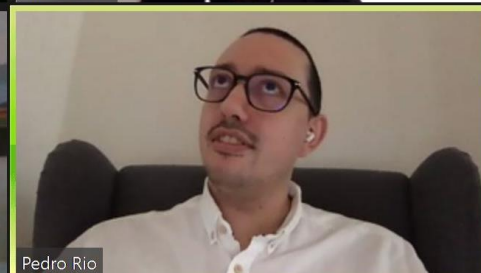
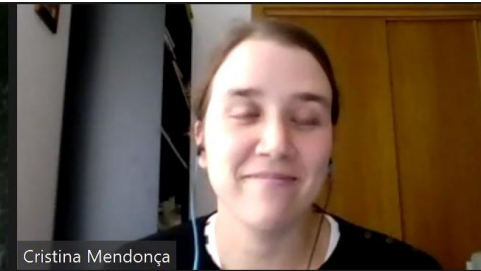

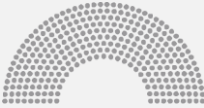



Social Physics and Complexity

Using big data to understand human interactions

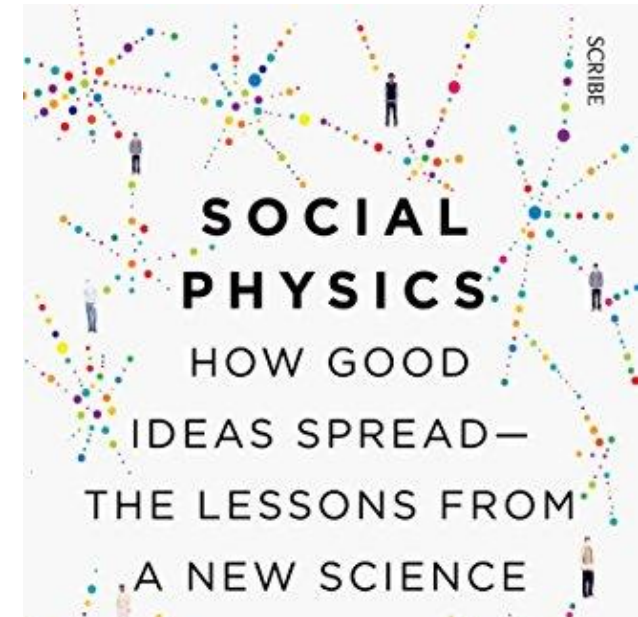


	QUESTIONS	DATA	TOOLS
HEALTH	 <p>Online vs. Offline Patterns Emergency Now-casting Antibiotic Over-prescription</p>	<p>Google Trends SNS24 Twitter ER acceptance /times SPMS e-prescriptions</p>	<p>Math Modelling ML Epidemiology</p>
POLICY	 <p>Political Decisions Gender Differences Agenda Setting Voting vs. Discourse</p>	<p>Media records Twitter Facebook Parliament data</p>	<p>NLP Networks Math Modelling Complex Systems</p>
BEHAVIOUR	 <p>Cognitive Biases Attitudes Towards Science Tracking Anxiety</p>	<p>Large scale surveys Behavioral experiments Twitter Facebook</p>	<p>Networks Math Modelling Psychology Information</p>

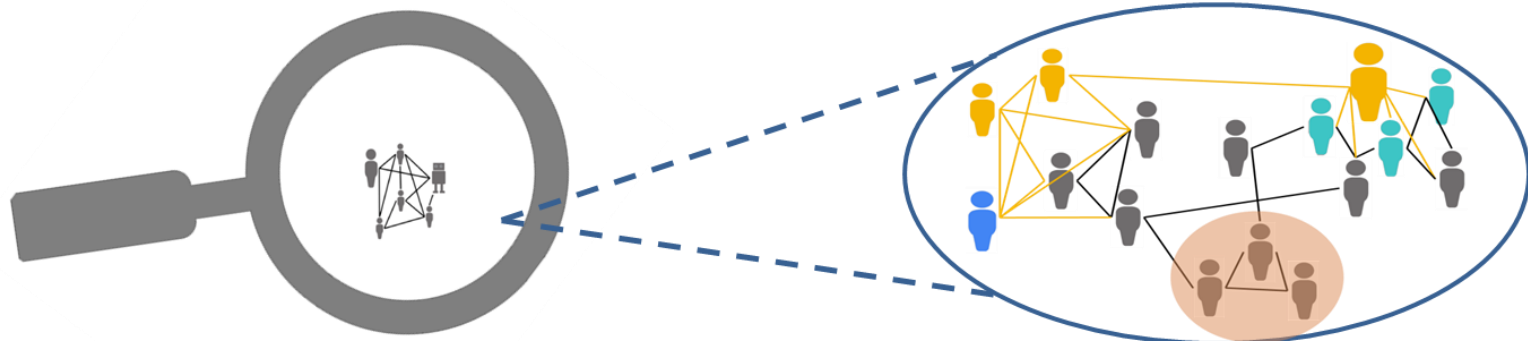
These problems—and a wide range of similar problems in the biological, medical, psychological, economic, and political sciences—are just too complicated to yield to the old nineteenth-century techniques which were so dramatically successful on two-, three-, or four-variable problems of simplicity. These new problems, moreover, cannot be handled with the statistical techniques so effective in describing average behavior in problems of disorganized complexity.

These new problems, and the future of the world depends on many of them, requires science to make a third great advance, an advance that must be even greater than the nineteenth-century conquest of problems of simplicity or the twentieth-century victory over problems of disorganized complexity. Science must, over the next 50 years, learn to deal with these problems of organized complexity.

Warren Weaver, 1947



Alex Pentland, 2014



From the Industrial Revolution to the Digital Revolution



First

Water and steam power is used to create mechanical production facilities.

Second

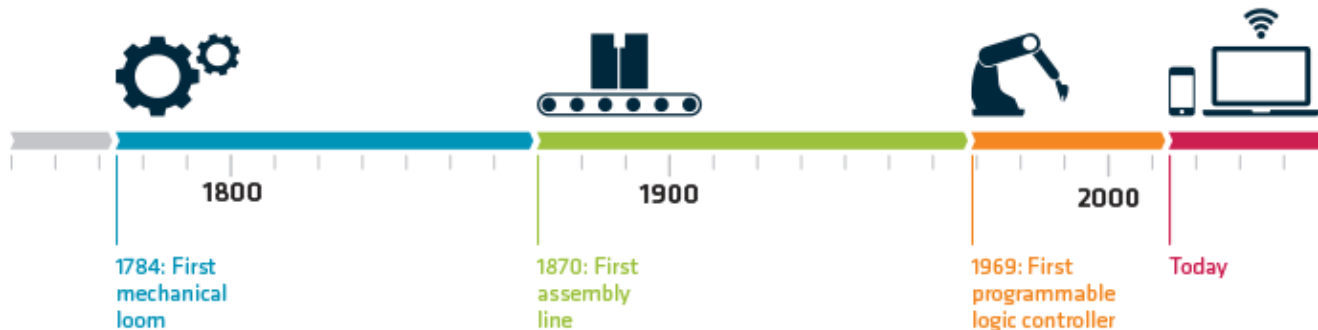
Electricity lets us create a division of labor and mass production.

Third

IT systems automate production lines further.

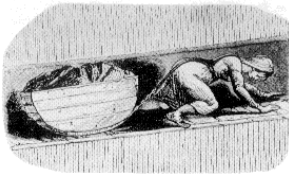
Fourth

IoT and cloud technology automate complex tasks.



Source: <https://mjolner.dk/2015/01/14/realizing-the-fourth-industrial-revolution/>

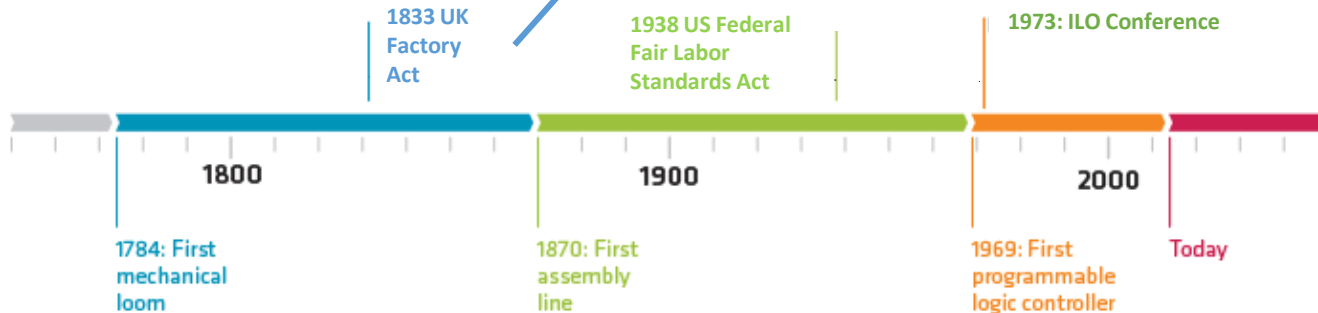
From the Industrial Revolution to the Digital Revolution



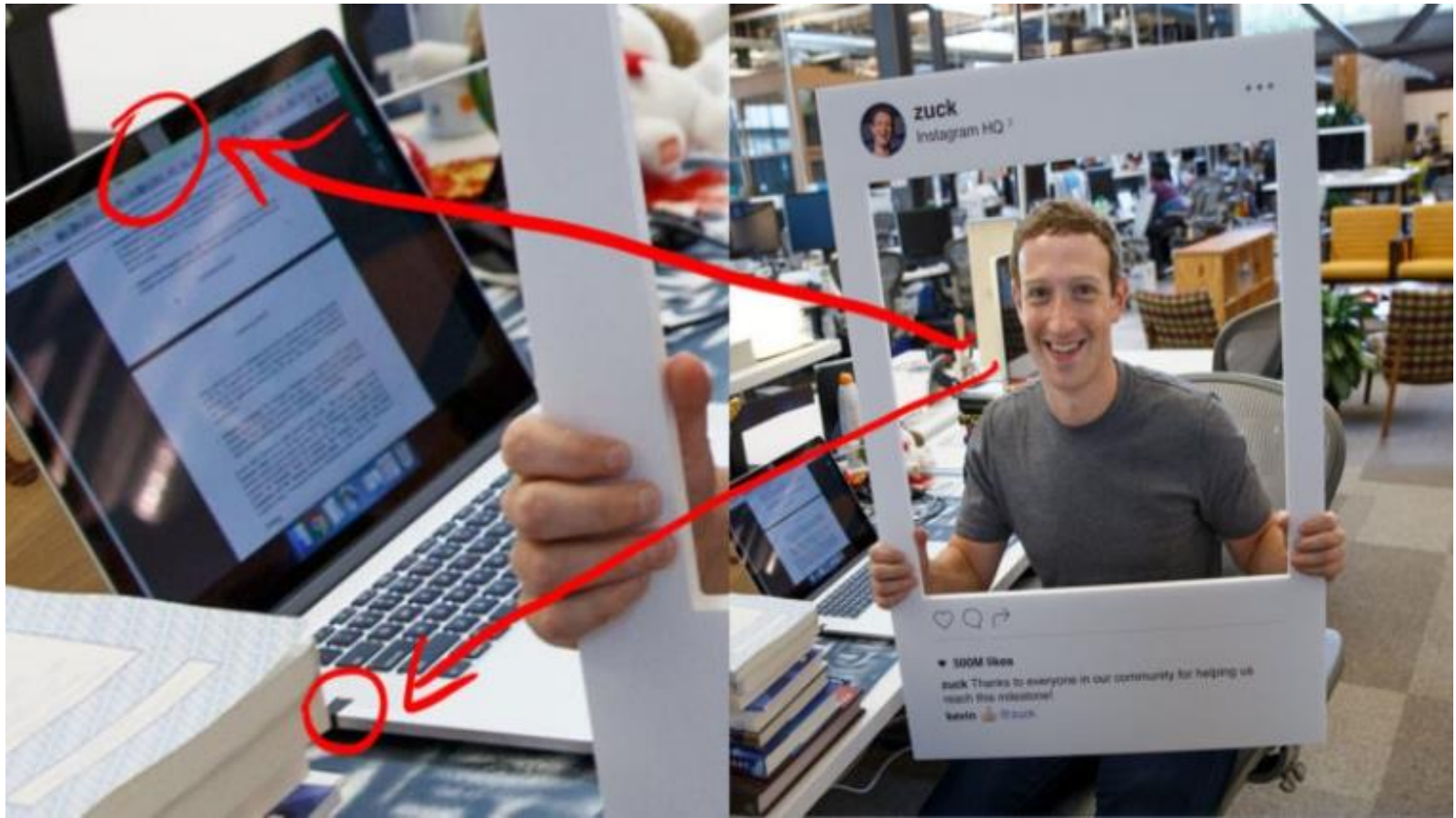
Child Labor in the Industrial Revolution



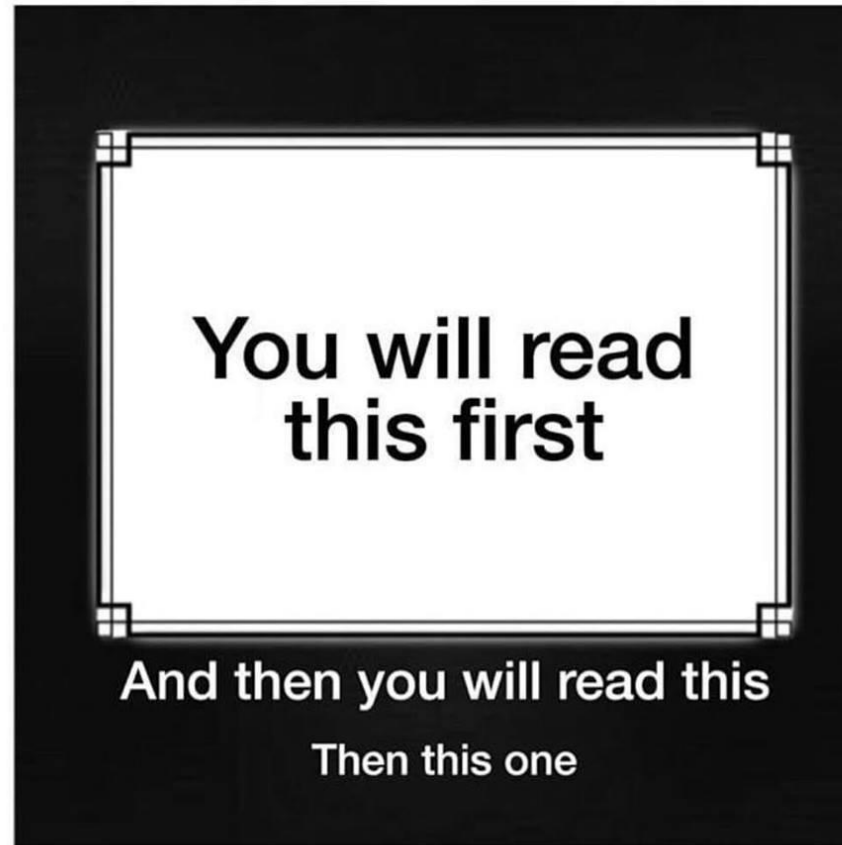
- 9 year old limit
- 9-13 yo should not work > 9h a day
- 13-18yo should not work > 12h a day
- Four inspectors

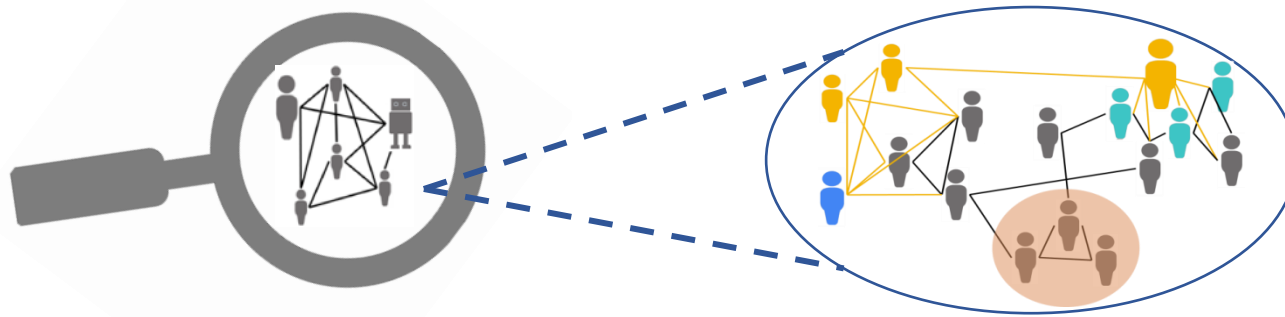


Source: <https://mjolner.dk/2015/01/14/realizing-the-fourth-industrial-revolution/>



And you will read this at the end





WE HAVE CREATED A “**MACROSCOPE**” BUT
WE DON’T KNOW HOW TO USE IT

why am i so

why am i so **tired**
why am i so **ugly**
why am i so **gassy**
why am i so **thirsty**
why am i so **angry**
why am i so **itchy**
why am i so **sad**
why am i so **hungry**
why am i so **emotional**
why am i so **bloated**

how to

how to **make slime**
how to **tie a tie**
how to **buy bitcoin**
how to **lose weight**
how to **draw**
how to **buy ripple**
how to **kiss**
how to **make pancakes**
how to **mine bitcoin**
how to **train your dragon**

como posso ser|

como posso ser **amigo de alguem**
como posso ser **feliz**
como posso ser **inteligente**
como posso ser **uma pessoa melhor**
como posso ser **salvo**
como posso ser **rico**
como posso ser **feliz sozinho**
como posso ser **um hacker**
como posso ser **popular no facebook**
como posso ser **cantora**

como é que se

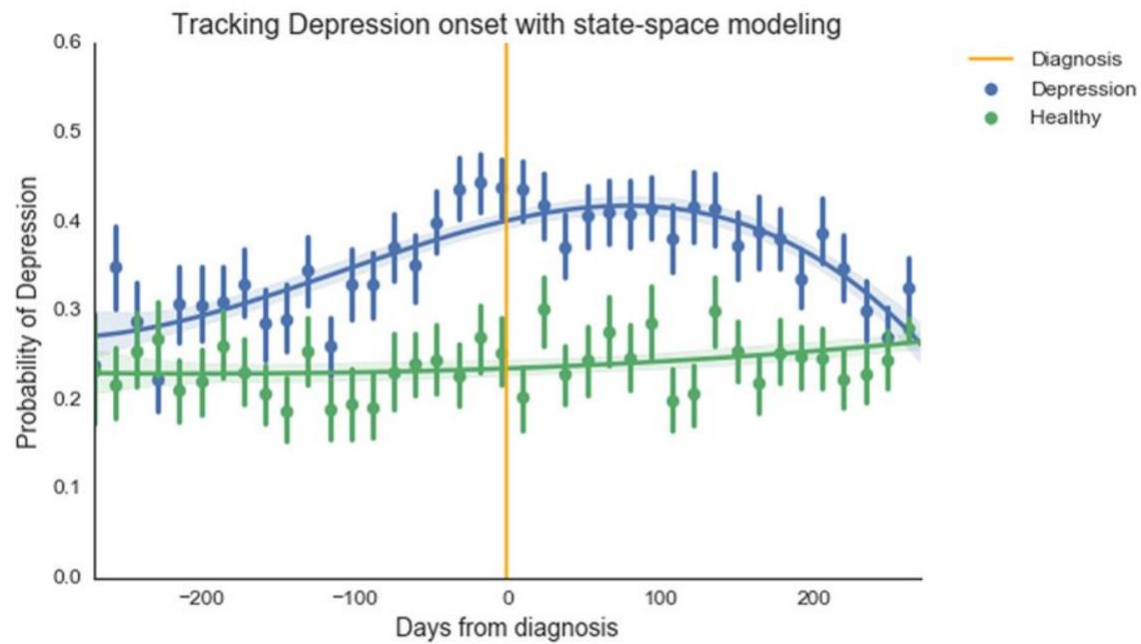
como é que se **beija**
como é que se **diz eu te amo**
como é que se **beija de lingua**
como é que se **engravidar**
como é que se **beija na boca**
como é que se **escreve**
como é que se **beija pela primeira vez**
como é que se **faz um facebook**
como é que se **faz um relatório**
como é que se **faz panquecas**

pourquoi je suis

pourquoi je suis **moche**
pourquoi je suis **triste**
pourquoi je suis **toujours fatigué**
pourquoi je suis **célibataire**
pourquoi je suis **toujours célibataire**
pourquoi je suis **devenu rebelle pdf**
pourquoi je suis **seule**
pourquoi je suis **toujours fatiguée**
pourquoi je suis **jalouse**
pourquoi je suis **triste sans raison**

comment faire

comment faire **du slime**
comment faire **un cv**
comment faire **des crepes**
comment faire **une dissertation**
comment faire **une capture d'écran**
comment faire **une bibliographie**
comment faire **un gateau**
comment faire **du caramel**
comment faire **de la glue**
comment faire **du pain**

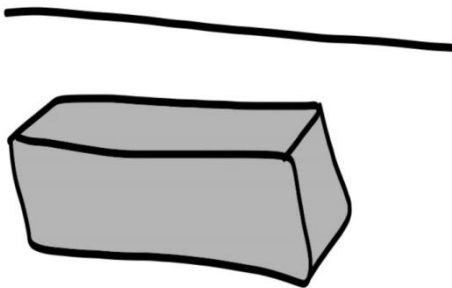


<https://www.nature.com/articles/s41598-017-12961-9>

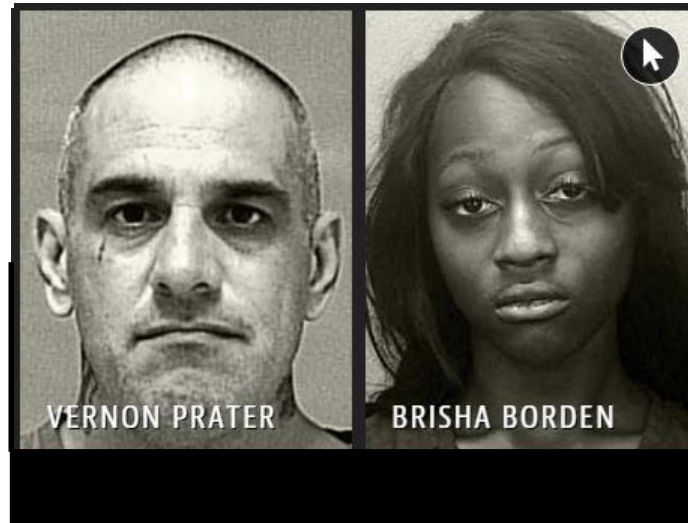
WILL Iot HELP US BECOME HAPPIER?

What if a picture knew it was making you feel calmer, more mindful...just happier?

Using sensors, apps & museums to enable wellbeing



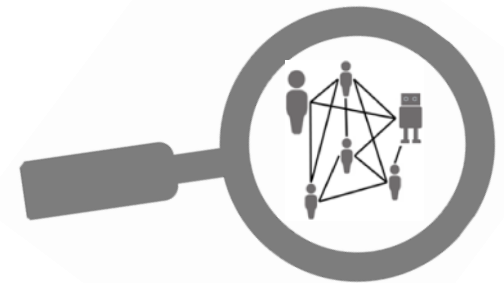
Punitive!



<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

In summary

- Human behavior is a complex problem
- We have growing data to tackle that complexity
- The data we produce is a reflection of our biases
- We should acknowledge those biases



Example: How does information spread on social networks?

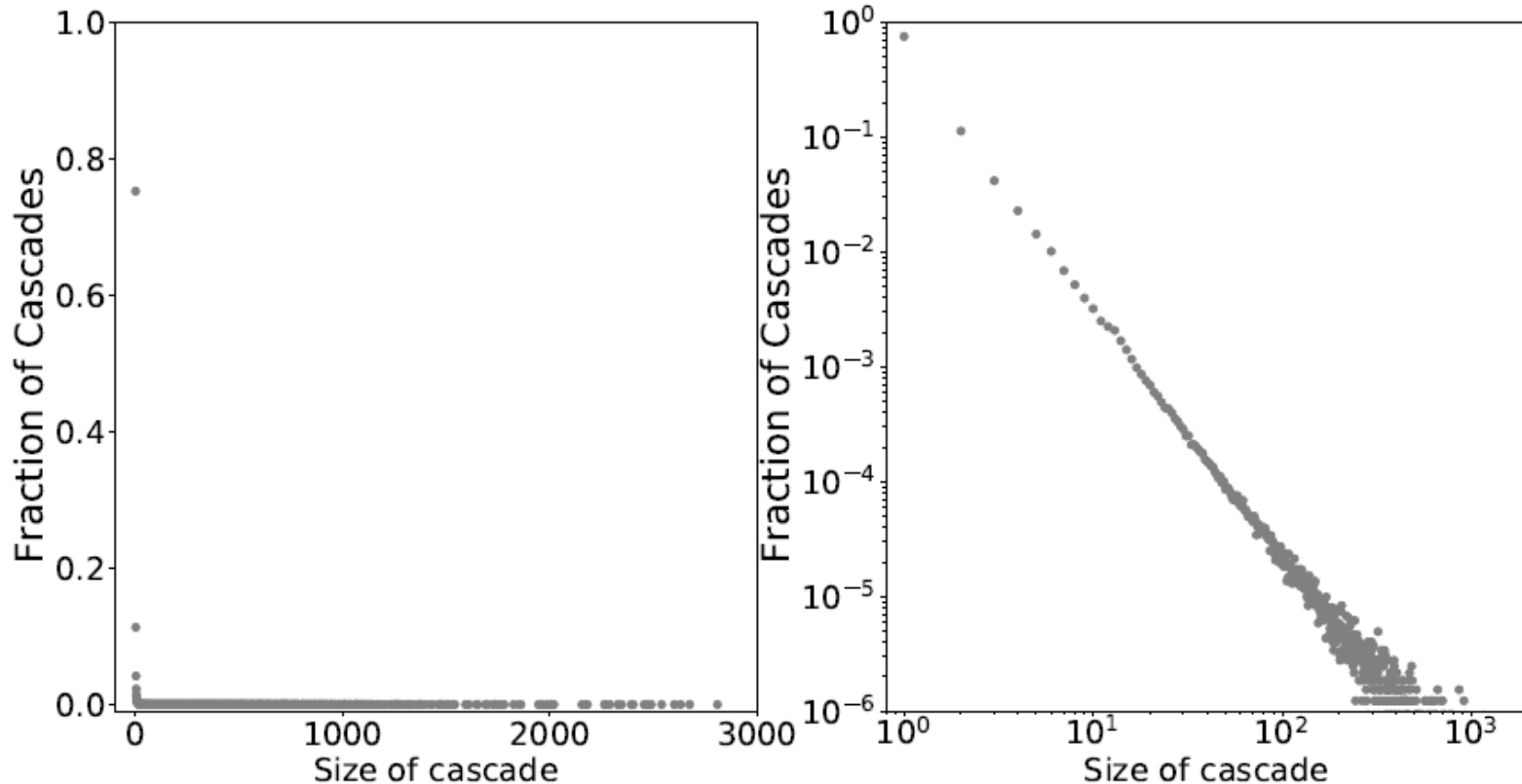
Twitter as a model system

Two important definitions

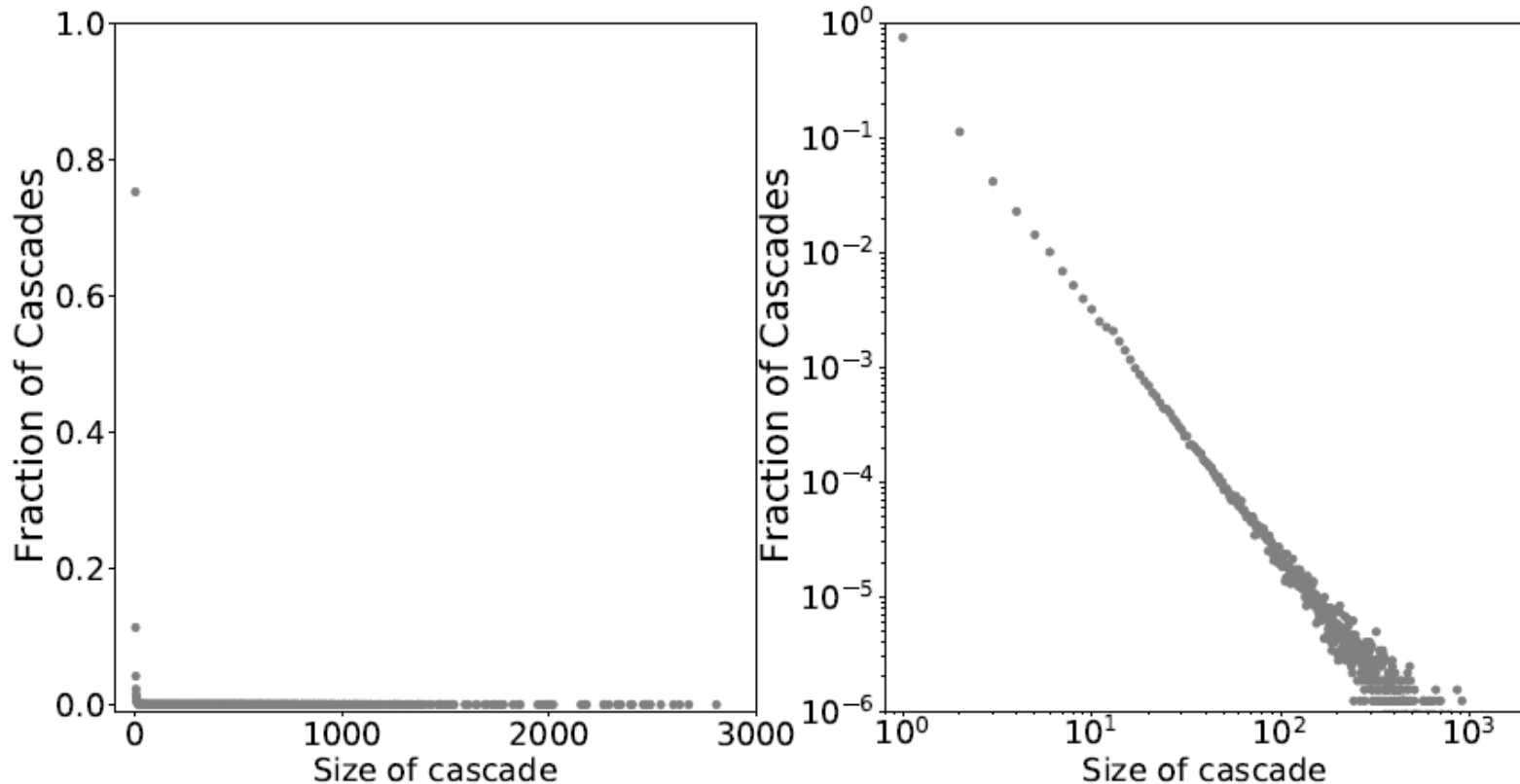
Tweet cascade: The set of all copies (re-tweets) of an original tweet.

We measure **Fitness** the rate at which the number of copies of a tweet grows.

Re-tweet cascade sizes follow a power law distribution



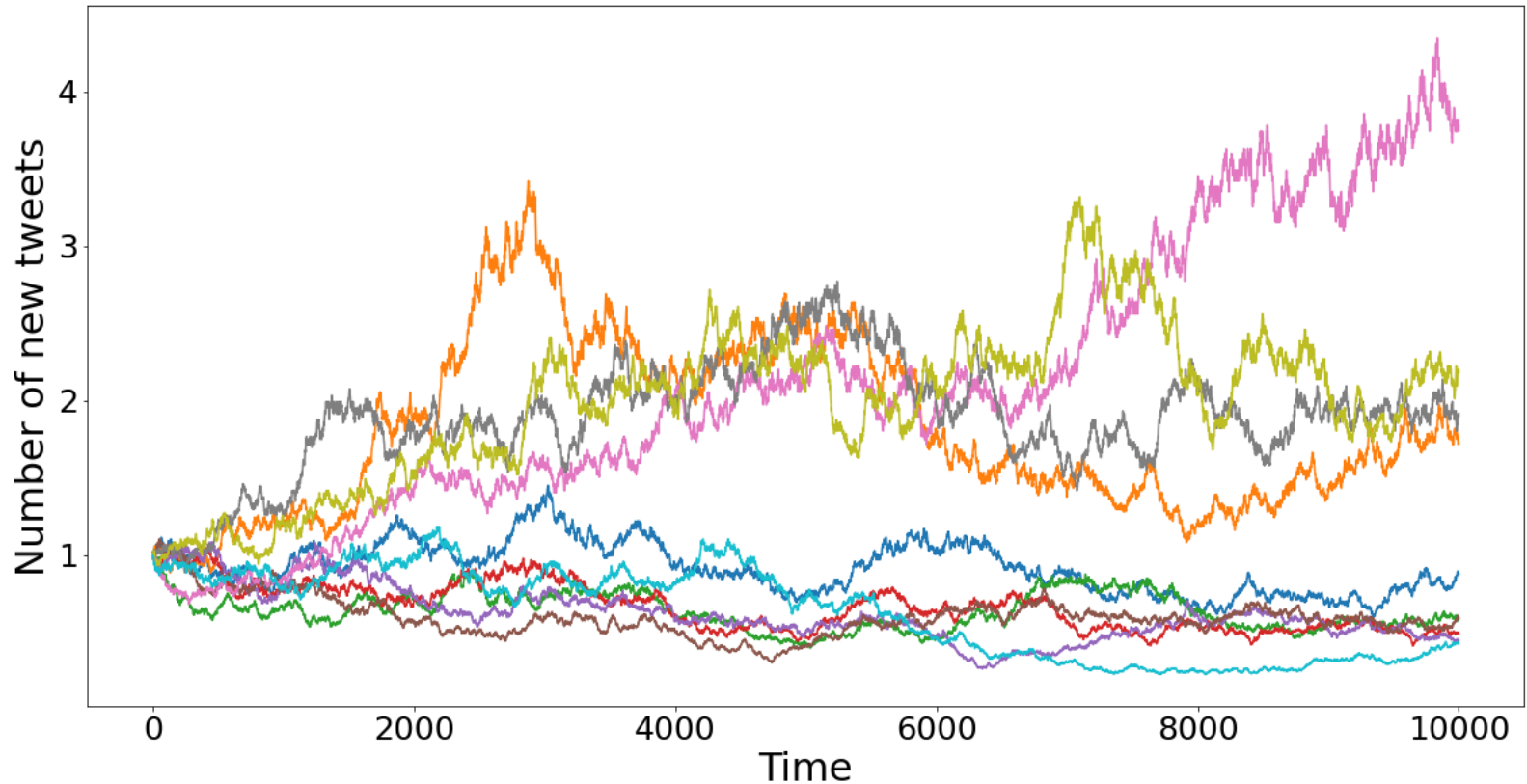
Re-tweet cascade sizes follow a power law distribution



... but does that mean the underlying process is also scale-free?



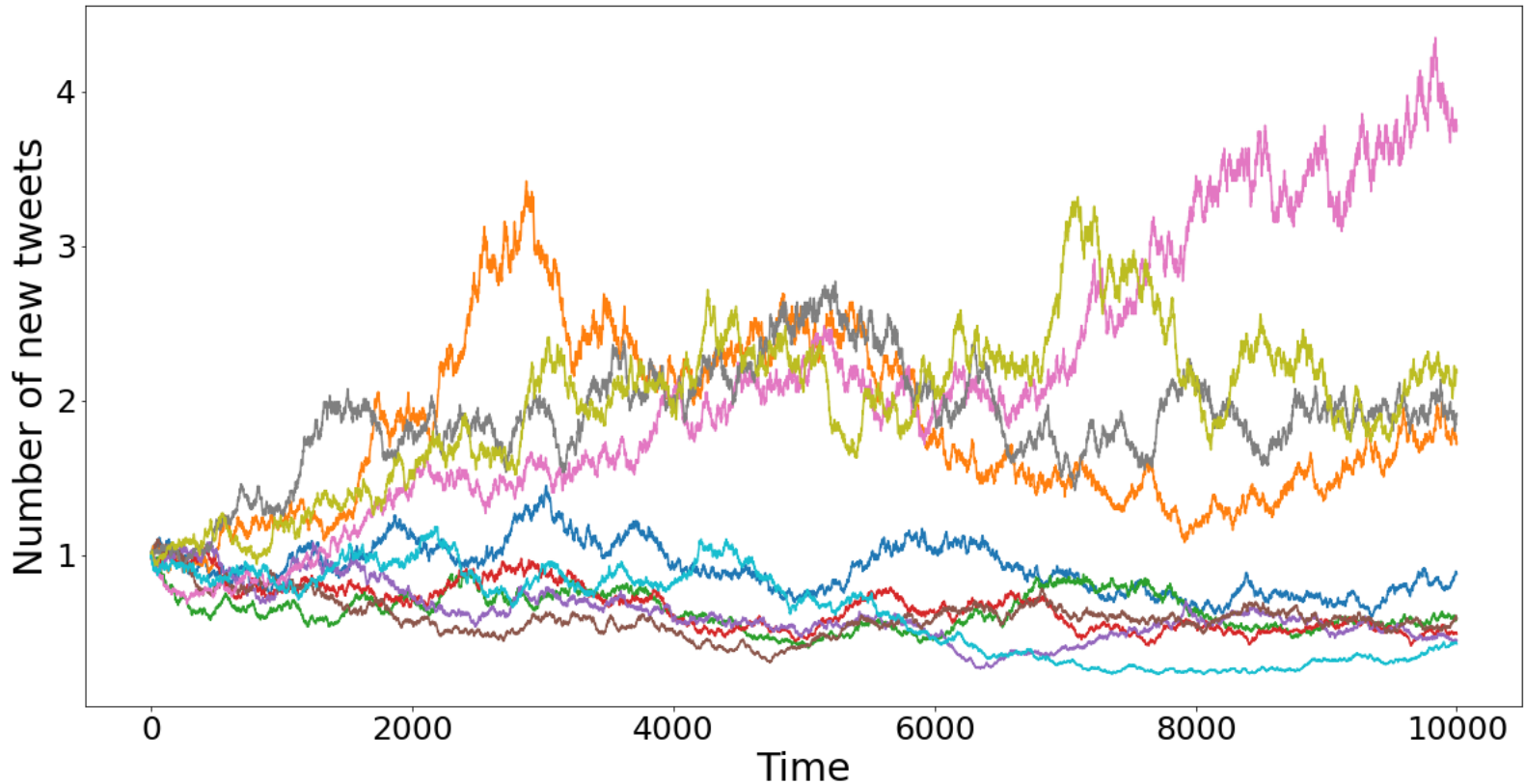
Cascade growth as a diffusion process



$$\frac{dx}{dt} = x\epsilon$$

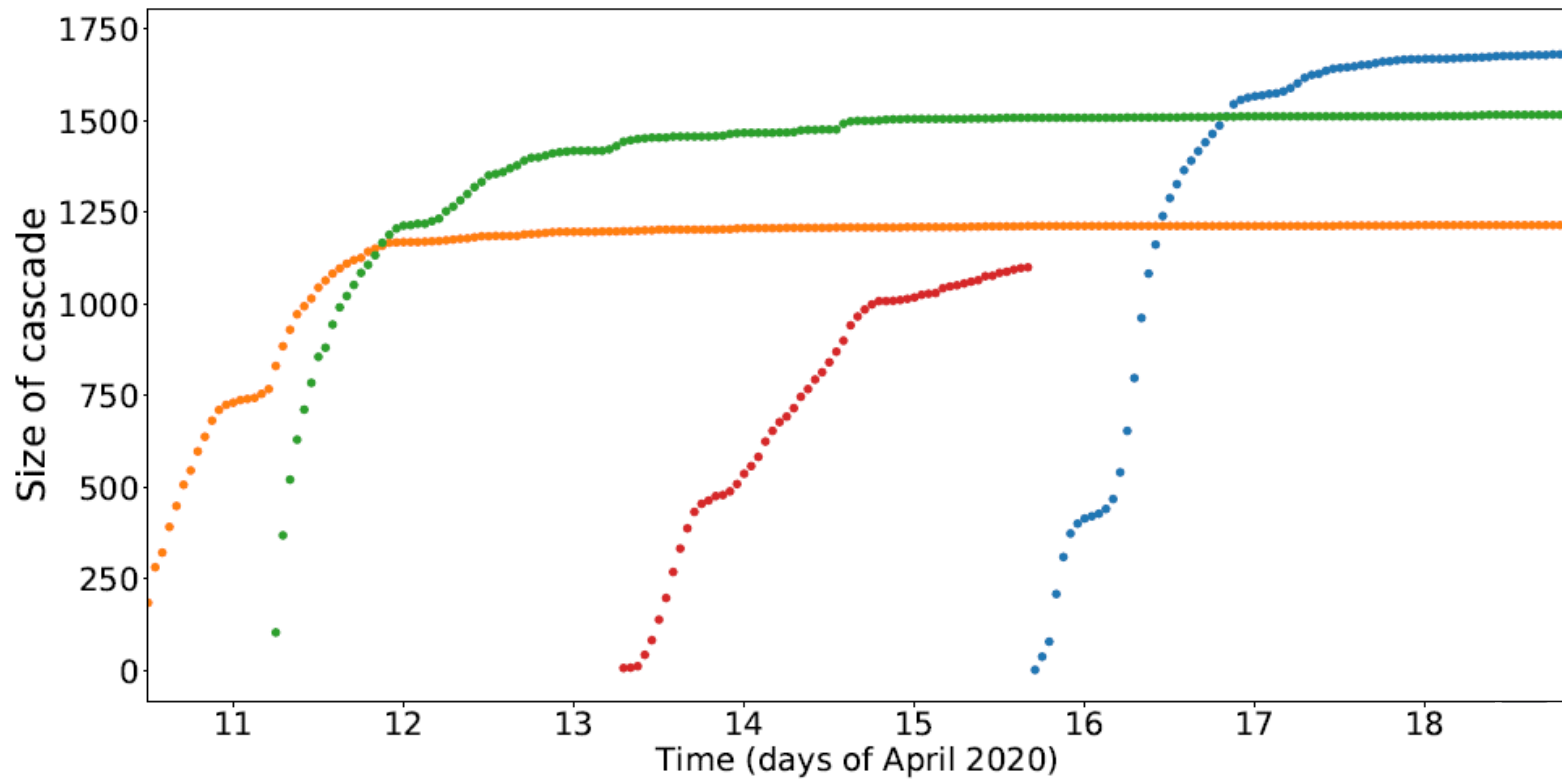
$$\frac{d \log(x)}{dt} = \log(x) + \log(\epsilon), \text{ where } \log(\epsilon) \text{ is a Gaussian random variable.}$$

Cascade growth as a diffusion process

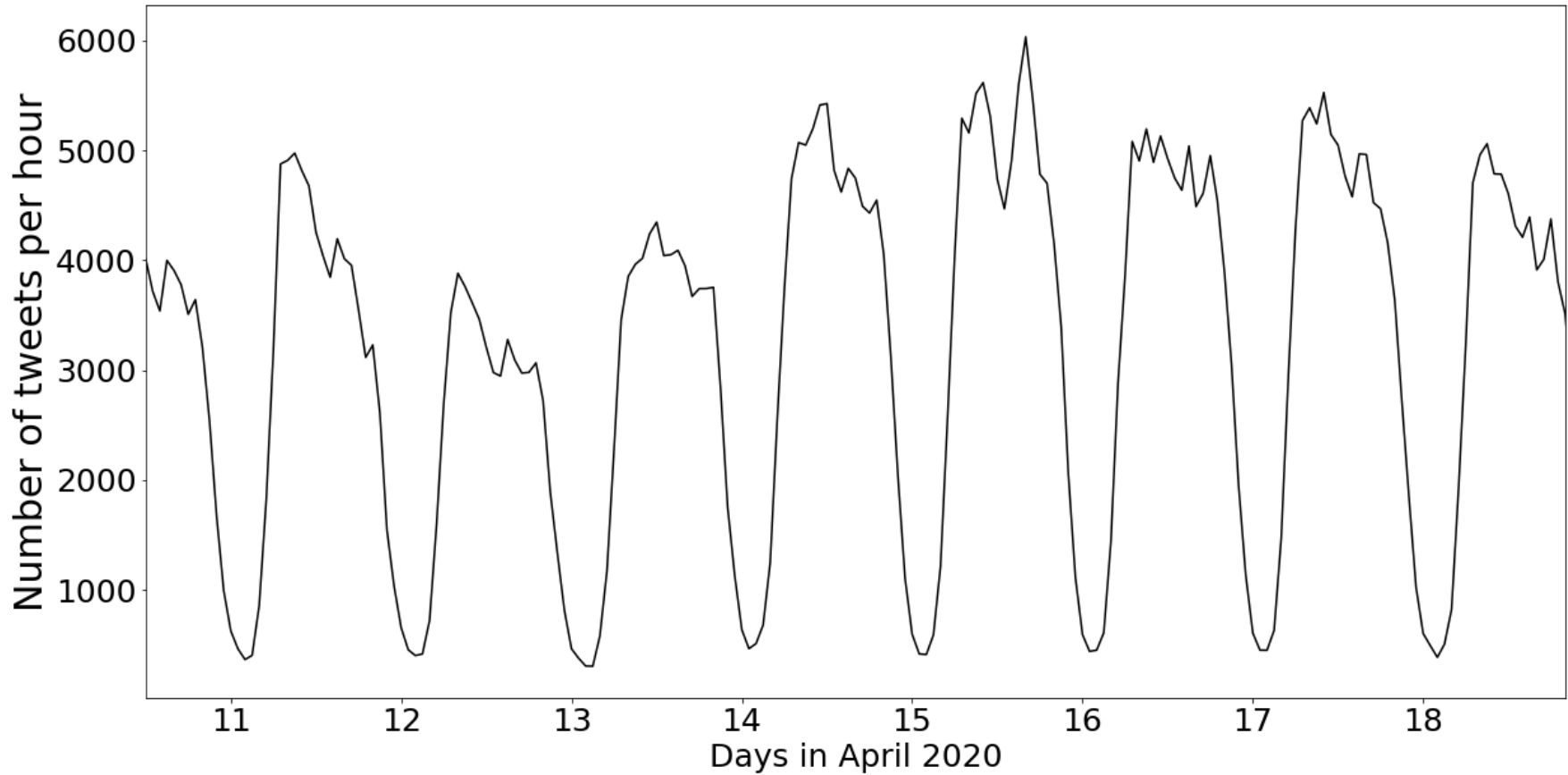


$$\frac{dx}{dt} = a \cdot x + x \cdot \epsilon$$

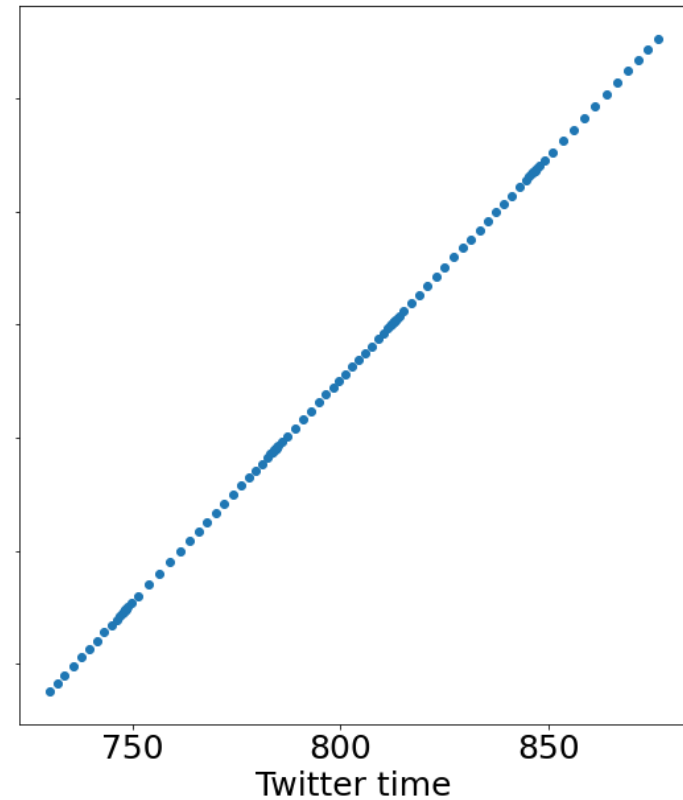
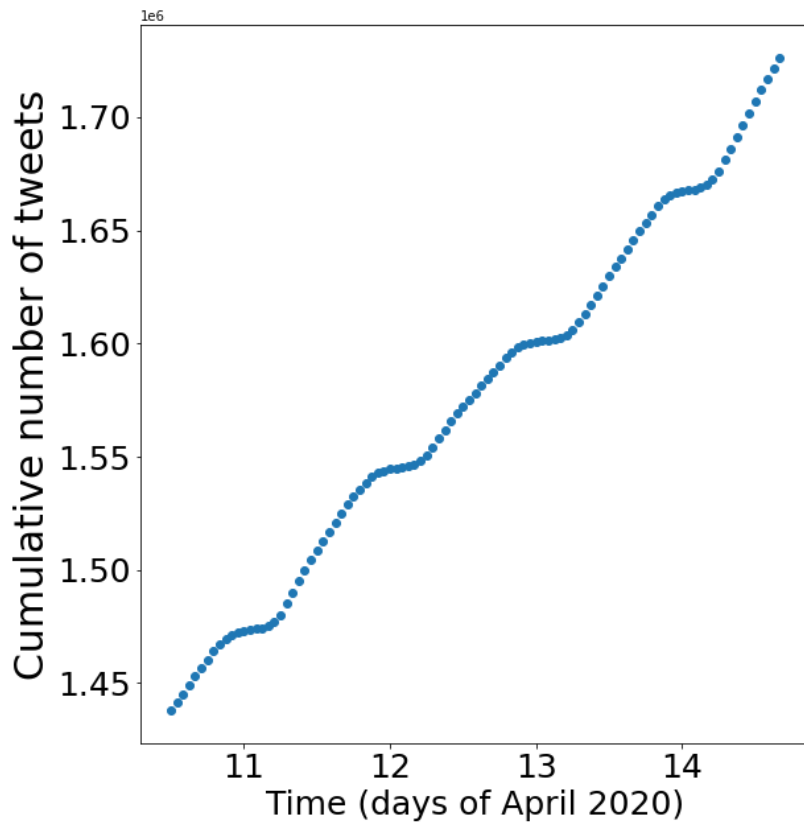
Cascade growth



Introducing Twitter time

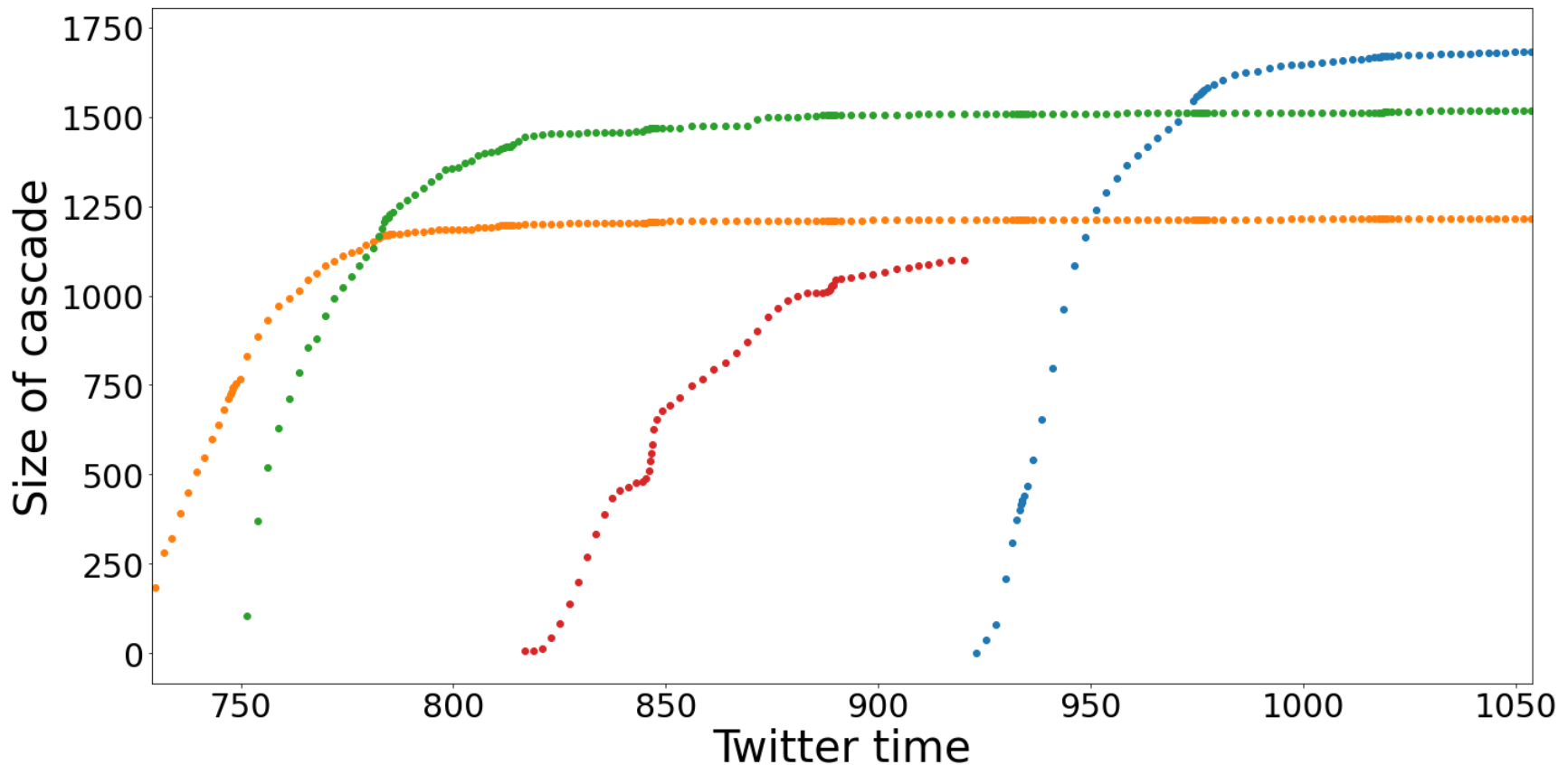


Introducing Twitter time



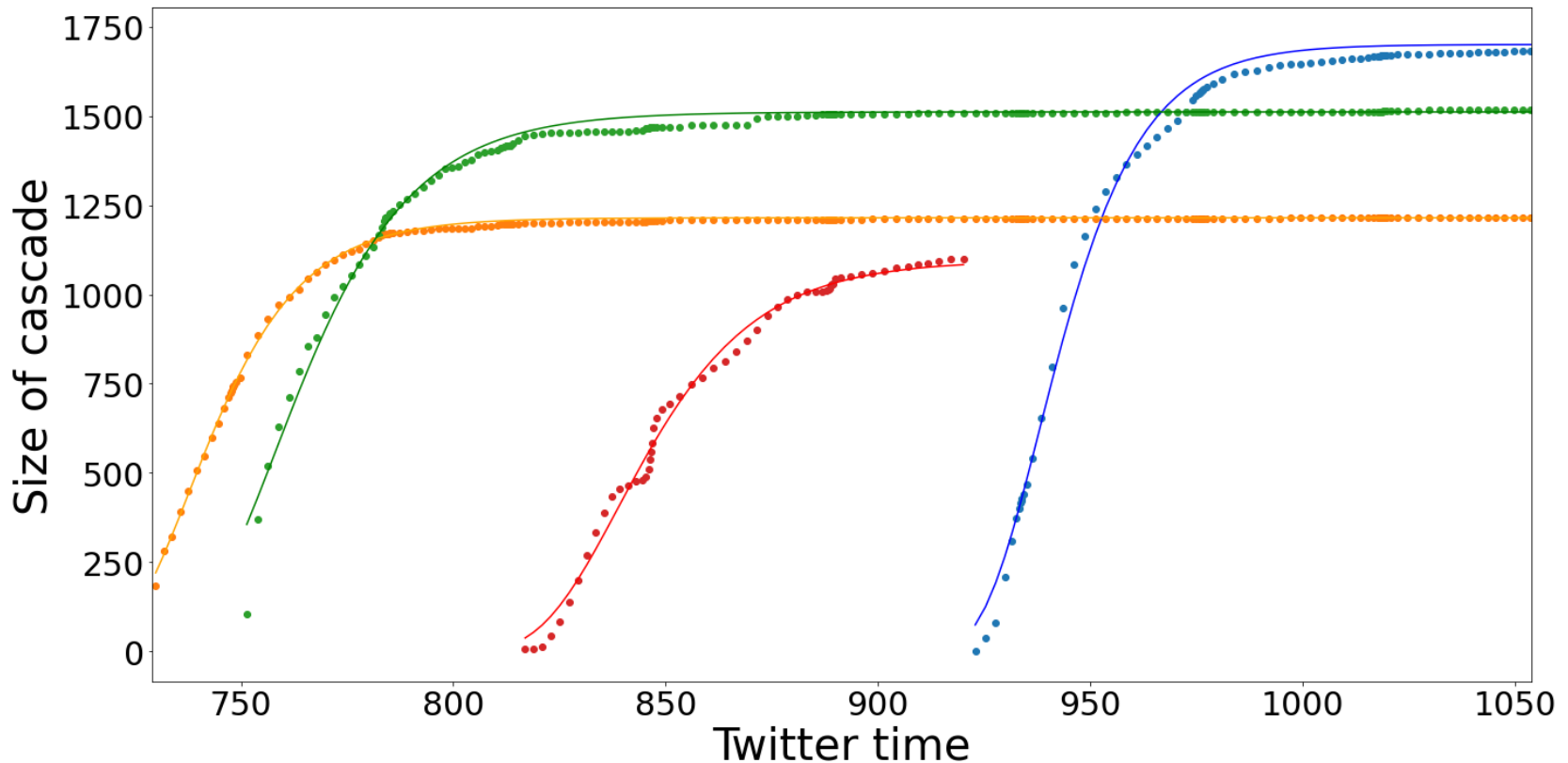
$$\frac{dT_t}{dt} = \frac{dN_{total}}{dt} \cdot C$$

The simplest model for cascade growth



$$\frac{dN}{dt} = N \cdot a \cdot e^{-g \cdot t} \epsilon$$

The simplest model for cascade growth



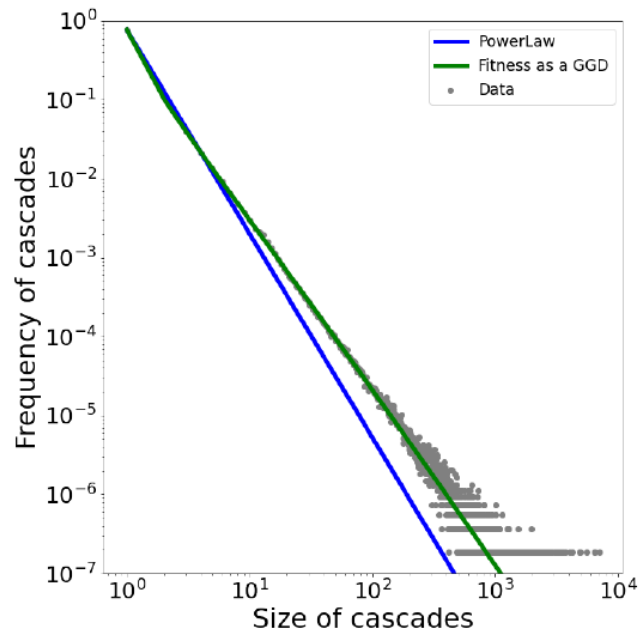
$$N(t) = N(0) \cdot e^{\frac{a}{g}} e^{a \cdot e^{-g \cdot t}}$$

$$\frac{dN}{dt} = N \cdot a \cdot e^{-g \cdot t}; \quad N(t) = N(0) \cdot e^{\frac{a}{g}} e^{a \cdot e^{-g \cdot t}}$$

$$\lim_{t \rightarrow +\infty} N(t) = e^{\frac{a}{g}}$$

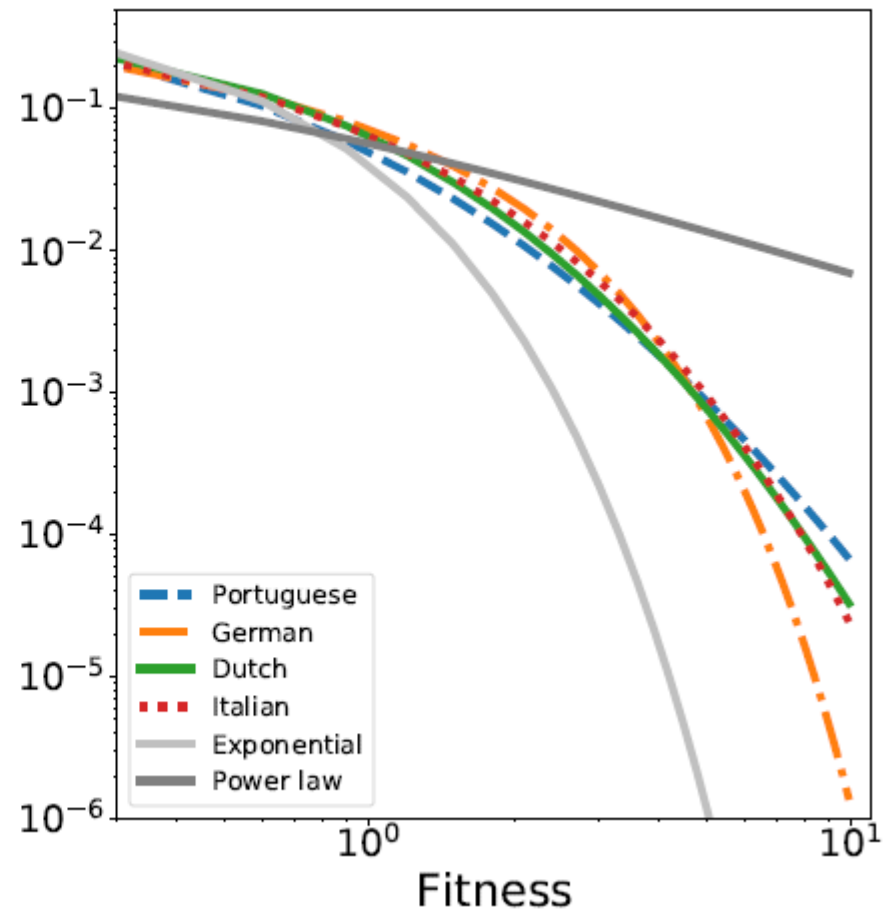
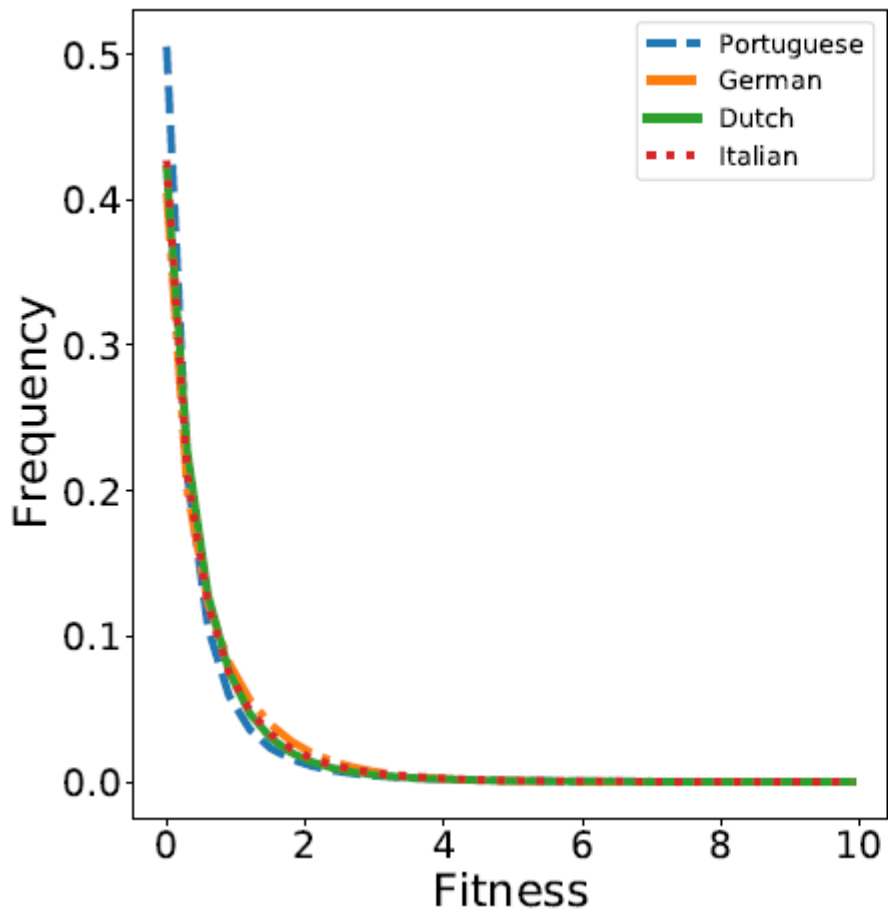
$$\lim_{t \rightarrow +\infty} \text{dist}(N) = \frac{1}{N} \cdot \text{dist}_{\frac{a}{g}}(\log(N))$$

FITTING THE GENERALIZED GAMMA DISTRIBUTION (GGD) TO FITNESS



$$f(x, a, c, \theta) = \frac{|c| \cdot x^{c \cdot a - 1} \cdot e^{-\frac{x^c}{\theta}}}{\theta^{c \cdot a} \cdot \Gamma(a)}$$

	a	c	top 20%
Portuguese	60	0.11	0.31
German	0.27	1.3	0.46
Dutch	65	0.14	0.36
Italian	1.0	0.69	0.36



Ongoing

- Add properties of the social network. Does it improve the fit? (numerical solutions and simulations)
- Compare different types of information – such as real *versus* fake news
- Include demographic and psychological characteristics of individuals
- ?

FARE - Fake News and Real People - Using Big Data to Understand Human Behaviour

<https://cordis.europa.eu/project/id/853566>

<https://tinyurl.com/SPAComplexity>



European Research Council

Established by the European Commission



lperfeito@lip.pt



joanagsa@lip.pt



@SpacLab @Dr_Lil @mjoanasa