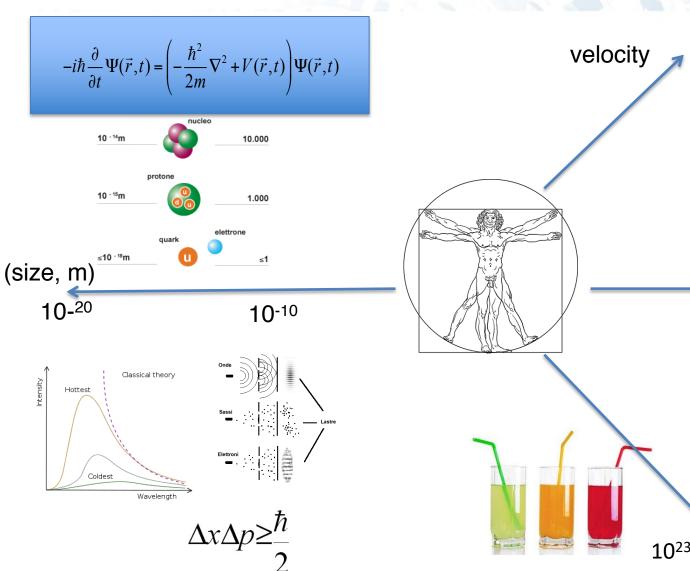


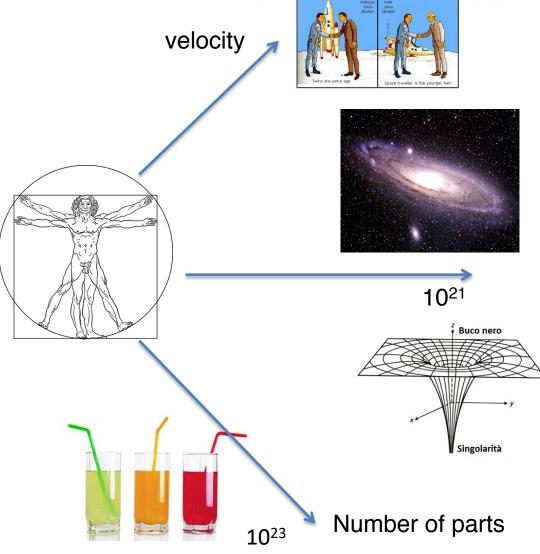
**Red Special** 

(un)fortunately very frequent approach....



#### Which Physics?



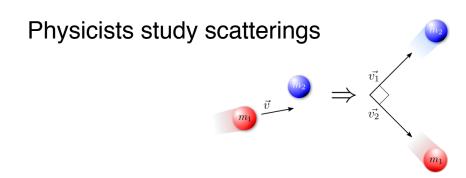




#### **Scatterings**



That's how we see ourselves...



To know what happens in a glass of water, I simply study the molecules' scatterings

Problem: We have 10<sup>23</sup>-10<sup>24</sup> of them

If every molecule were a rice grain, to match the number of molecules in a glass of water we should collect the harverst of 27 milions of years

(at the present rate of 739 milions of tons/year)



#### **Statistical Physics**



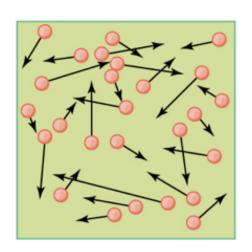
Ludwig Boltzmann

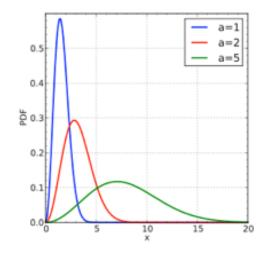


Joshua Willard Gibbs



James Clerk Maxwell





$$PV = \frac{1}{3}Nm\overline{v^2}$$

$$E = \frac{3}{2}k_BT$$

$$S = k_B \ln W$$

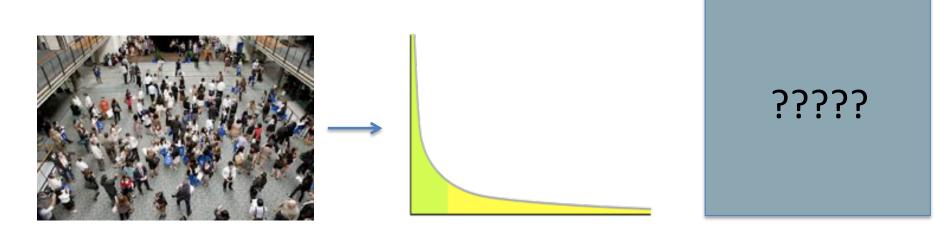


#### **Statistical Physics and Complexity**

What can we learn?

People are not described by 2 variables only, and react differently.

All we know, is the geometry of interaction





### **Internet shape**

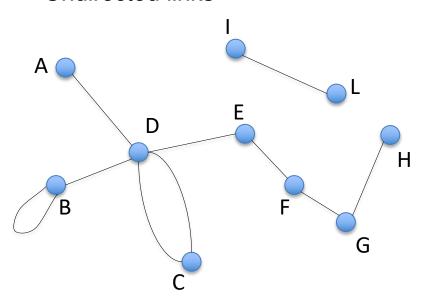




#### **Graph Theory**

#### undirected

#### **Undirected links**

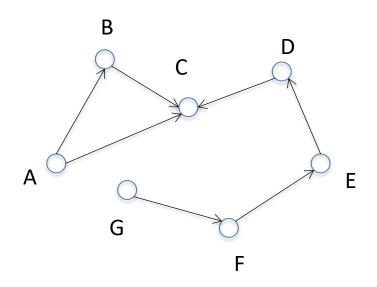


#### **Example**

Collaborations (movies, papers) **Protein Interactions** Internet

#### directed

#### **Directed links**



#### **Example**

**WWW** Phone calls Metabolic reactions



#### **Properties of complex networs**

Networks are game changer once we understand their properties

1 scale invariance that is highly heterogeneous

2 small world that is easy to travel

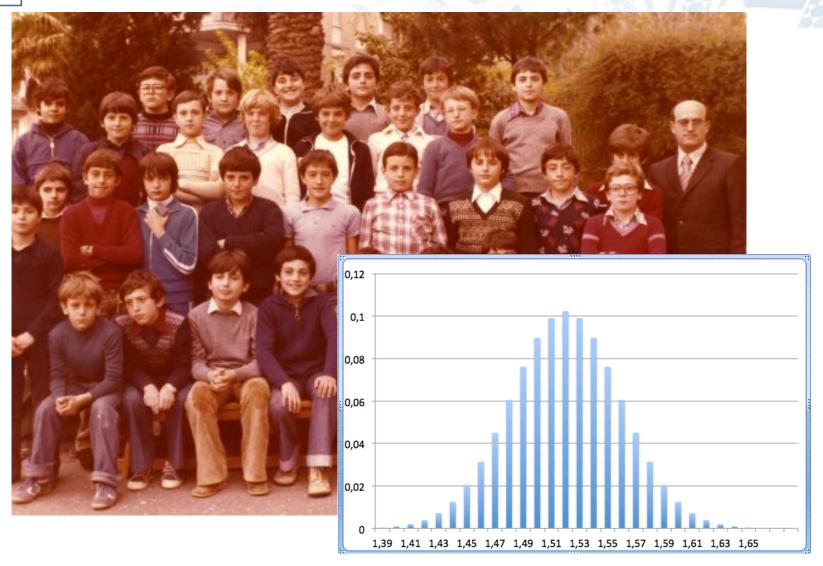
3 highly clustered that is community are present

4 centrality made that is some are more important than others

And much more



### 1) 12-years old students





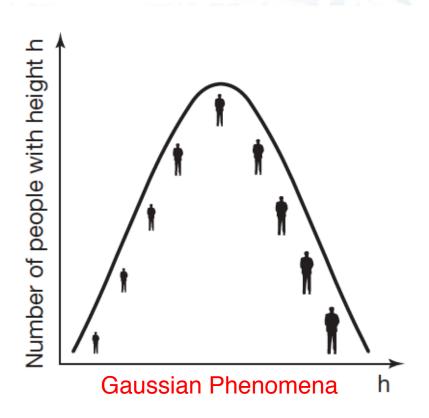


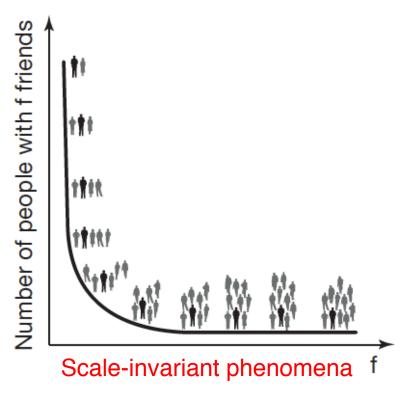
### 1) NYSE





#### 1) Scale invariance





Nobody is 3mm or 20Km tall

But somebody has 10000 more friends on Facebook than us!



#### 2) Small world

The beginning of a typical chain (#111) in the Nebraska Study.



STARTING PERSON

As a crude beginning, we thought it best to draw our starting persons from a distant city, so we chose Wichita, Kansas for our first study and Omaha, Nebraska for our second. (From Cambridge, these cities seem vaguely 'out there,' on the Great Plains or somewhere.) To obtain our sample, letters of solicitation were sent to residents in



1st REMOVE

Self-employed friend in Council Bluffs, lower these cities asking them to participate in a study of social contact in American society. The target person in our first study lived in Cambridge and was the wife of a divinity school student. In the second study carried out in

This tells the person who receives the folder exactly who sent it to him. The roster also has another practical effect; it prevents endless looping of the folder through participants who have already served as links in the chain, because each participant can see exactly what sequence of persons has led up to his own participation.

In addition to the document, the folder contains a stack of 15 business reply, or "tracer" cards. Each person receiving the folder takes out a card, fills it in, returns it to us, and sends the remaining cards along with the document to the next link.

Several other features of the procedure need to be emphasized. First, each



was the wife of a student living in Cambridge. Four days after the folders were sent to a group of starting persons in Kansas, an instructor at the Episcopal Theological Seminary approached our target person on the street. "Alice," he said, thrusting a brown folder toward her, "this is for you." At first she thought he was simply returning a folder

the median at five [see illustration setts? Part of the excitement of experiabove]. A median of five intermediate mental social psychology is that it is all persons is, in certain ways, impressive, so new we often have no way of knowing considering the distances traversed. Rewhether our techniques will work or cently, when I asked an intelligent friend simply turn out to be wispy pipe dreams.

No. of Completed Chains



Total no. of Chains, 44

No. of Intermediaries needed to reach Target Person

> require 100 intermediate persons or more to move from Nebraska to Sharon. Many people make somewhat similar estimates, and are surprised to learn that only five intermediaries will-on the average-suffice. Somehow it does not accord with intuition. Later, I shall try to explain the basis of the discrepancy

of mine how many steps he thought it



5th REMOVE

would take, he estimated that it would

giving us the illusion that the chains are shorter than they really are. There is a certain decay in the number of active chains over each remove, even when they do not drop out because they reach the target person. Of 160 chains that started in Nebraska, 44 were completed and 126 dropped out. These chains die before completion because on each remove a certain proportion of participants simply do not cooperate and fail to send on the folder. Thus, the results we obtained on the distribution of chain



6th REMOVE





7th REMOVE

Participants indicated on the reply cards whether they were sending the folder on to a friend, a relative, or an acquaintance. In the Kansas Study, 123 sent the folder to friends and acquaintances, while only 22 sent it to relatives. Cross-cultural comparison would seem useful here. It is quite likely that in societies which possess extended kinship systems, relatives will be more heavily represented in the communication network than is true in the United States. In American society, where extended kinship links are not maintained, ac-



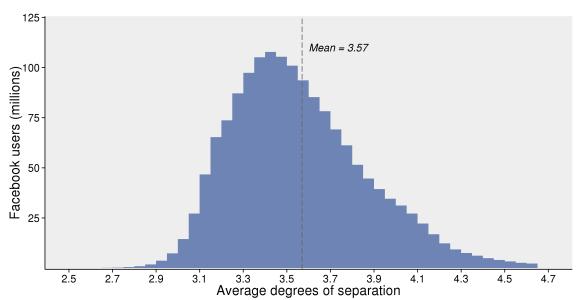


### 2) Small world

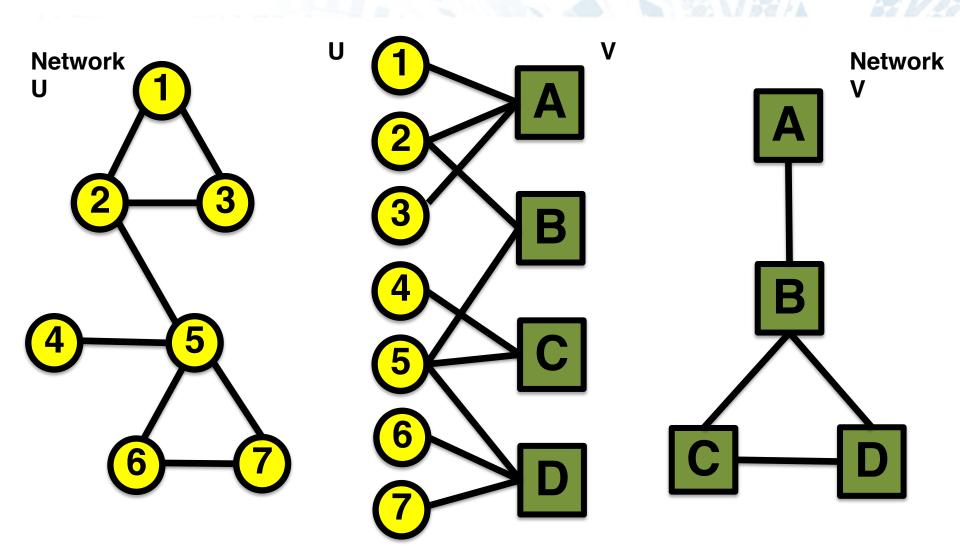
On average we are at six Degrees of separation



#### Much less actually



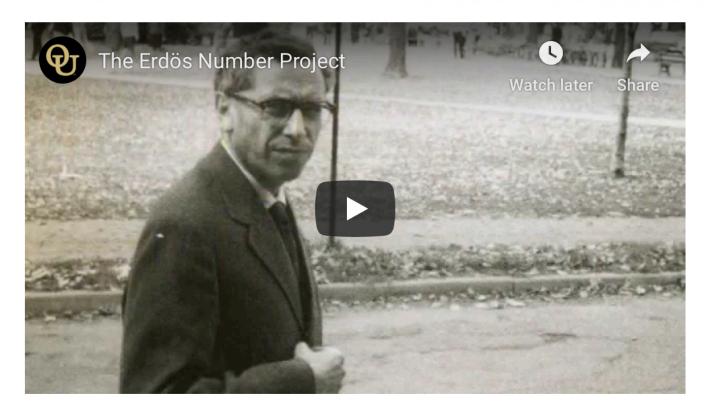








#### 3) Bipartite networks, Erdos number



Read Aug. 1, 2014 News at OU article on the popularity of this website.

#### **The Erdös Number Project**

This is the website for the Erdös Number Project, which studies research collaboration among mathematicians.

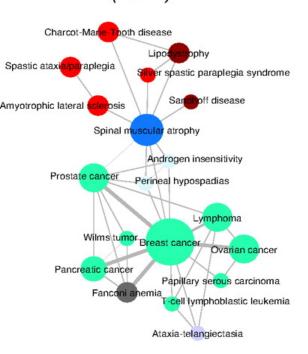


#### 3) Bipartite networks, Bacon number

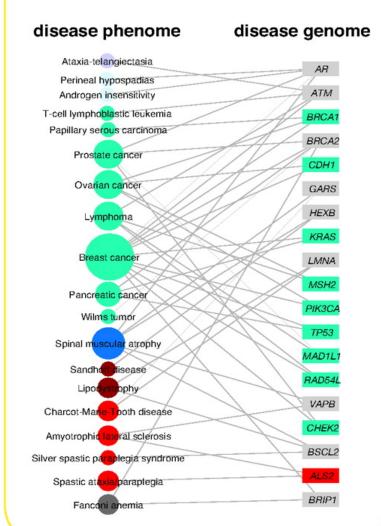




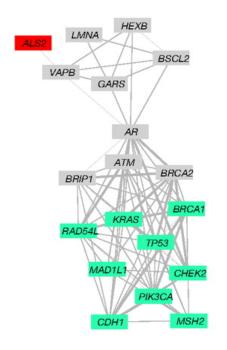
## Human Disease Network (HDN)



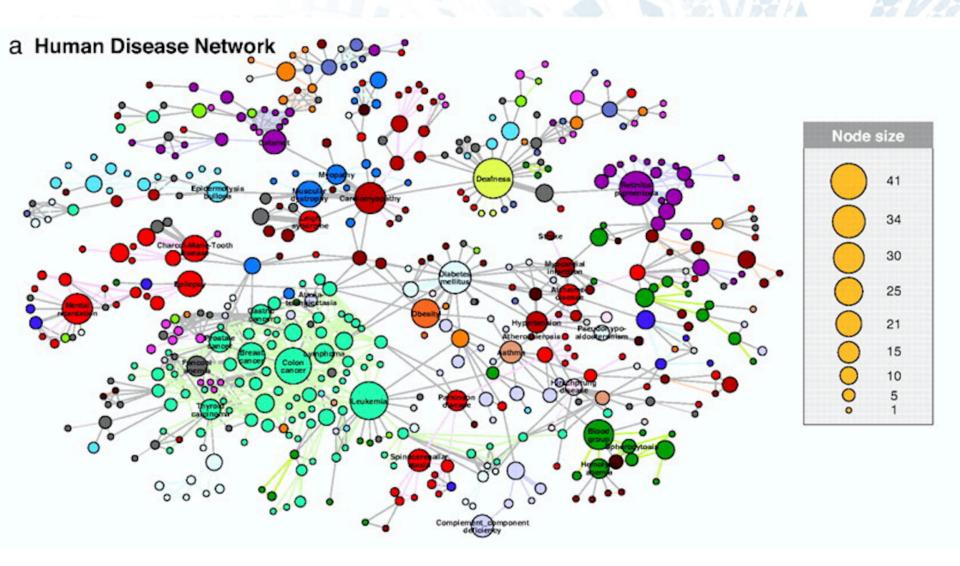
#### **DISEASOME**



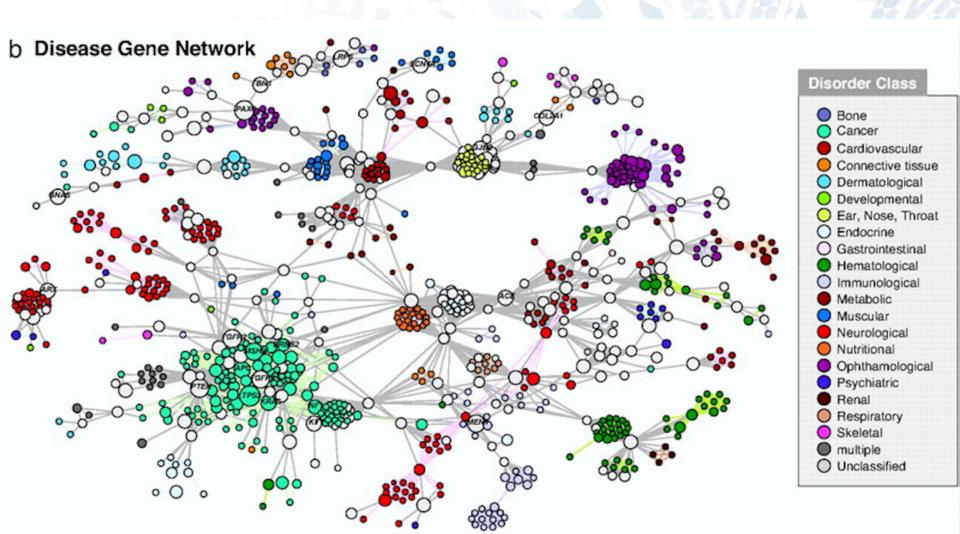
## Disease Gene Network (DGN)





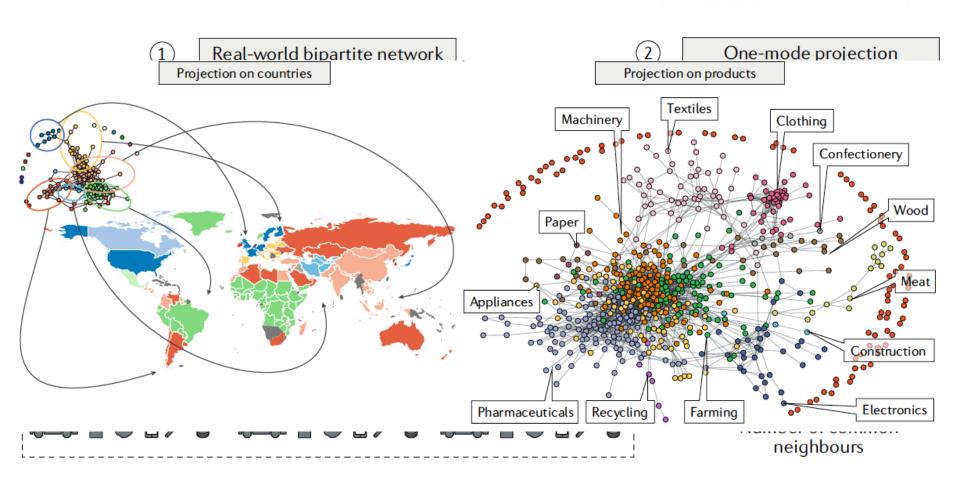








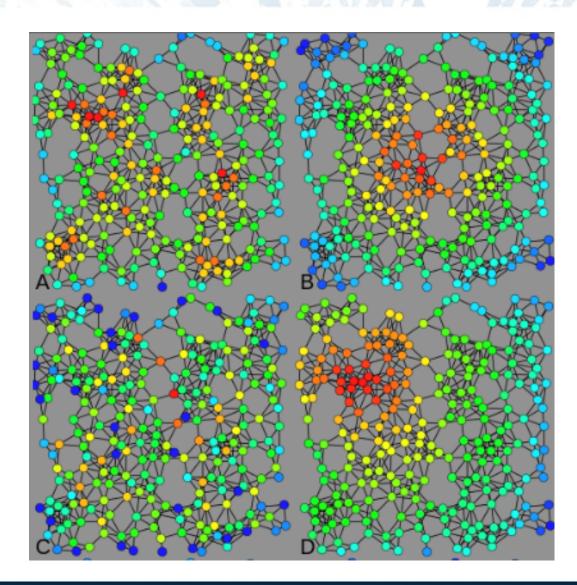
#### **World Trade Web**





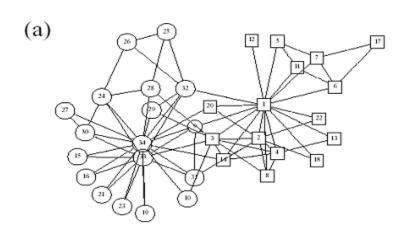
### 4) Centrality

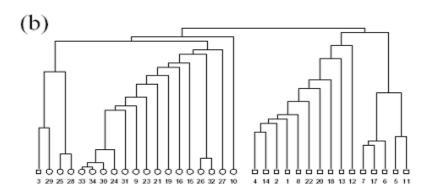
- **A** Degree Centrality
- **B** Distance Centrality
- **C** Betweenness Centrality
- **D** Eigenvector Centrality

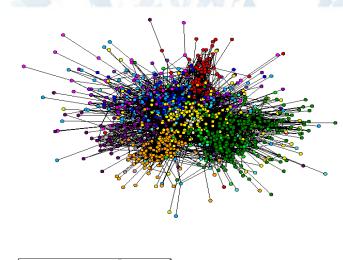


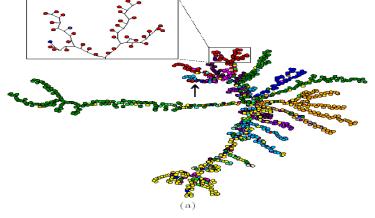


### 4) Clustering



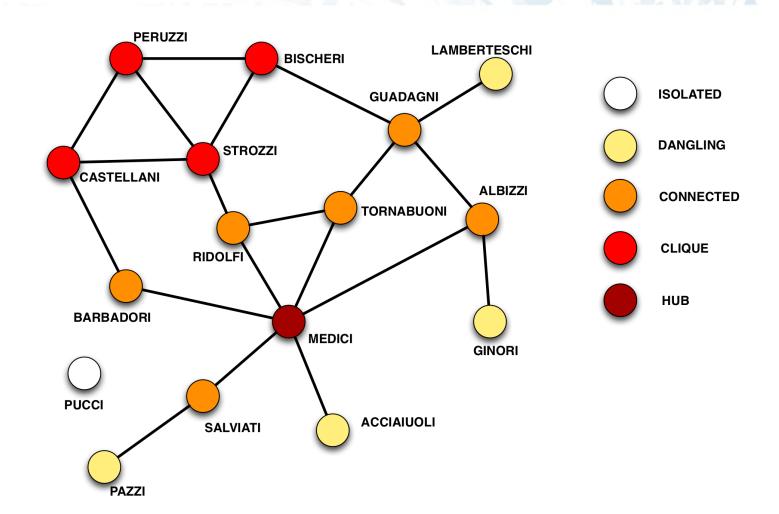








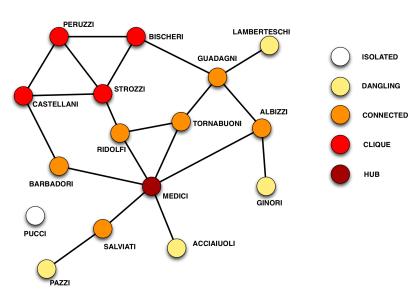
### 4) Centrality in politics



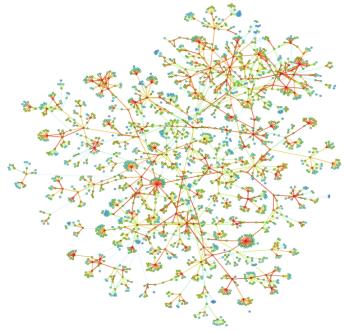


#### Our interactions are a network





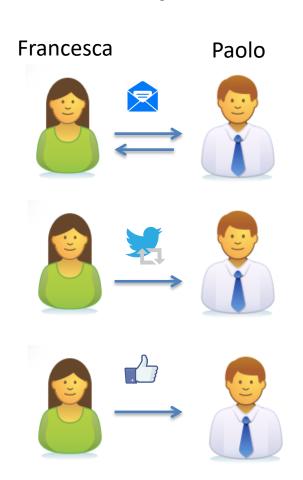






#### **Networks and user profiling**

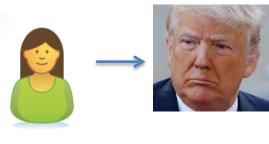
#### Friendship

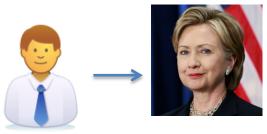


#### Goods



#### **Political Opinions**







### **Profiling is quite effective**





# With great powers comes great responsibility

(Spiderman)



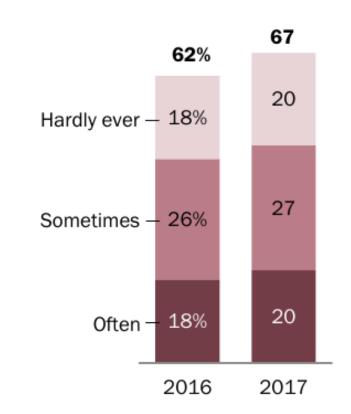




In 2016, roughly 50% of Americans aged 18-29 used online platforms as their primary source of news while 27% watched the news on television and 5% read print newspapers[Niklewicz 2017];

As of August 2017, two-thirds (67%) of US adults report that they get at least some of their news on social media[Shearer and Gottfried 2017];

% of U.S. adults who get news from social media sites ...



@GuidoCaldarelli



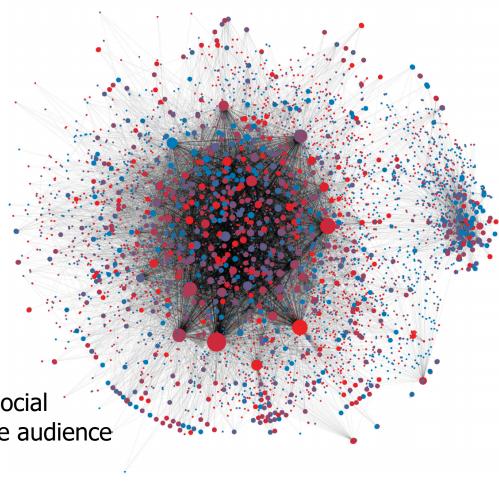


## **Twitter networks**

Retweet network for the #SB277 hashtag, about a California law on vaccination;

- The node size represents influence;
- The color represents bot scores:
  - red nodes: bot accounts;
  - blue nodes: humans.

Social bots can alter the perception of social media influence, artificially enlarging the audience of some topics[Ferrara et al. 2017]





## Latest results

# Data mining has revealed previously unknown Russian Twitter troll campaigns

Trolls left forensic fingerprints that cybersecurity experts used to find other disinformation campaigns both in the US and elsewhere.

by Emerging Technology from the arXiv October 11, 2018

arxiv.org/abs/1810.01466: Unsupervised Machine Learning of Open Source Russian Twitter Data Reveals Global Scope and Operational Characteristics



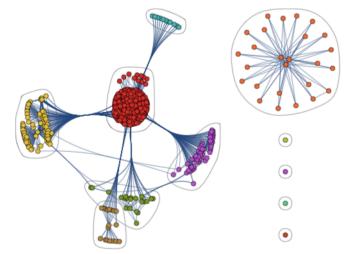
my2017resoult
sometimesitsokto
thingsortaughtalschool
makemehatevouinonephra
my2017biggesthop
my2017byggesthop
my2017byggesthop

General statement (glatiante general place) in the control of the

redat marks he most hold gates followed and gates of the second ga





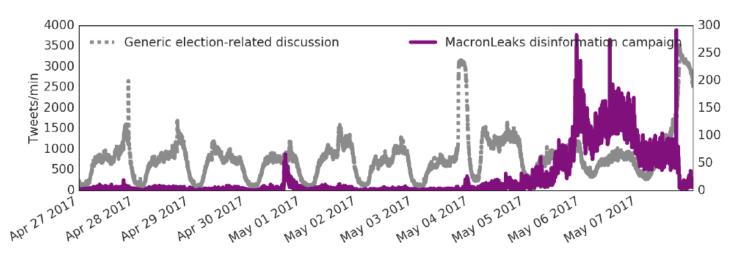




## Manipulation of public opinion:

- Societal debate interference: misinformation may strongly influence individuals' beliefs on several topics, (e.g effectiveness of vaccination[Schmidt et al. 2018]);
- **Disinformation** in political campaign: massive diffusion of fake news on social media during the 2017 French presidential elections [Ferrara 2017] and the 2016 US presidential elections [Shao et al. 2018].

Timeline of the tweets volume generated every minute during April 27, 2017 through May 7, 2017 (French presidential election date).



Purple line: tweets on MacronLeaks;

Gray line: tweets on generic election-related discussion

@GuidoCaldarelli

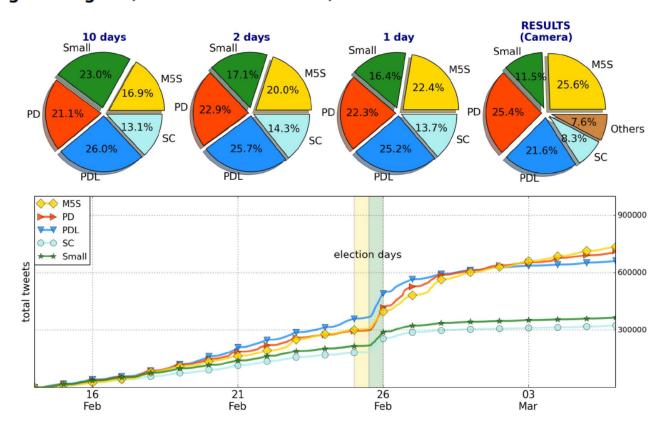




## **Study of Twitter votes**

## A Multi-Level Geographical Study of Italian Political **Elections from Twitter Data**

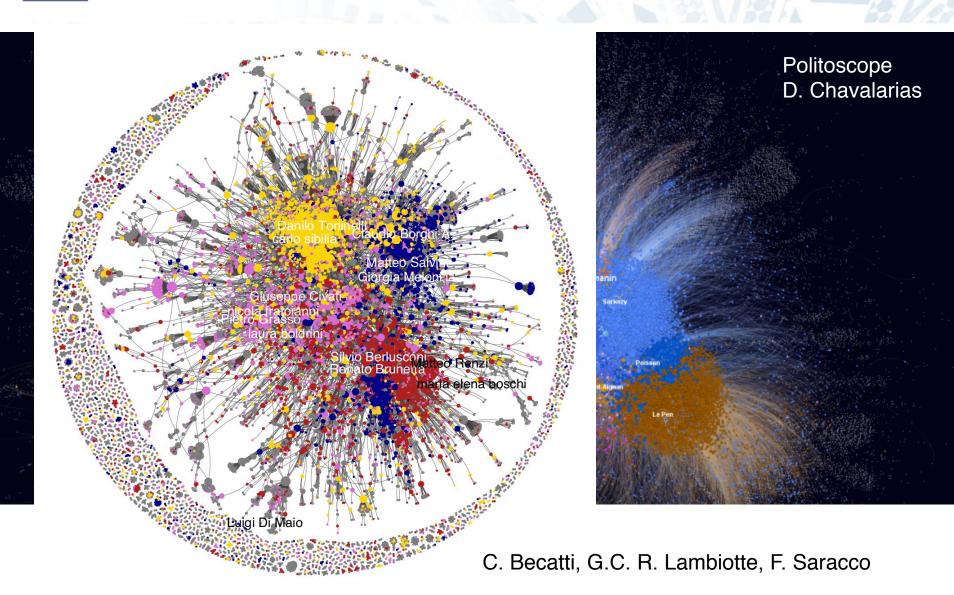
Guido Caldarelli<sup>1,2,3,4</sup>, Alessandro Chessa<sup>1,4</sup>, Fabio Pammolli<sup>1,5</sup>, Gabriele Pompa<sup>1</sup>, Michelangelo Puliga<sup>1,4</sup>\*, Massimo Riccaboni<sup>1,6</sup>, Gianni Riotta<sup>1,7</sup>







## **Centrality in political debate (France-Italy)**





# **Applications**

## Other than Economics and Medicine:

**Intelligence:** we can track down the network of terrorists

**Finance:** for systemic risk in the interbank market

**Brain:** for diagnosis and analysis of diseases



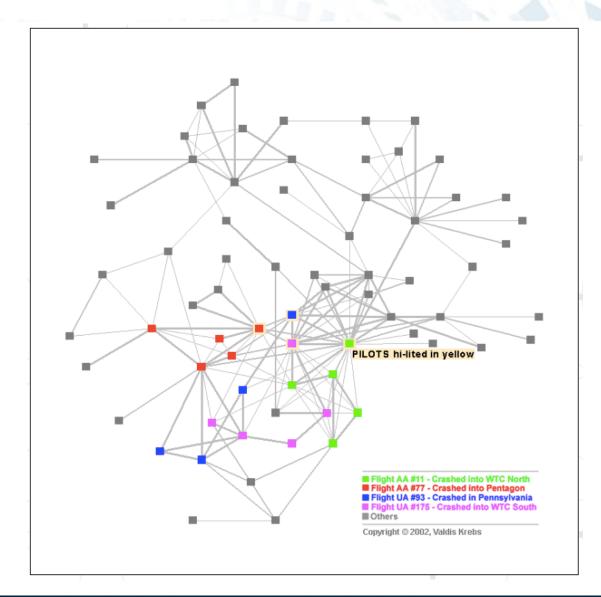
## **Intelligence and Networks**

- Incompleteness the inevitability of missing nodes and links that the investigators will not uncover.
- 2. Fuzzy boundaries the difficulty in deciding who to include and who not to include.
- 3. Dynamic these networks are not static, they are always changing. Instead of looking at the presence or absence of a tie between two individuals, Sparrow suggests looking at the waxing and waning strength of a tie depending upon the time and the task at hand.

Twin Towers atack 11 September 2001



## **Network of attackers**





## **Attackers**

#### THE HIJACKERS ...

#### American Airlines 11

Crashed into WTC (north)



Mohamed Atta (Egyptian) Received pilot training



Waleed M. Alshehri (Saudi) Commercial pilot



Wail Alshahri (Saudi) Possible pilot training



Satam al-Suqami (Nationality unknown)

No picture available Abdulaziz Alomari\* (Saudi) Possible pilot training

#### American Airlines 77

Crashed into Pentagon



Khalid al-Midhar (Nationality unknown) Received pilot training



Majed Moqed (Nationality unknown)



Salem Alhamzi\* (Saudi) Possible pilot training



Nawaf Alhamzi\* (Saudi)



Hani Hanjour (Saudi)

#### UnitedAirlines 175

Crashed into WTC (south)



Marwan al-Shehhi (United Arab Emirates) Received pilot training

No picture available Fayez Ahmed (Believed to be Saudi)



Ahmed Alghamdi (Possibly Saudi)



Hamza Alghamdi (Believed to be Saudi) Possible pilot training



Mohald Alshehri (Nationality unknown) Possible pilot training

#### United Airlines 93

Crashed in Pennsylvania



Ziad Jarrah (Lebanese) Received pilot training



Ahmed Alhaznawi (Saudi)



Ahmed Alnami (Nationality unknown)



Saeed Alghamdi\* (Seems to be Saudi)

\*Disputed identity

#### AND HOW THEY WERE CONNECTED

#### Attended same technical college

Hamburg, Germany Mohamed Atta Marwan al-Shehhi Ziad Jarrah

#### Took flight classes together

Pilot schools in Florida Mohamed Atta

Mohamed Atta Marwan al-Shehhi

Pilot schools In San Diago

Khalid al-Midhar Nawaf Alhamzi

#### Bought flight Boug tickets using ticket same address

Mohamed Atta\*
 Marwan al-Shehhi
 Abdulaziz Alomari\*

\* Also used same credit card

- Waleed M. Alshehri
   Wail Alshahri
- Fayez Ahmed Mohald Alshehri

Ahmed Alghamdi
 Hamza Alghamdi

#### Known to be together in week before attacks

Stayed together in a Florida motel

Mohamed Atta Marwan al-Shehhi

Attended a gym in Maryland (Sept 2-6), also seen dining together

Khalid al-Midhar Majed Moqed Salem Alhamzi Nawaf Alhamzi Hani Hanjour

#### Bought flight tickets together

Mohamed Atta Ziad Jarrah Ahmed Alhaznawi

Picked up tickets bought earlier in Baltimore

Khalid al-Midhar Majed Moqed

Bought from the same travel agent in Florida

Ahmed Alnami Saeed Alghamdi

#### Last known address

Hollywood, Florida Marwan al-Shehhi Waleed M. Alshehri Wail Alshahri Ziad Jarrah Hani Hanjour

Other cities in Florida Mohamed Atta Fayez Ahmed

Fayez Ahmed Ahmed Alghamdi Mohald Alshehri Khalid al-Midhar Ahmed Alhaznawi Ahmed Alnami Saeed Alghamdi

#### Outside Florida

Satam al-Suqami Hamza Alghamdi Abdulaziz Alomari Majed Moqed Salem Alhamzi Nawaf Alhamzi







## **Attackers**

Degrees		Betweenness		Closeness	
Mohamed Atta	0.334	Nawaf Alhazmi	0.571	Mohamed Atta	
Marwan Al-Shehhi	0.318	Mohamed Atta	0.537	Nawaf Alhazmi	
Hani Hanjour	0.227	Hani Hanjour	0.507	Hani Hanjour	
Nawaf Alhazmi	0.158	Marwan Al-Shehhi	0.500	Marwan Al-Shehhi	
Ziad Jarrah	0.116	Saeed Alghamdi*	0.480	Ziad Jarrah	
Ramzi Bin al-Shibh	0.081	Hamza Alghamdi	0.429	Mustafa al-Hisawi	
Said Bahaji	0.080	Waleed Alshehri	0.429	Salem Alhazmi*	
Hamza Alghamdi	0.076	Ziad Jarrah	0.424	Lotfi Raissi	
Saeed Alghamdi*	0.064	Mustafa al-Hisawi	0.424	Saeed Alghamdi*	
Lotfi Raissi	0.049	Abdul Aziz Al-Omari*	0.419	Abdul Aziz Al-Omari*	
Zakariya Essabar	0.033	Satam Suqami	0.414	Hamza Alghamdi	
Agus Budiman	0.031	Fayez Ahmed	0.414	Ramzi Bin al-Shibh	
Khalid Al-Mihdhar	0.030	Ahmed Al Haznawi	0.409	Said Bahaji	
Mounir El Motassadeq	0.026	Nabil al-Marabh	0.404	Ahmed Al Haznawi	
Mustafa al-Hisawi	0.016	Raed Hijazi	0.400	Zakariya Essabar	
Nabil al-Marabh	0.015	Lotfi Raissi	0.396	Agus Budiman	
Rayed Abdullah	0.012	Mohand Alshehri*	0.396	Khalid Al-Mihdhar	
Satam Suqami	0.011	Khalid Al-Mihdhar	0.391	Ahmed Alnami	
Waleed Alshehri	0.010	Ramzi Bin al-Shibh	0.391	Mounir El Motassadeq	
Abdul Aziz Al-Omari*	0.007	Salem Alhazmi*	0.387	Fayez Ahmed	
Abdussattar Shaikh	0.004	Ahmed Alghamdi	0.387	Mamoun Darkazanli	
Ahmed Al Haznawi	0.004	Said Bahaji	0.371	Zacarias Moussaoui	
Ahmed Alnami	0.002	Rayed Abdullah	0.367	Ahmed Khalil Al-Ani	
Fayez Ahmed	0.000	Abdussattar Shaikh	0.360	Abdussattar Shaikh	
	Mohamed Atta Marwan Al-Shehhi Hani Hanjour Nawaf Alhazmi Ziad Jarrah Ramzi Bin al-Shibh Said Bahaji Hamza Alghamdi Saeed Alghamdi* Lotfi Raissi Zakariya Essabar Agus Budiman Khalid Al-Mihdhar Mounir El Motassadeq Mustafa al-Hisawi Nabil al-Marabh Rayed Abdullah Satam Suqami Waleed Alshehri Abdul Aziz Al-Omari* Abdussattar Shaikh Ahmed Al Haznawi Ahmed Alnami	Mohamed Atta Marwan Al-Shehhi D.318 Hani Hanjour Nawaf Alhazmi D.158 Ziad Jarrah Ramzi Bin al-Shibh Said Bahaji D.080 Hamza Alghamdi Saeed Alghamdi* D.076 Saeed Alghamdi* D.049 Zakariya Essabar Agus Budiman Khalid Al-Mihdhar Mounir El Motassadeq Mustafa al-Hisawi Nabil al-Marabh Rayed Abdullah Satam Suqami Waleed Alshehri Abdul Aziz Al-Omari* Abdussattar Shaikh Ahmed Al Haznawi Ahmed Al Haznawi Alsa 1. 0.002	Mohamed Atta Marwan Al-Shehhi Hani Hanjour Nawaf Alhazmi Ziad Jarrah Ramzi Bin al-Shibh Said Bahaji Hongar Saeed Alghamdi Saee	Mohamed Atta         0.334         Nawaf Alhazmi         0.571           Marwan Al-Shehhi         0.318         Mohamed Atta         0.537           Hani Hanjour         0.227         Hani Hanjour         0.507           Nawaf Alhazmi         0.158         Marwan Al-Shehhi         0.500           Ziad Jarrah         0.116         Saeed Alghamdi*         0.480           Ramzi Bin al-Shibh         0.081         Hamza Alghamdi         0.429           Said Bahaji         0.080         Waleed Alshehri         0.429           Hamza Alghamdi         0.076         Ziad Jarrah         0.429           Hamza Alghamdi         0.076         Ziad Jarrah         0.429           Hamza Alghamdi*         0.064         Mustafa al-Hisawi         0.424           Lotfi Raissi         0.049         Abdul Aziz Al-Omari*         0.419           Zakariya Essabar         0.033         Satam Suqami         0.414           Khalid Al-Mihdhar         0.031         Fayez Ahmed         0.414           Khalid Al-Mihdhar         0.030         Ahmed Al Haznawi         0.409           Mustafa al-Hisawi         0.016         Raed Hijazi         0.400           Nabil al-Marabh         0.400         Nabil al-Marabh <t< td=""></t<>	



## **Financial Networks**



#### **ARTICLE**

Received 15 Jun 2016 | Accepted 23 Dec 2016 | Published 21 Feb 2017

DOI: 10.1038/ncomms14416

**OPEN** 

ws point ple, ference —

Pathways towards instability in financial networks

Marco Bardoscia<sup>1,2</sup>, Stefano Battiston<sup>1</sup>, Fabio Caccioli<sup>3,4</sup> & Guido Caldarelli<sup>2,5,6</sup>

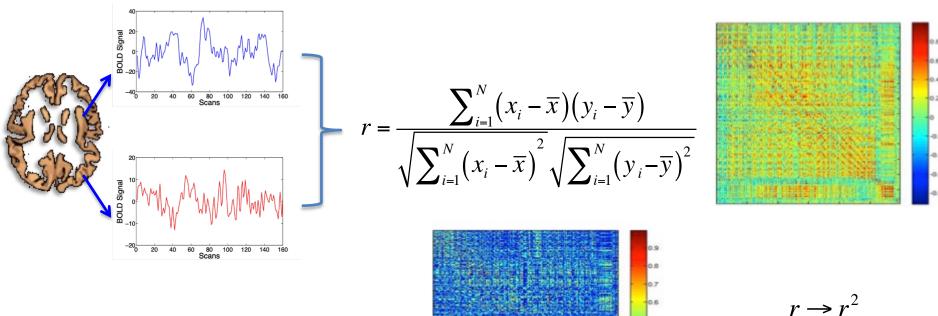






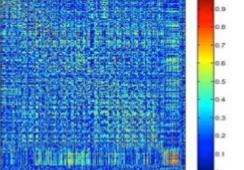
# **Research question: Brain Networks**

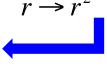
We can extract information on Functional Magnetic Resonance Imaging



R. Mastrandrea, F. Piras, A. Gabrielli, G. Caldarelli, G. Spalletta, T. Gili

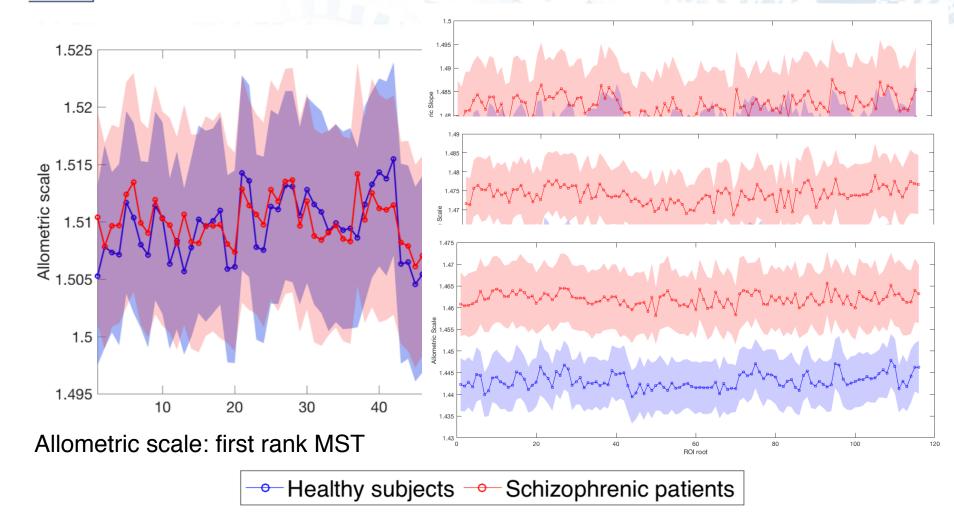
arXiv:1901.08521







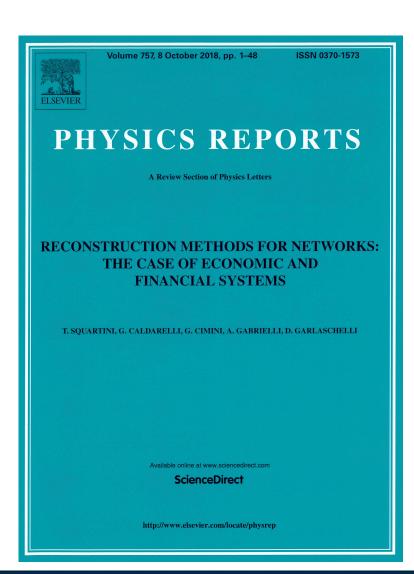
# Research question: Brain Networks



Can we use Networks for Disease Diagnosis?

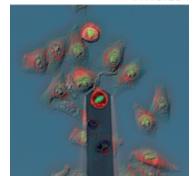


## **Statistical Physics of Networks**



REVIEWS

PHYSICS



**TECHNICAL REVIEWS** 

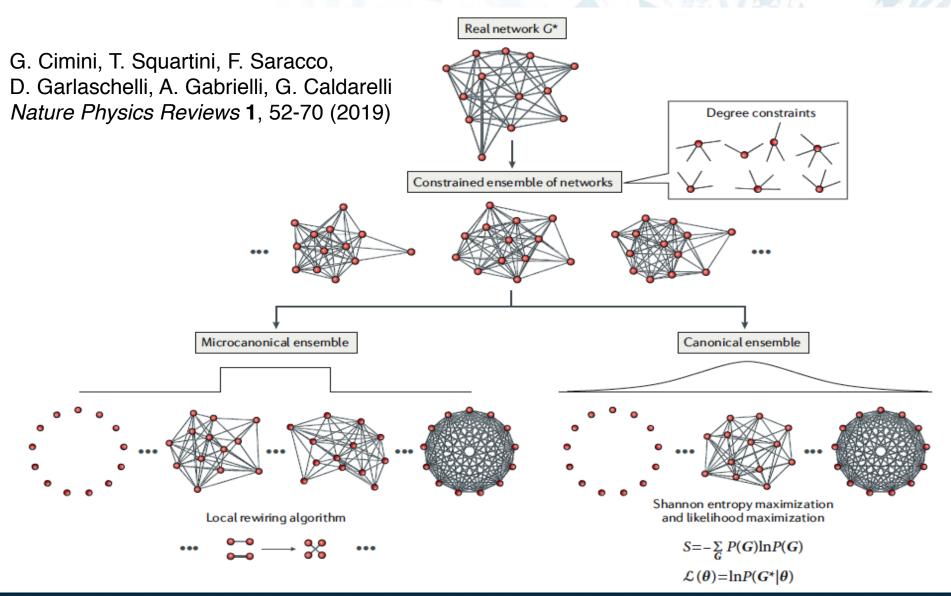
# The statistical physics of real-world networks

Giulio Ciminio 1.2, Tiziano Squartini 1, Fabio Saracco 1, Diego Garlaschelli 1.3, Andrea Gabrielli 1.2 and Guido Caldarelli 1.2.4.5 \*

Abstract | In the past 15 years, statistical physics has been successful as a framework for modelling complex networks. On the theoretical side, this approach has unveiled a variety of physical phenomena, such as the emergence of mixed distributions and ensemble non-equivalence, that are observed in heterogeneous networks but not in homogeneous systems. At the same time, thanks to the deep connection between the principle of maximum entropy and information theory, statistical physics has led to the definition of null models for networks that reproduce features of real-world systems but that are otherwise as random as possible. We review here the statistical physics approach and the null models for complex networks, focusing in particular on analytical frameworks that reproduce local network features. We show how these models have been used to detect statistically significant structural patterns in real-world networks and to reconstruct the network structure in cases of incomplete information. We further survey the statistical physics models that reproduce more complex, semilocal network features using Markov chain Monte Carlo sampling, as well as models of generalized network structures, such as multiplex networks, interacting networks and simplicial complexes.



# Reconstructing from partial information





# Concluding

### **COMPLEX NETWORKS**

ALLOW TO MEASURE (QUANTITATIVELY) SEVERAL RELATIONS

DESCRIBES DYNAMICAL PROCESSES AND TIME SERIES

MODELS THE EMERGENCE OF COMPLEXITY



# A Societal challenge

# nature physics

## correspondence

G. Caldarelli, S. Wolf, Y. Moreno Nature Physics 14 870 (2018).

# Physics of humans, physics for society

To the Editor — Today, the massive use of information and communication technologies (ICT) has made it possible to attach a traceable set of data to almost any person. We argue that these data provide the opportunity to build a 'physics of society': describing a society composed of many interacting heterogeneous entities (people, businesses, institutions) - as a physical system. While important ethical implications have to be taken into account, the benefits in developing such physics of society would be tremendous. Indeed, it could help understanding, anticipating and forecasting future societal trends and human behavioural responses, and their associated uncertainty1; or address societal challenges in which globally networked risks play a role23. A case in point is modern epidemiology and its success in predicting the large-scale spreading of infectious diseases4.

like to find quantities similar to pressure the integral of the particles' impulse — but, since individuals do not follow a well-behaved Maxwell-Boltzmann distribution function, this could simply be unachievable.

Yet, despite all of the above, we argue that physics can still play a pivotal role in the quest to find regularities in societal dynamics based on patterns of human interactions. Sociotechnical systems where human interactions are partly mediated by ICT exhibit emerging dynamics and are more often than not out of equilibrium. They also involve many temporal and spatial scales, are governed by nonlinear effects and can adapt to external and internal perturbations. We have roughly all the ingredients of a truly out-of-equilibrium complex system, for the study of which physics provides methods and tools.

How should we proceed? The first step, in

would say necessary to address social problems. An example is data-driven stochastic microsimulation of the Zika virus epidemic<sup>10</sup>, which helps our understanding of a global and socially relevant dynamics, and at the same time contributes to the more traditional problem of characterizing diffusion in disordered media.

However, the pursuit of a physics for society would bear fruit only if physicists were ready to leave their traditional comfort zone and establish unconventional collaborations with researchers in other disciplines — such as computer and social scientists, and economists, in addition to mathematicians. We believe that such vision and agenda can only be realized with the support of international organizations like the European Commission, which triggered the development of the field with its Future and Emerging Technology





## **NETWORKS@IMTLUCCA.IT**



**ISC-CNR** 



**Elsewhere** 



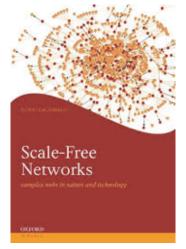




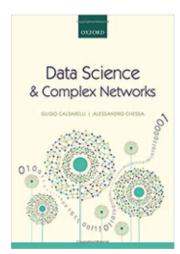




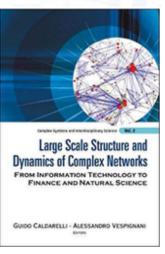
## **Bibliography**



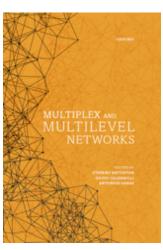
**OUP 2007** 



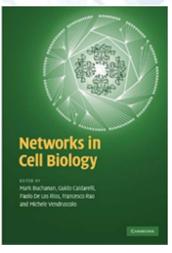
OUP 2016



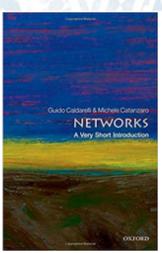
WSP 2007



**OUP 2019** 



CUP 2010



**OUP 2012** 

Nature Physics Reviews 1, 52-70 (2019)

Physics Reports 757, 1-47 (2018)

PNAS **115** 6548-6553 (2018)

Nature Physics 14 870 (2018)

PNAS 114, 3035-3039 (2017)

Nature Communications 8 14416 (2017)

PNAS 113 10031-10035 (2016)

PNAS 113 554-559 (2016)

Nature Communications 6,8627 (2015)

Nature Physics 9, 125 (2013)

Nature Physics 9, 123 (2013)