

SPAC – Social Physics and Complexity @LIP

November 8th 2021



SCIENCE AND COMPLEXITY

By WARREN WEAVER

Rockefeller Foundation, New York City



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.

SCIENCE AND COMPLEXITY

By WARREN WEAVER

Rockefeller Foundation, New York City



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.

QUESTIONS

DATA

TOOLS

HEALTH

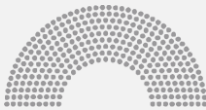


Online vs. Offline Patterns
Emergency Now-casting
Antibiotic Over-prescription

Google Trends
SNS24
Twitter
ER acceptance /times
SPMS e-prescriptions

Math Modelling
ML
Epidemiology

POLICY



Political Decisions
Gender Differences
Agenda Setting
Voting vs. Discourse

Media records
Twitter
Facebook
Parliament data

NLP
Networks
Math Modelling
Complex Systems

BEHAVIOUR



Cognitive Biases
Attitudes Towards Science
Tracking Anxiety

Large scale surveys
Behavioral experiments
Twitter
Facebook

Networks
Math Modelling
Psychology
Information



Fake News



Pathogens



Humans



Hosts

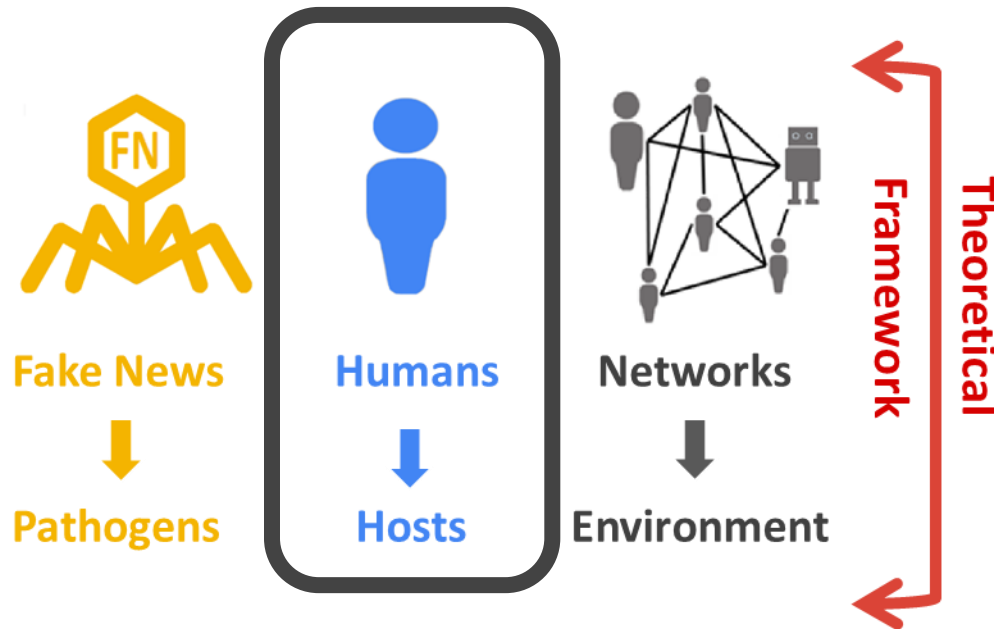


Networks



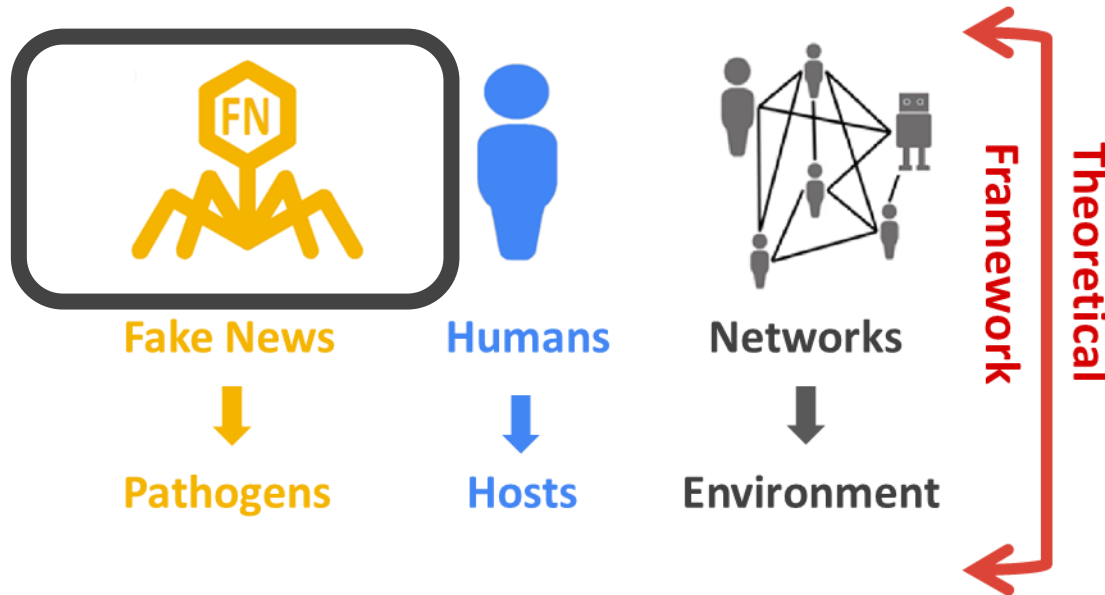
Environment





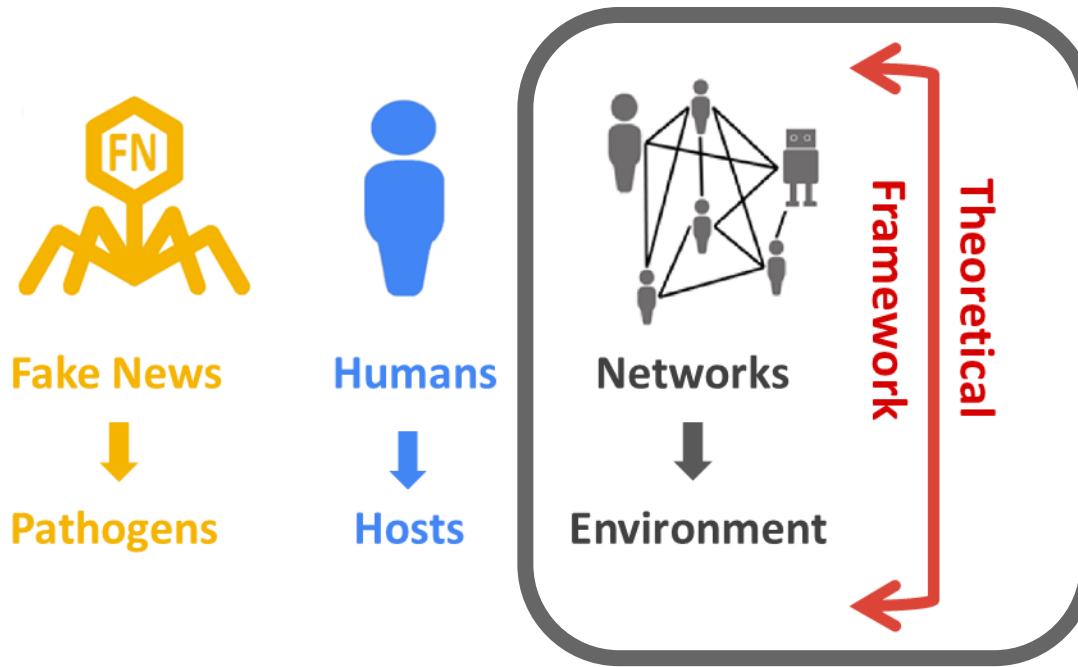
1- INDIVIDUAL CHARACTERISTICS THAT PROMOTE “DISEASE” SUSCEPTIBILITY

Frederico Francisco, Simone Lackner, Cristina Mendonça, Ângela Rijo



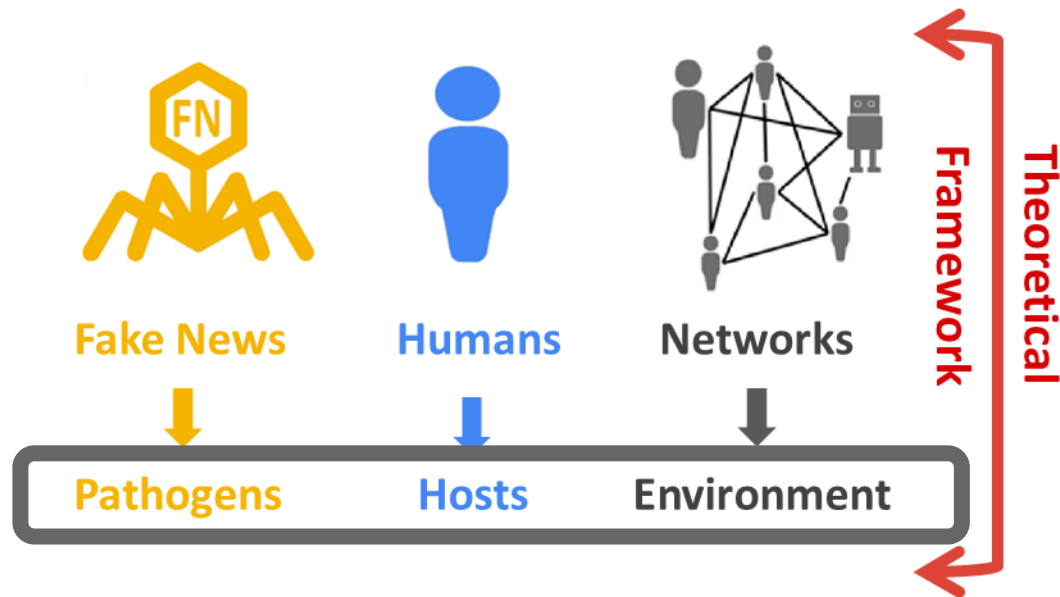
2- DISINFORMATION, TRACKING, AND BEHAVIORAL TARGETING

Íris Damião, Alex Davidson (former member), Cristina Mendonça, Lília Perfeito, José Reis



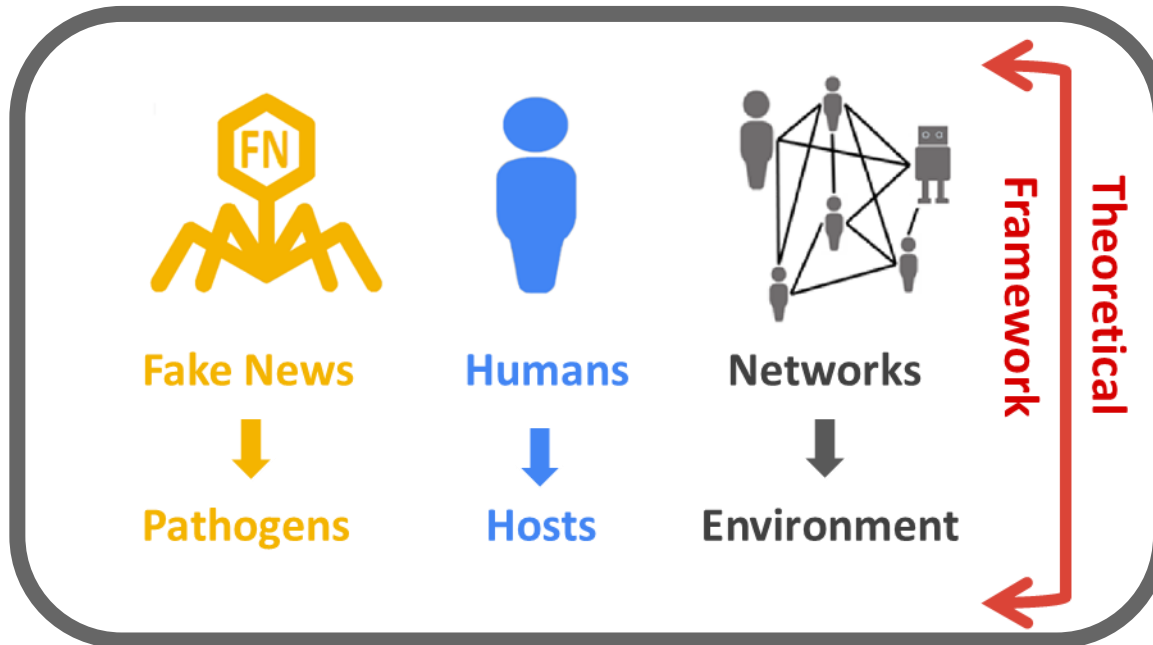
3- INFORMATION FLOW ON NETWORKS

Paulo Almeida, Pedro Duarte, Tiago Miranda, Lília Perfeito



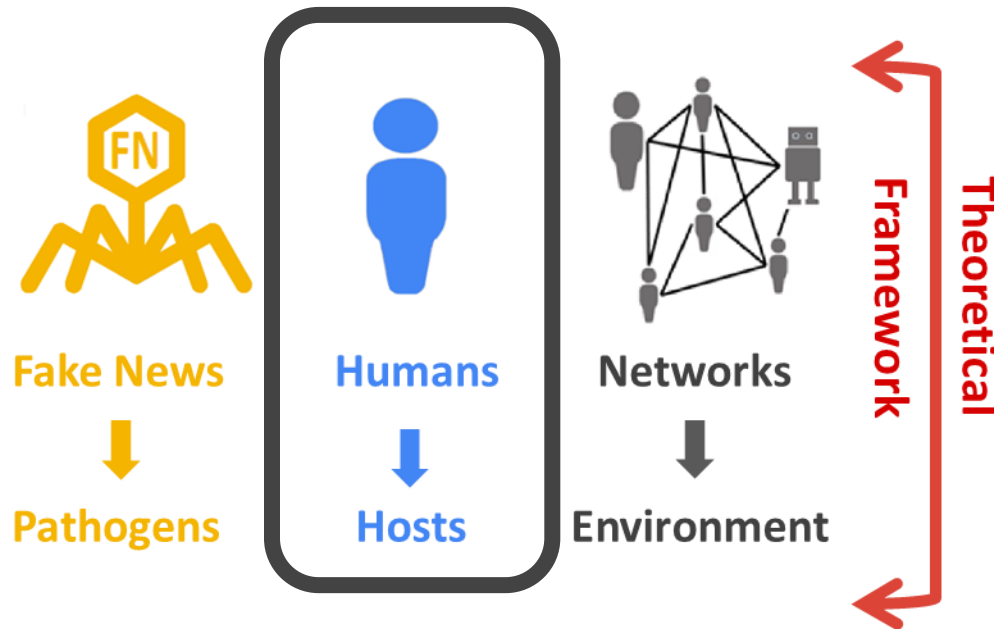
4 - DYNAMICS OF NON-INFLUENZA RESPIRATORY VIRUSES (NIRVS)

João Loureiro, Sara Mesquita, Lília Perfeito, Pedro Rio, Eleonora Tulumello, Irma Varela



5- PROTECTING PERSONAL DATA WITH MULTI-PARTY COMPUTATION

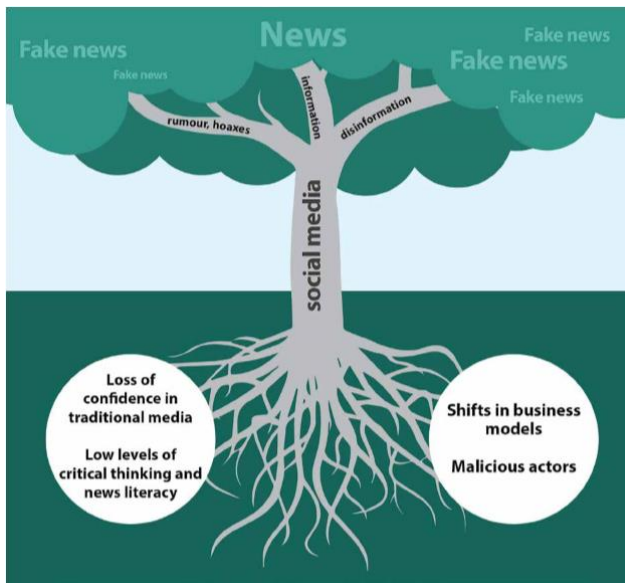
Paulo Almeida, Hugo Cachitas, Alex Davidson (former member)



INDIVIDUAL CHARACTERISTICS THAT PROMOTE “DISEASE” SUSCEPTIBILITY

Frederico Francisco, Simone Lackner, Cristina Mendonça, Ângela Rijo

Individual susceptibility



[UNESCO's World Trends Report](#)

Question:

What individual characteristics explain irrational judgment and behavior (e.g., fake news sharing)?

Objectives:

- a. **Identify** individual characteristics that influence negative judgment / behavior.
- b. **Test** whether individual characteristics improve our “disease” models.

Individual susceptibility

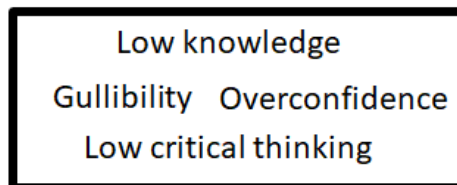
Datasets:

- a. Already existing large-scale surveys (demographics, attitudes, opinions, perception of risk, etc.).

Survey	Start	Frequency	Countries	Sample size
Eurobarometer	1974	Twice a year	EU countries	~1,000 per country
European Social Survey	2001	Every two years	30+ countries	~800 – 1,500 per country
General Social Survey	1972	Every two years	United States	~1,500 – 6,000

- b. Surveys / experiments we design ourselves, e.g.:

Individual characteristics (survey)



Linked to



Individual susceptibility

Tools:

- A. Statistics & data visualization.
- B. Agent-based modeling.
- C. Integration of this individual-level data with environment and pathogen data.



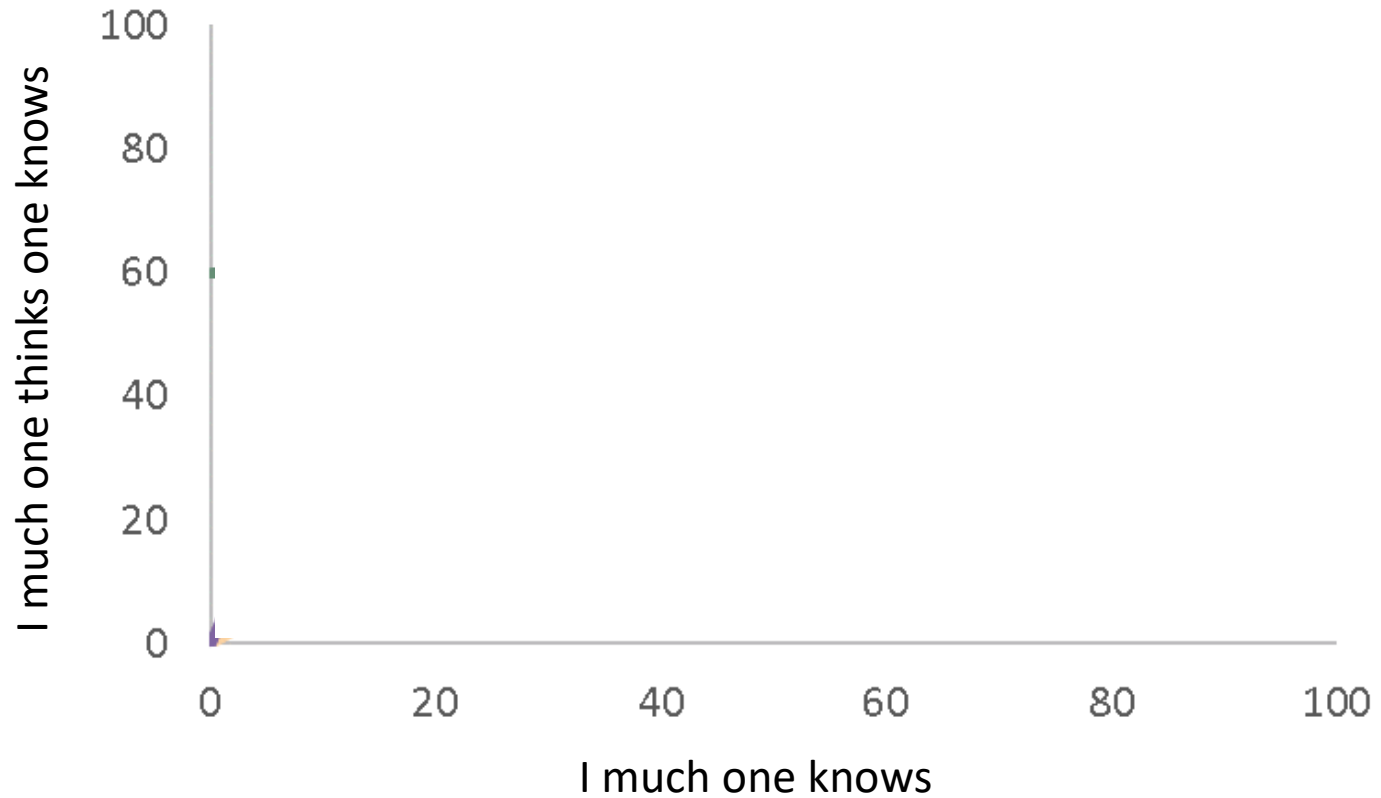
February 7 at 1:00 PM · 🌐



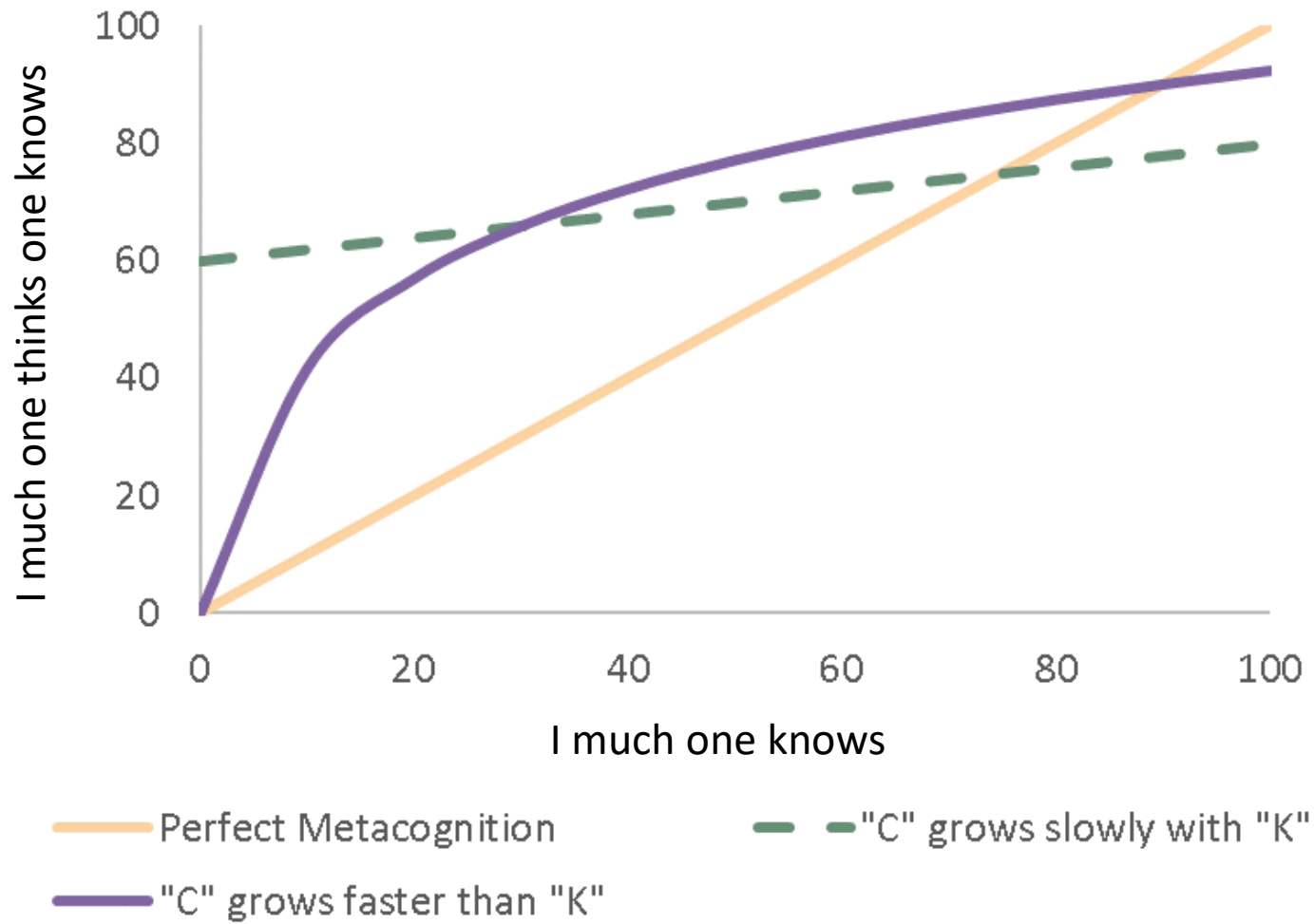
Hay naku ! So it's been around then ? 😞 & How did they know about a Virus we knew nothing about ? 😞 Well ,This is definitely a Government Conspiracy lysol labels as proof the government knew in advance about the Corona virus .

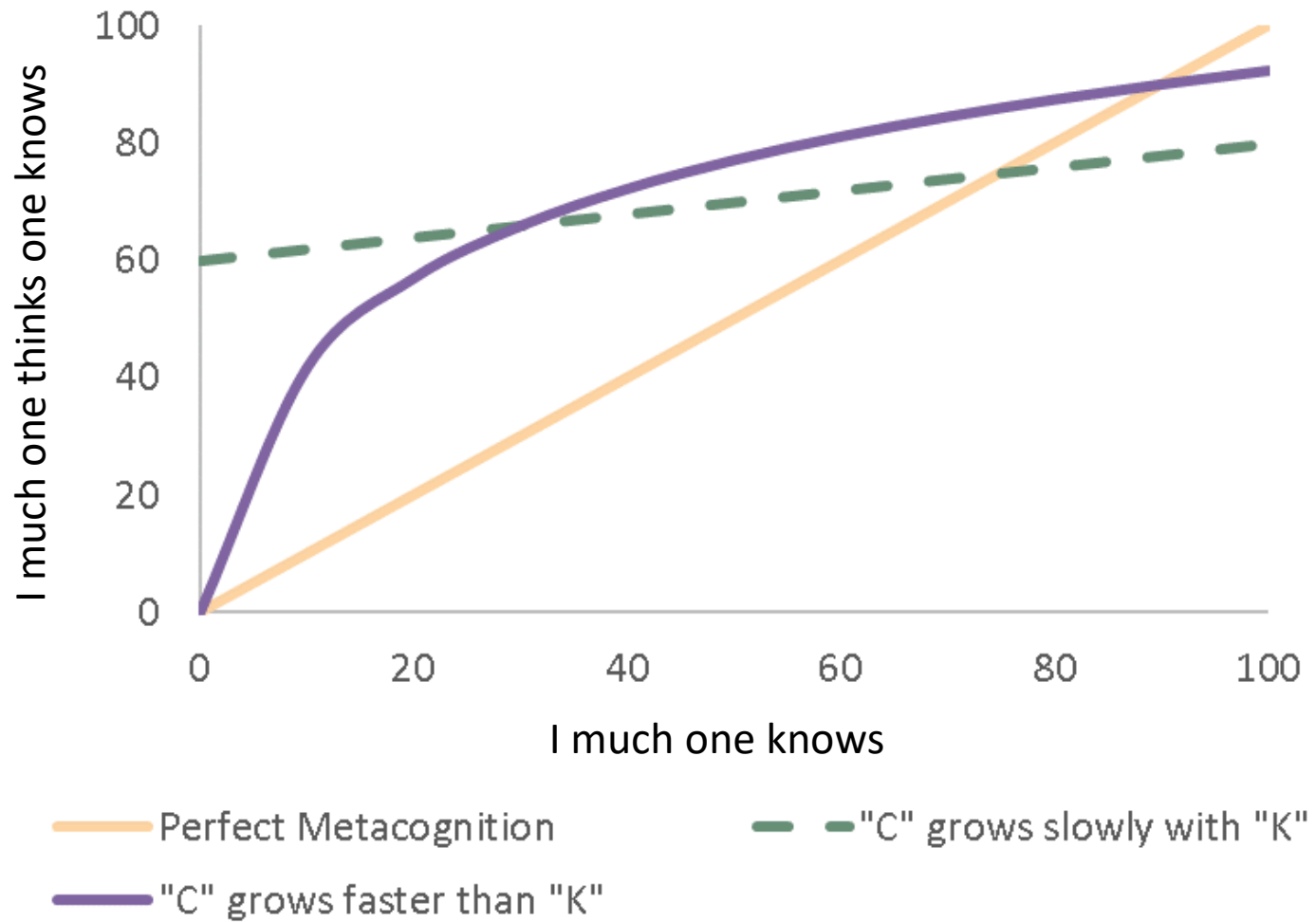
Corona Virus is nothing new.





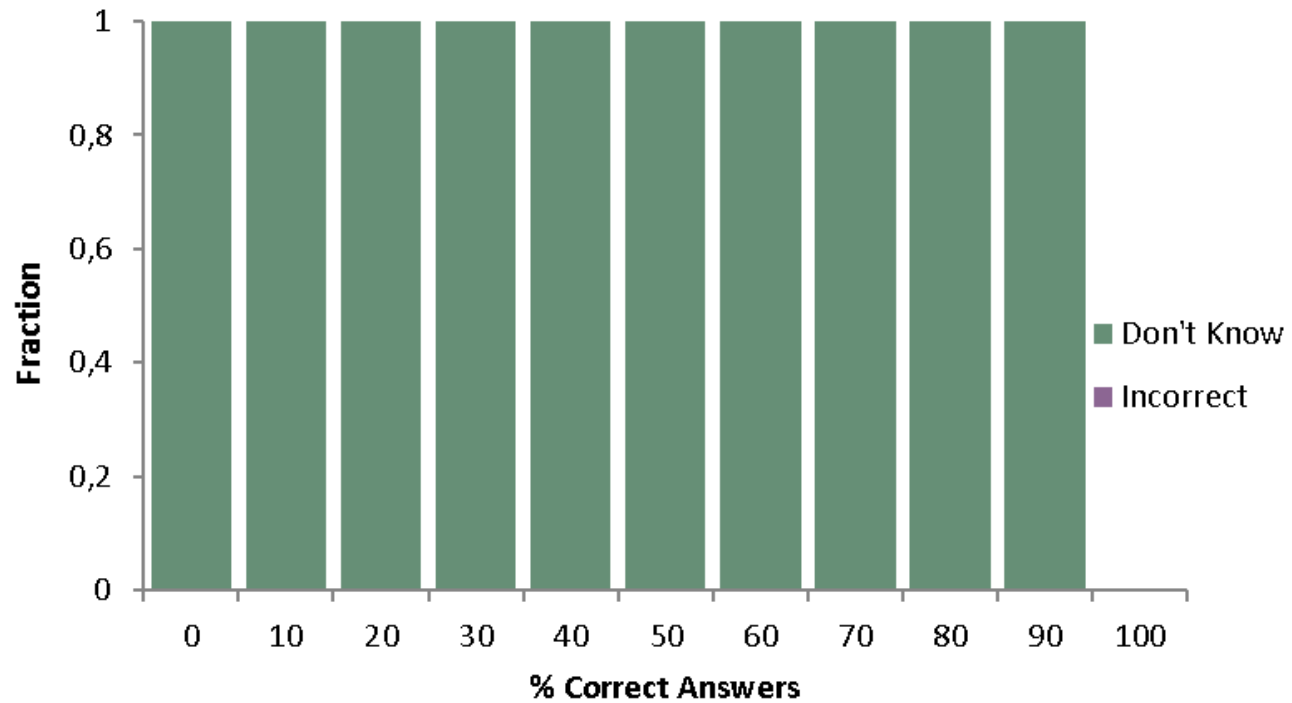
Overconfidence peaks at intermediate knowledge levels



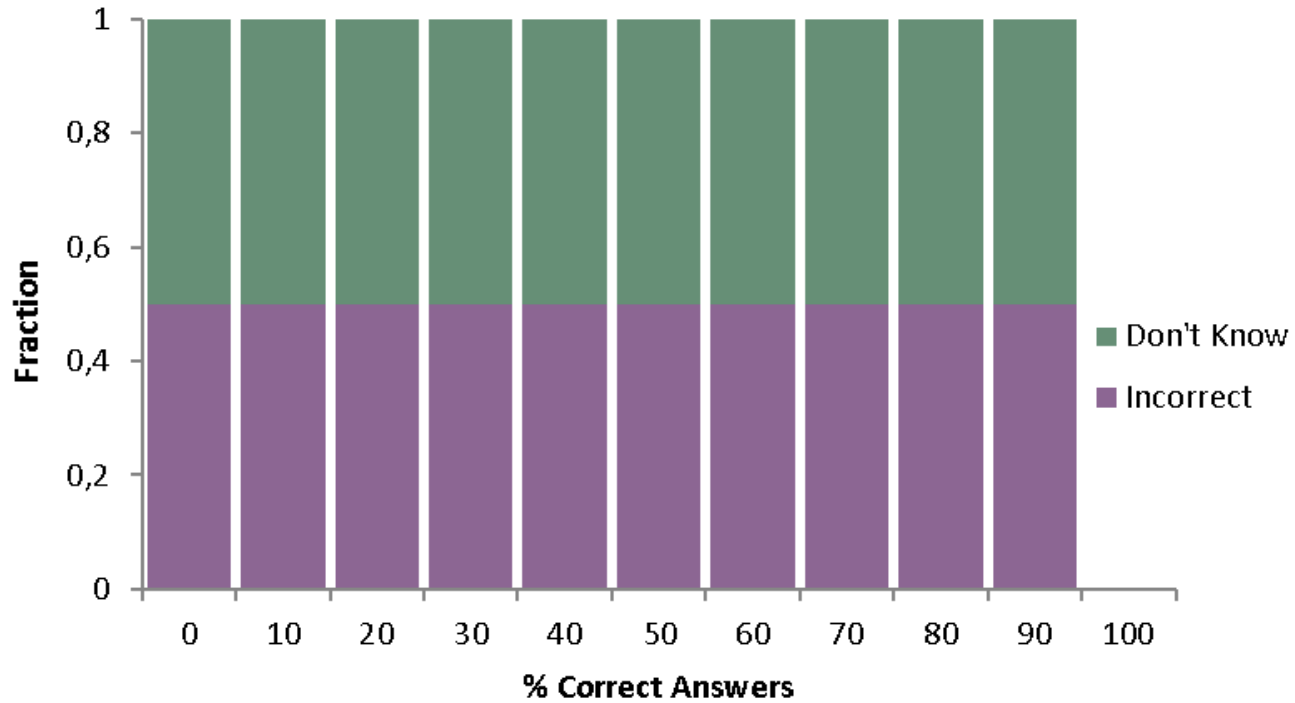


Overconfidence peaks at intermediate knowledge levels

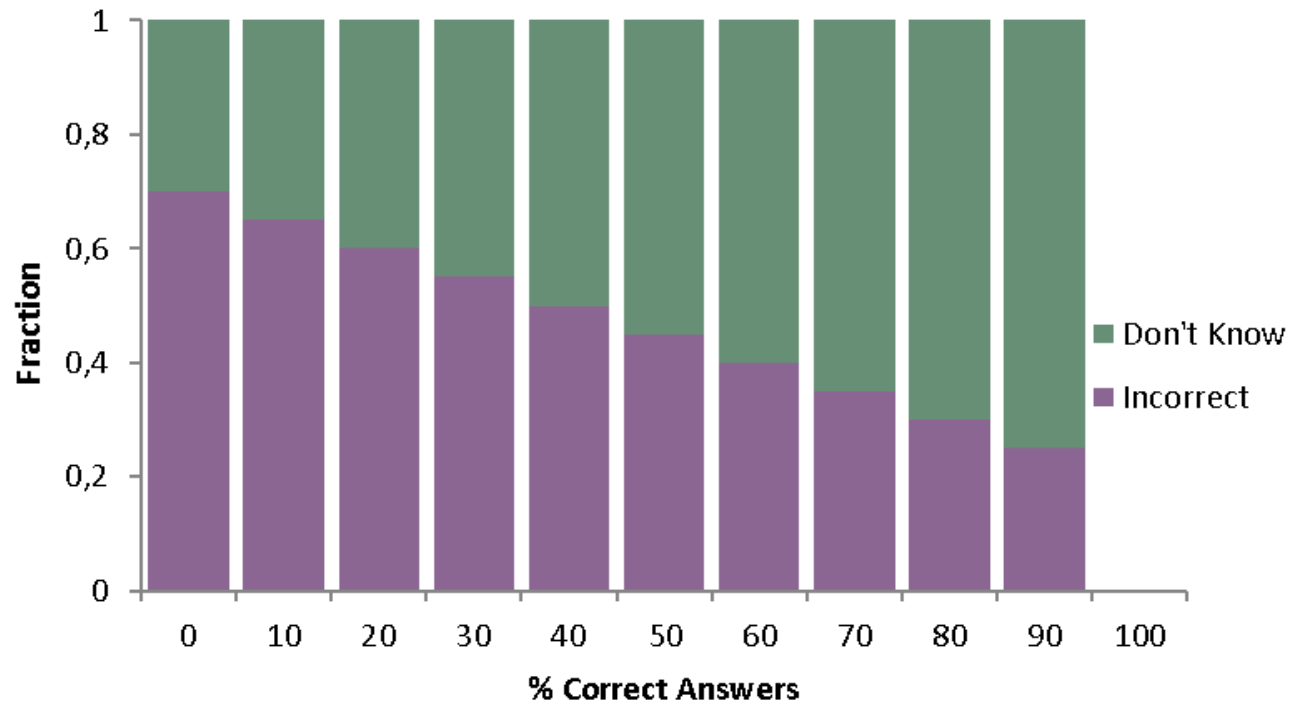
Perfect metacognition

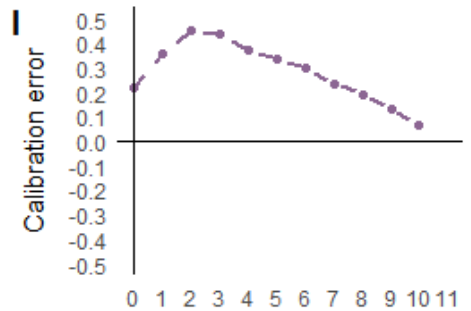
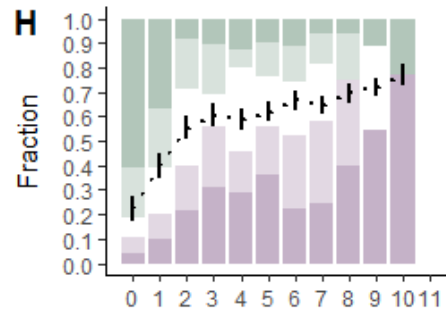
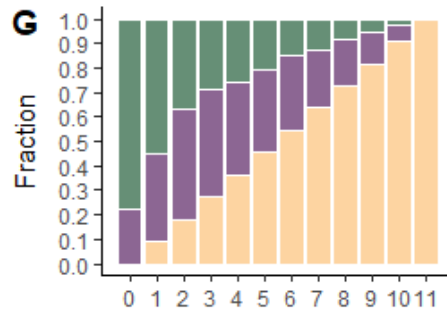
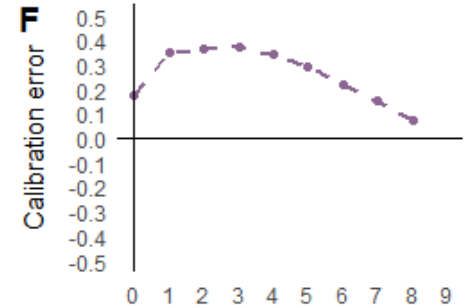
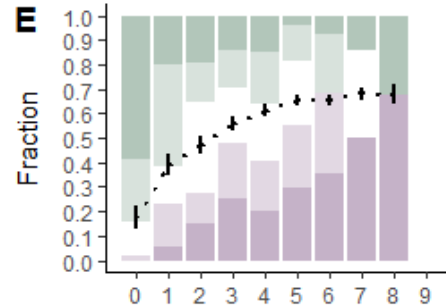
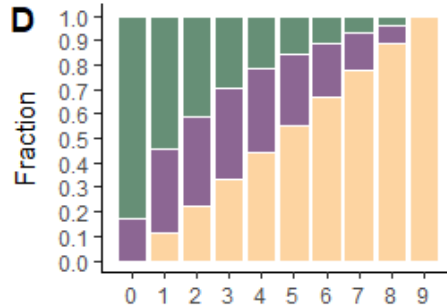
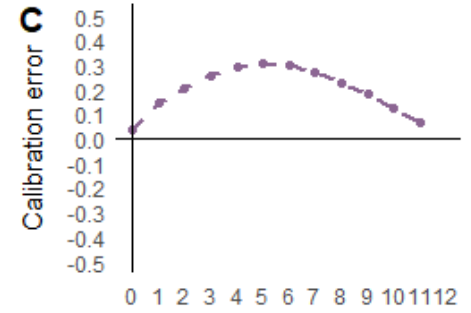
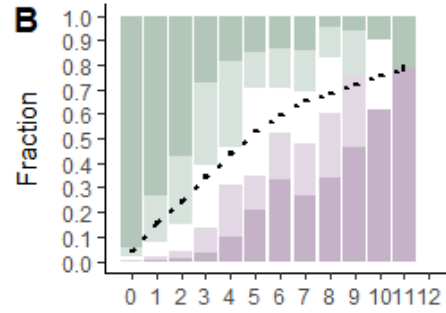
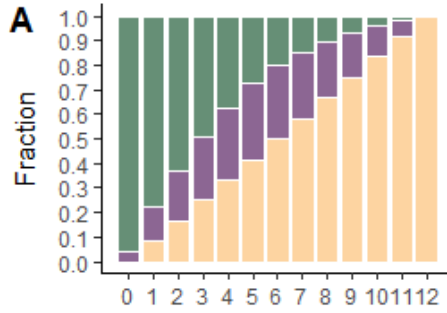


Random answering



DKE prediction



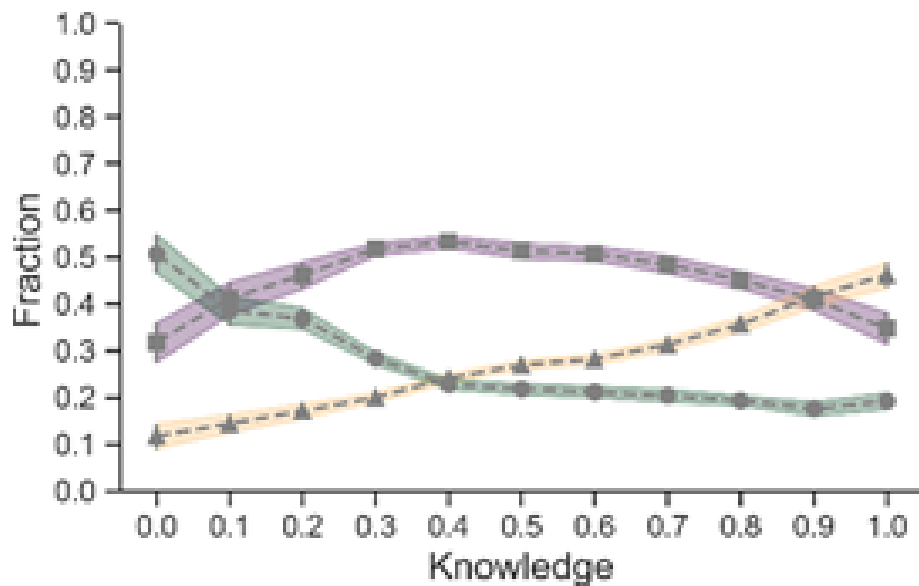


Correct Don't Know Incorrect Confidence

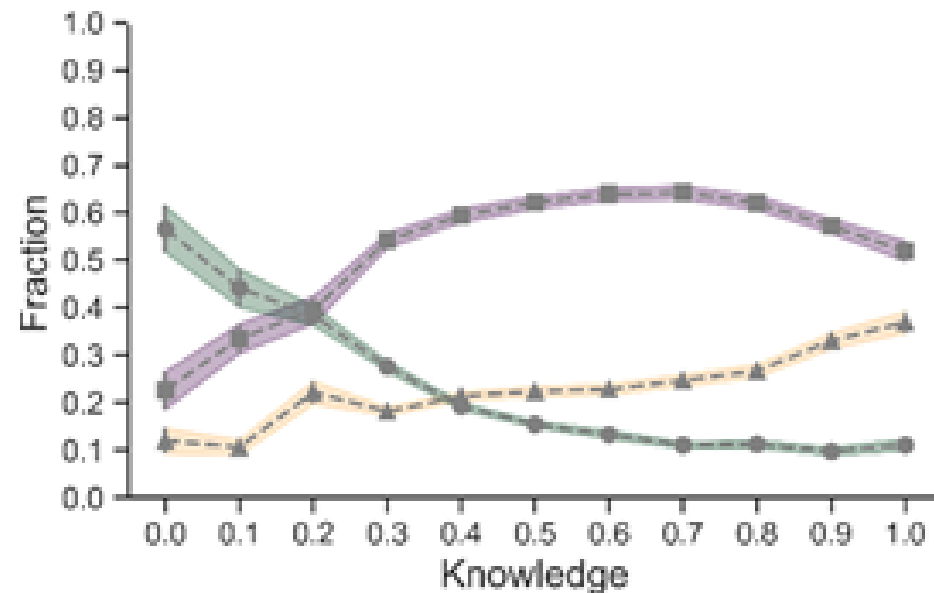


This has implications for attitudes towards science

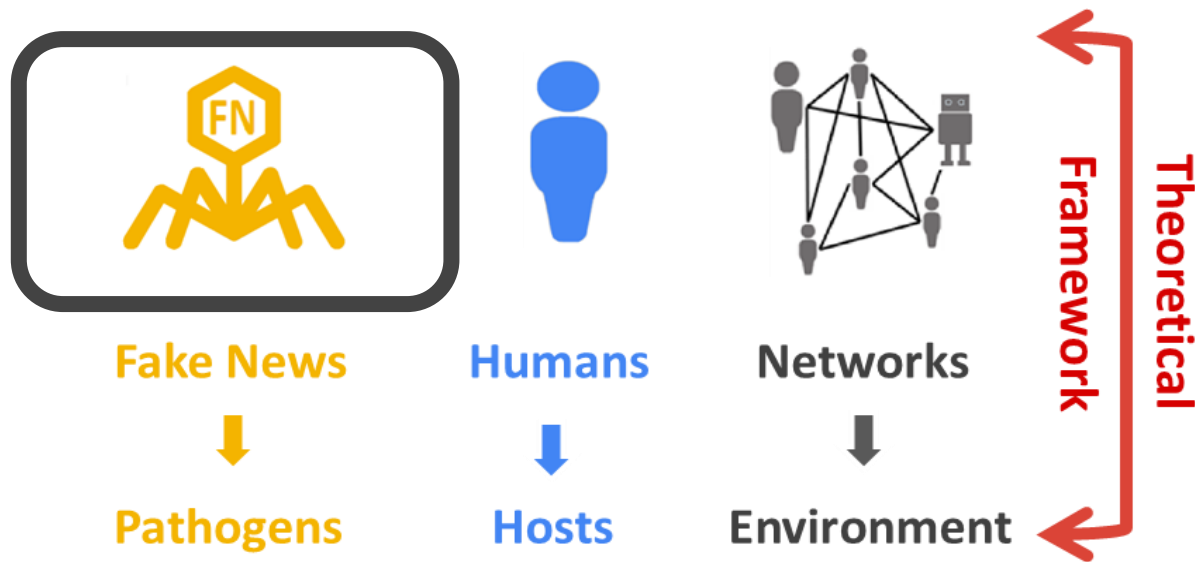
We depend too much on science
and not enough on faith



Because of their knowledge,
scientific researchers have a power
that makes them dangerous



Agree Disagree Don't Know



DISINFORMATION, TRACKING, AND BEHAVIORAL TARGETING

Newsweek repeatedly ran “sponsored” articles advertising colloidal silver as a way to prevent or treat COVID-19

One sponsored Newsweek article falsely claimed that “Black Americans can help protect themselves from COVID-19” by using a colloidal silver brand

WRITTEN BY [ERIC HANANOKI](#)

PUBLISHED 09/07/21 11:31 AM EDT

SHARE COMMENT



General Objective:

- a. **Identify** disinformation articles from known fact checkers
- b. **Measure** tracking and third-party content in disinformation websites
- c. **Test** whether differential tracking and content targeting occurs in disinformation websites

Building a fake news database

Main data sources:

- EU v. Disinfo;
- Google Fact Check tools + ClaimReview Schema;

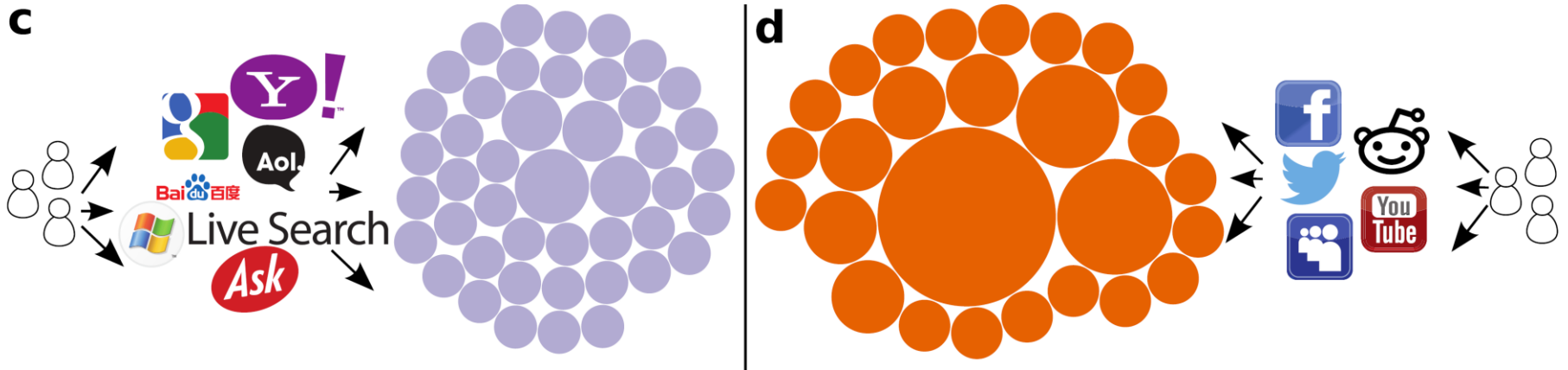
Challenges:

- Low adoption of ClaimReview;
- Mostly political topics;

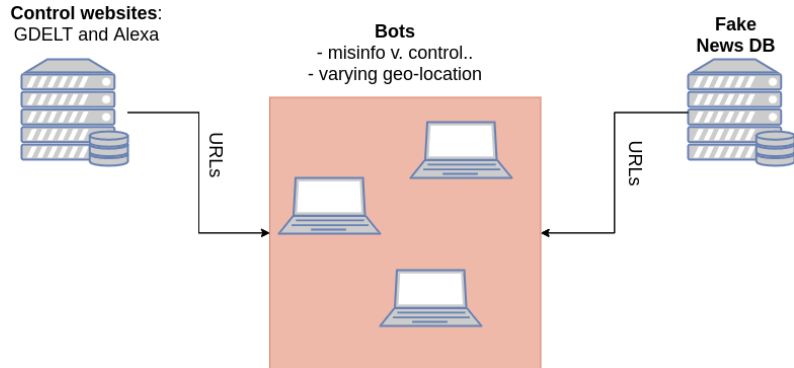


Measuring online social bubbles

Dimitar Nikolov, Diego F.M. Oliveira, Alessandro Flammini and
Filippo Menczer



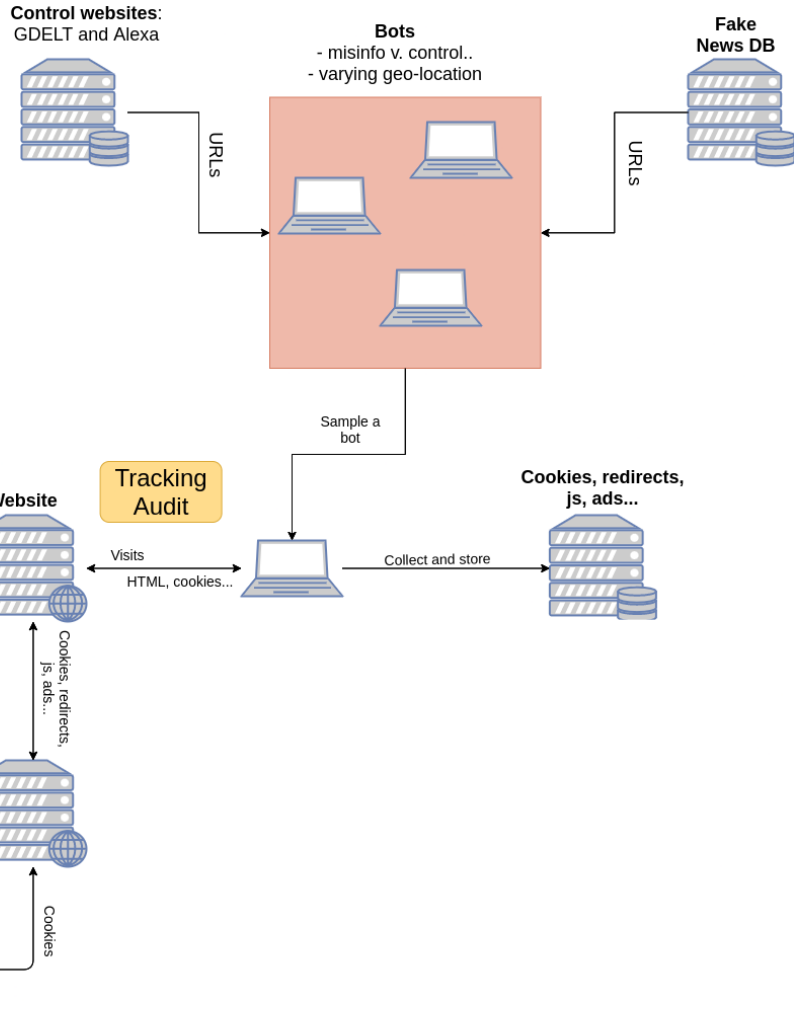
Adapted from <https://peerj.com/articles/cs-38/>



STEP1 : Create browsing HISTORIES

2 virtual agents (OpenWPM) visited 500 websites:

- **Treatment:** random sample of fake news websites from our DB
- **Control:** random sample stratified by top-level domain from Alexa's top 1M websites worldwide
- Same geo-location: USA, Miami
- Ran sequentially /parallel



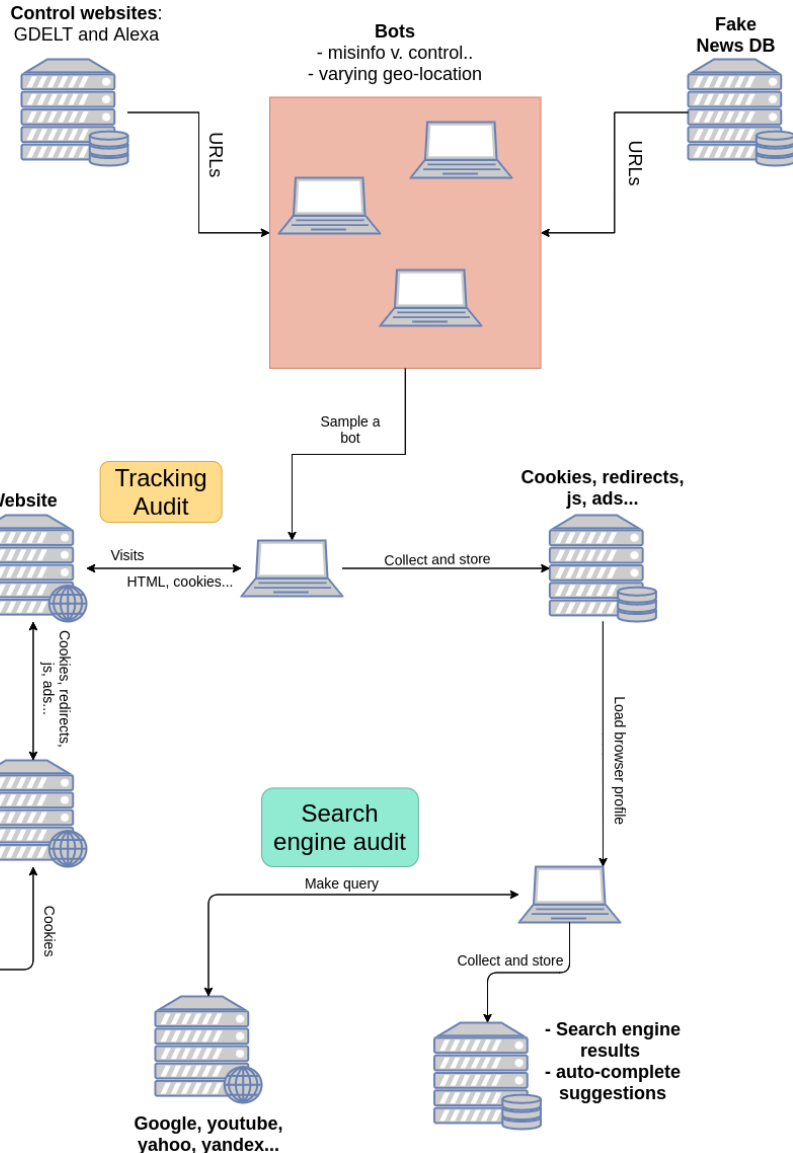
STEP1 : Create browsing HISTORIES

2 virtual agents (OpenWPM) visited 500 websites:

- **Treatment:** random sample of fake news websites from our DB
- **Control:** random sample stratified by top-level domain from Alexa's top 1M websites worldwide
- Same geo-location: USA, Miami
- Ran sequentially /parallel

STEP2 : TRACKING AUDIT

- Collect all HTTP-requests, redirect, responses, and javascript code along with metadata;
- Identify ads and store the raw html code;
- Identify third-parties;
- Collect cookies.



STEP1 : Create browsing HISTORIES

2 virtual agents (OpenWPM) visited 500 websites:

- **Treatment:** random sample of fake news websites from our DB
- **Control:** random sample stratified by top-level domain from Alexa's top 1M websites worldwide
- Same geo-location: USA, Miami
- Ran sequentially /parallel

STEP2 : TRACKING AUDIT

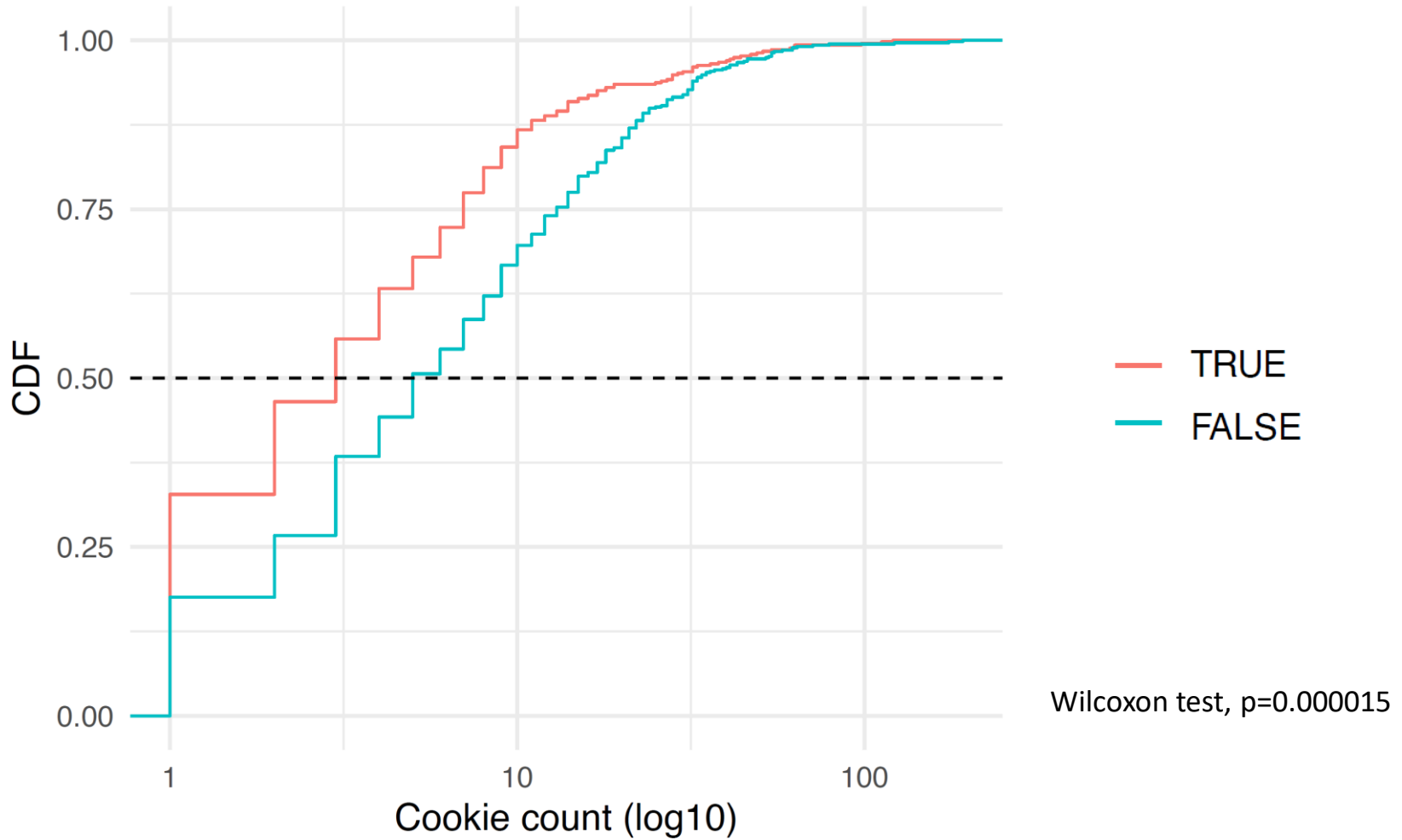
- Collect all HTTP-requests, redirect, responses, and javascript code along with metadata;
- Identify ads and store the raw html code;
- Identify third-parties;
- Collect cookies.

STEP3 : SEARCH ENGINE AUDIT

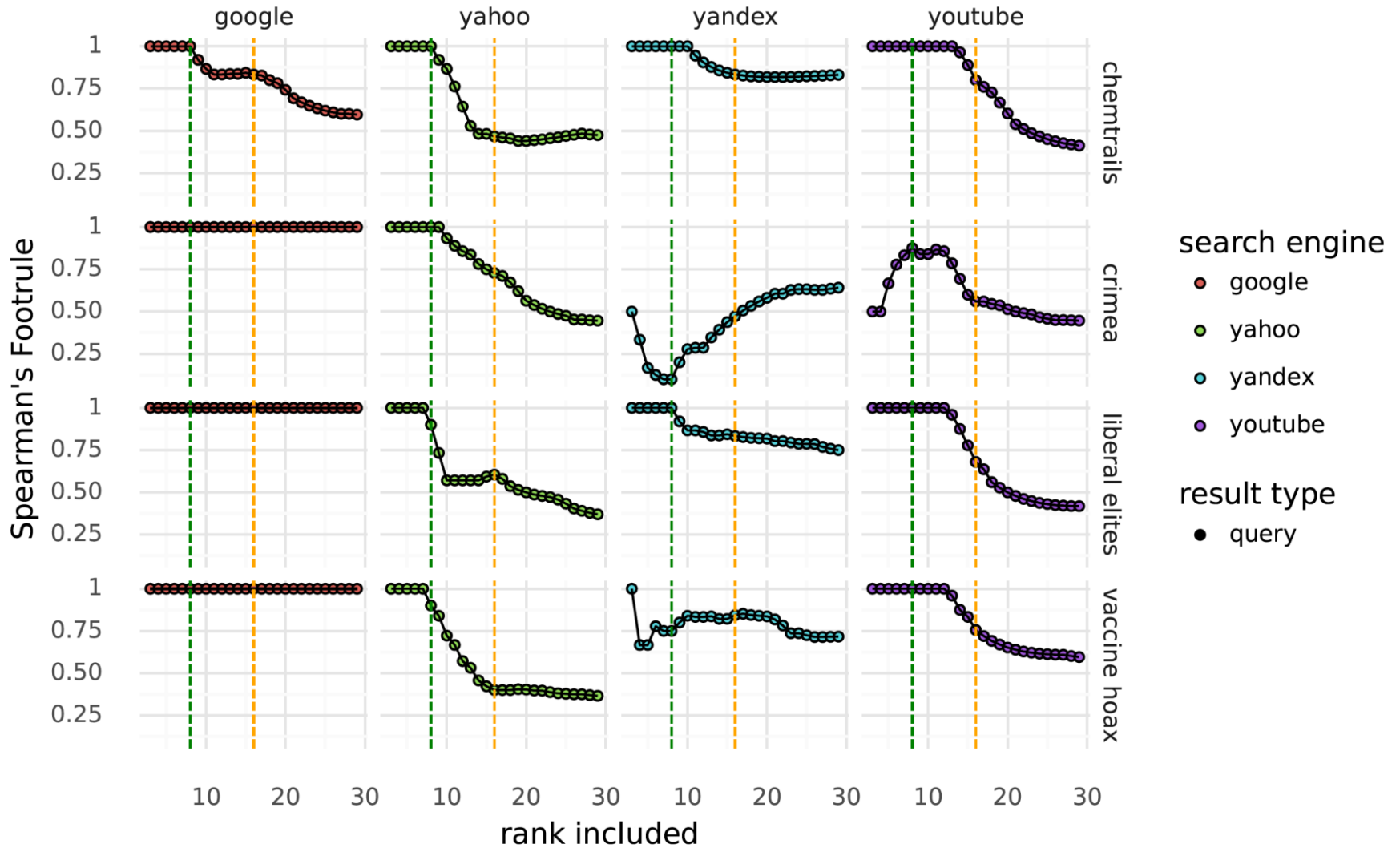
Each bot visited 6 search engines and typed the same queries. We collect:

- Auto-complete suggestions (where available)
- Ranked results of the first 10 pages/scrolls

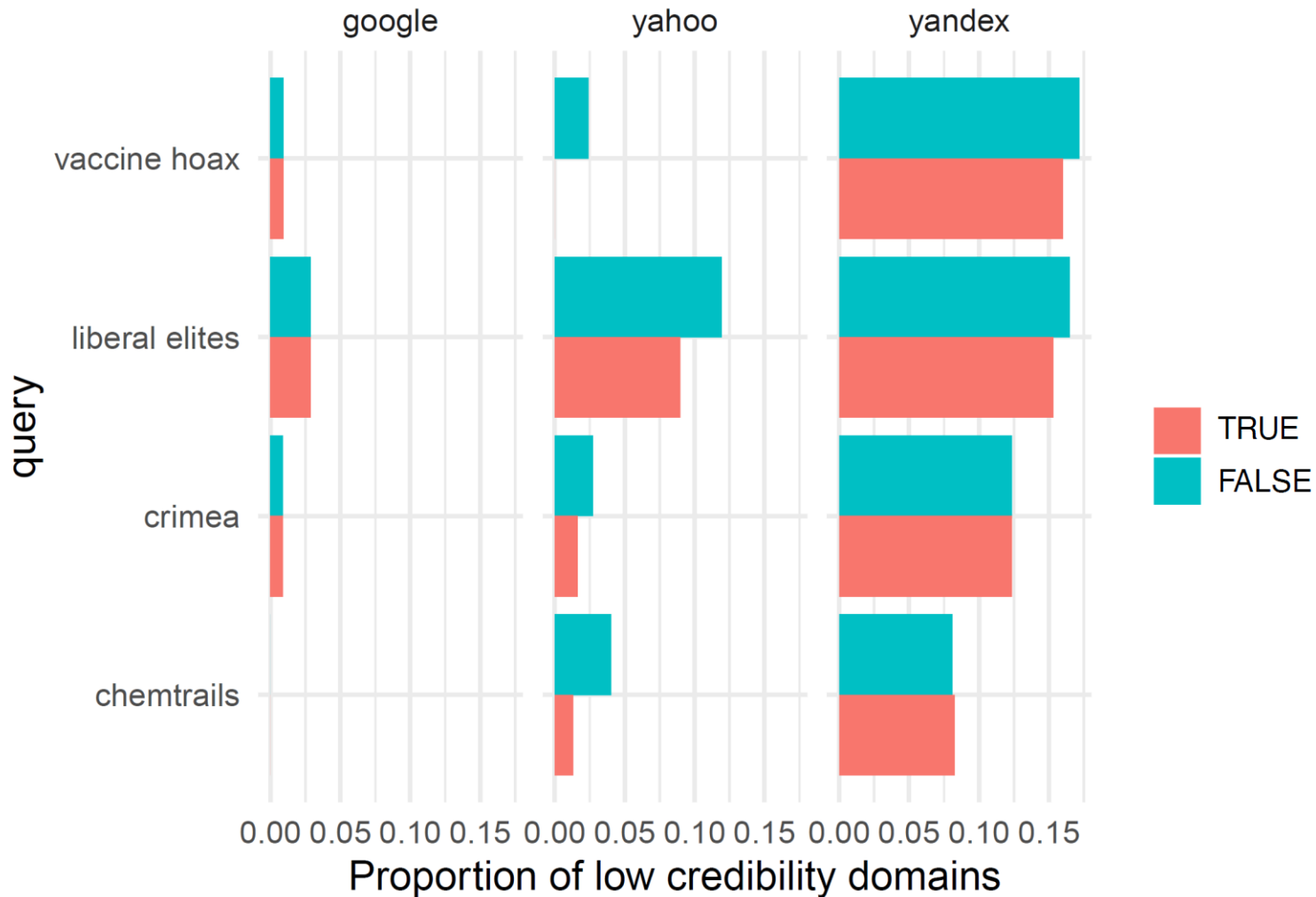
Disinformation websites tend to place more cookies



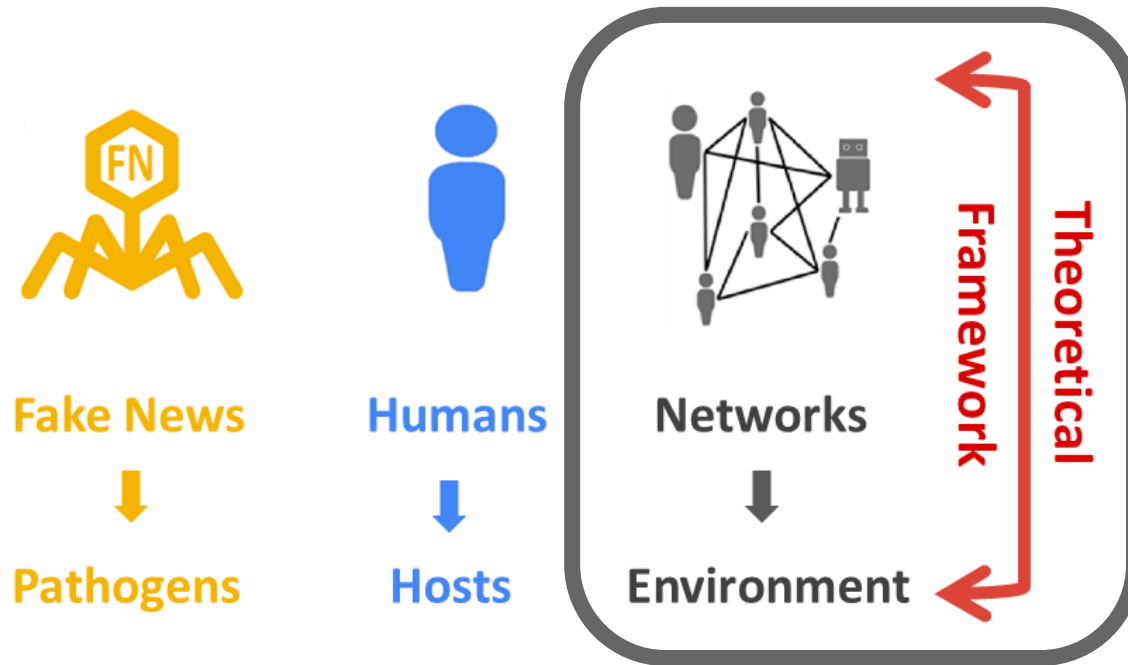
Profiles influence search results (both query and auto-complete)



Profiles that visited disinformation websites are directed to poorer quality content

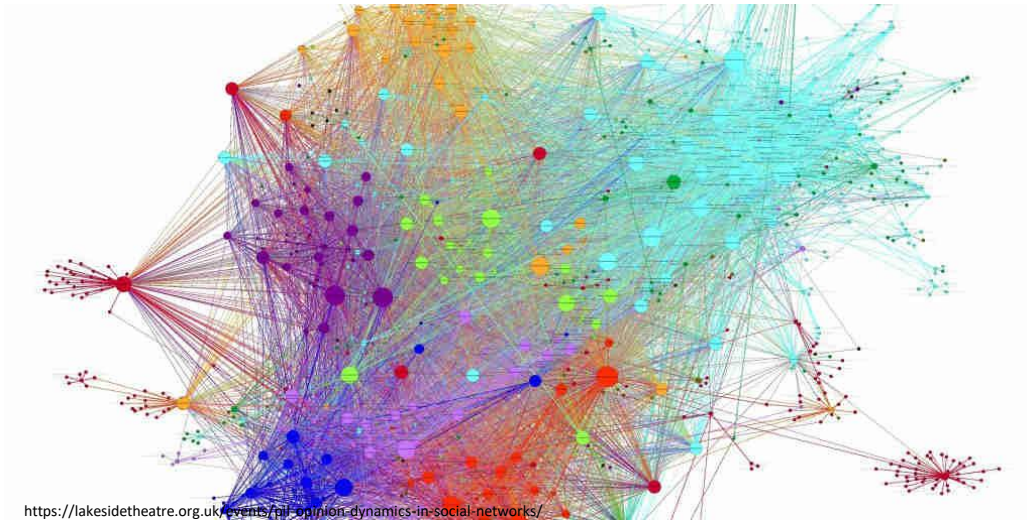


1. Designed and implemented an experiment to audit tracking and its influence on search engine results;
2. Disinformation websites tend to place more cookies;
3. Profiles/ browsing history influence search results (both query and auto-complete);
4. Some search engines might be more worrisome than others;
5. Profiles that visited disinformation websites are directed to poorer quality content;
 - Running bots in parallel
 - Increasing geo-locations
 - Analyzing ads (from Google)



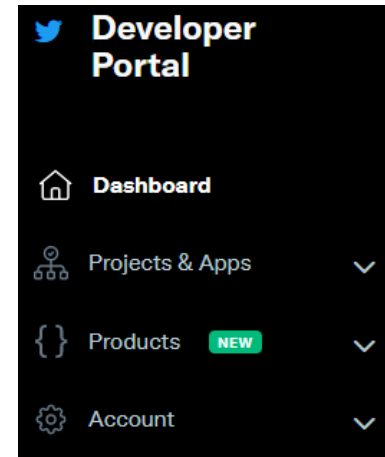
INFORMATION FLOW ON NETWORKS

Paulo Almeida, Pedro Duarte, Lília Perfeito



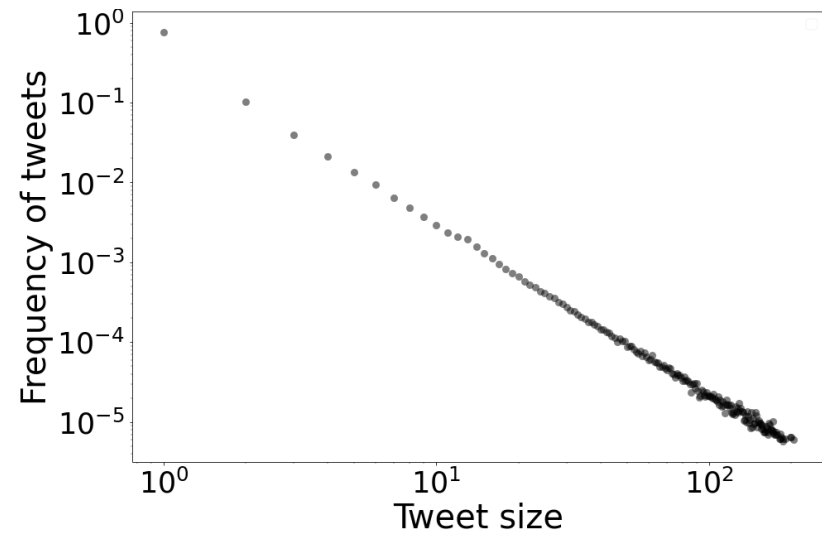
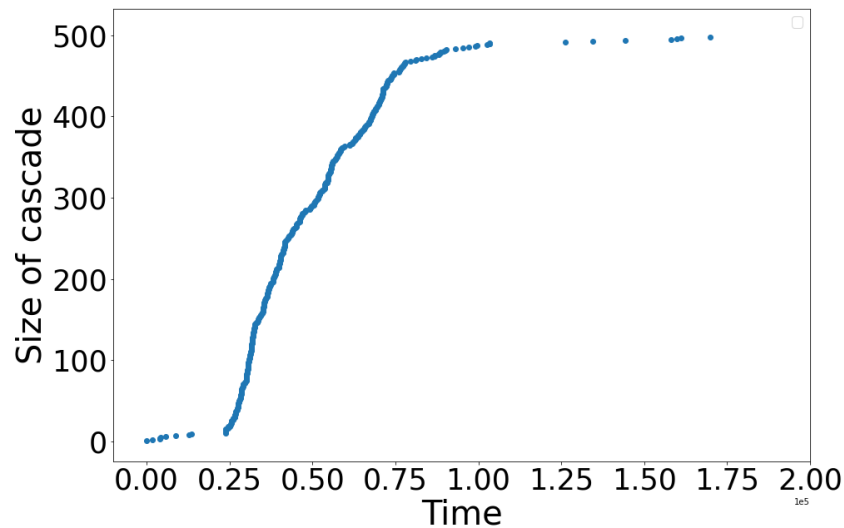
Questions

- Can we estimate the fitness of a tweet?
- What is the contribution of the network?
- What is the contribution of individual heterogeneity?
- Why do some tweets spread further than others?



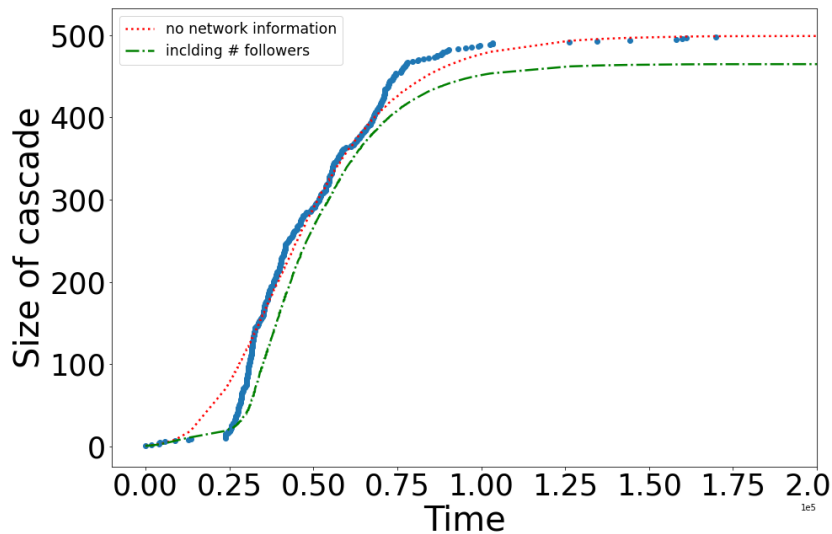
Data:

- Tweets and user profiles (from Twitter API)
- Feed into database (being built)
- Group tweets by content (cascades)



Tools:

- Model process:
 - Analytical model - diffusion, population dynamics, epidemiological, etc.
 - Simulations
- Fit observations and compare parameters
- Estimate the effect of different features of tweets, network and users



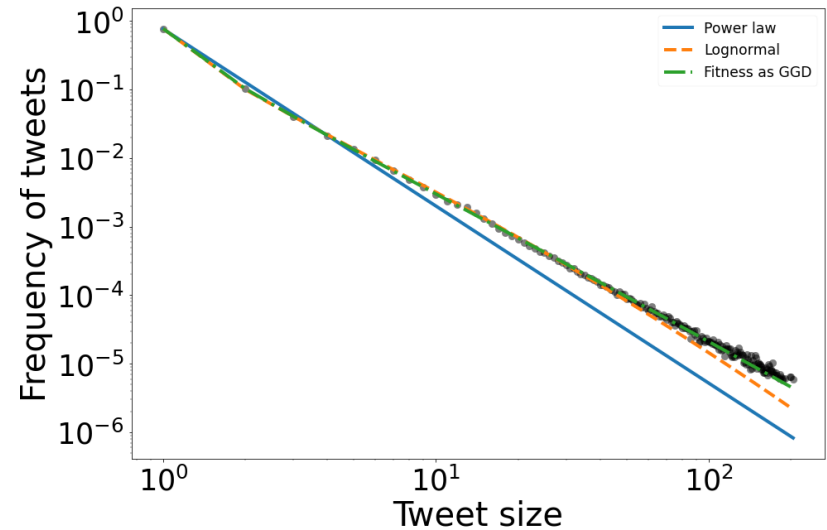
We do not require the network to explain the cascade size distribution ...

$$N(t) = e^{\frac{a}{s}} e^{a \cdot e^{-s \cdot t}} \quad f_N(N) = f_{a/g}(\log(N)) \cdot \frac{1}{N}$$

... but it is there. What is the contribution of the network?

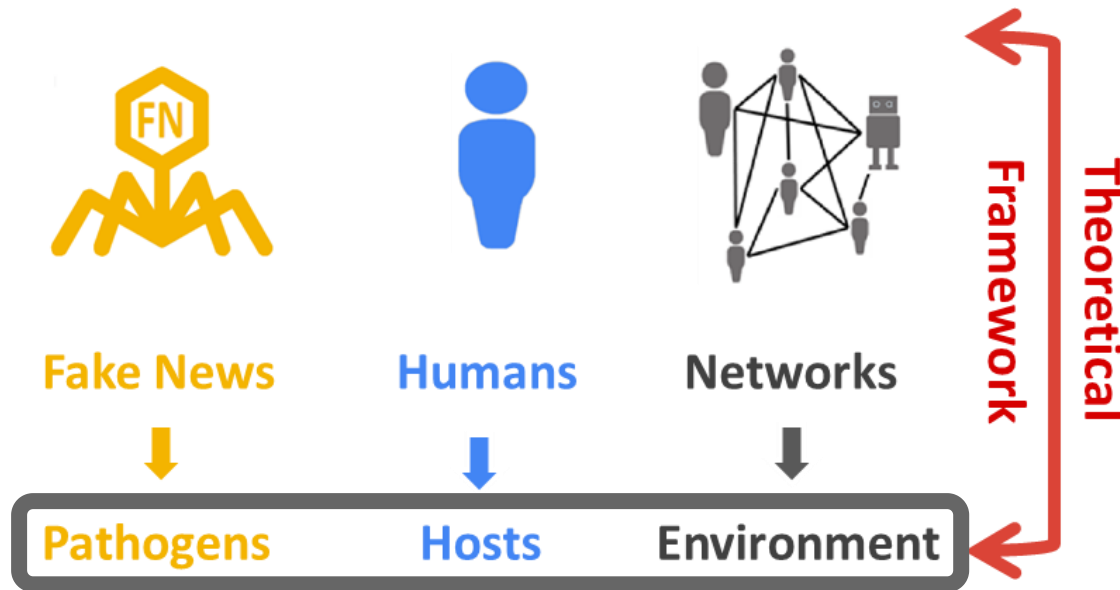
$$\Delta N = N \cdot a \cdot e^{-s \cdot t} \cdot f(t)$$

Simulations in development ...



To be added to the framework:

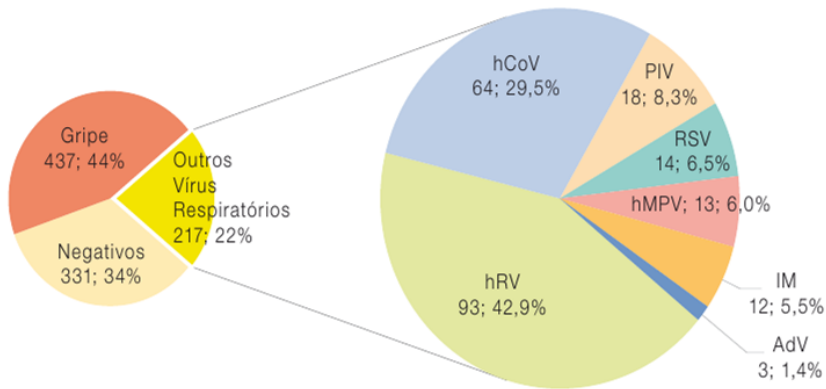
- Classification of tweets in topics, sentiment, fake&real news, etc.
- Heterogeneity in users – demographic, psychological properties, etc.



DYNAMICS OF NON-INFLUENZA RESPIRATORY VIRUSES (NIRVS)

Flu season & NIRVs

n=985



INSA National flu surveillance program. Season 2018/2019

General Objectives:

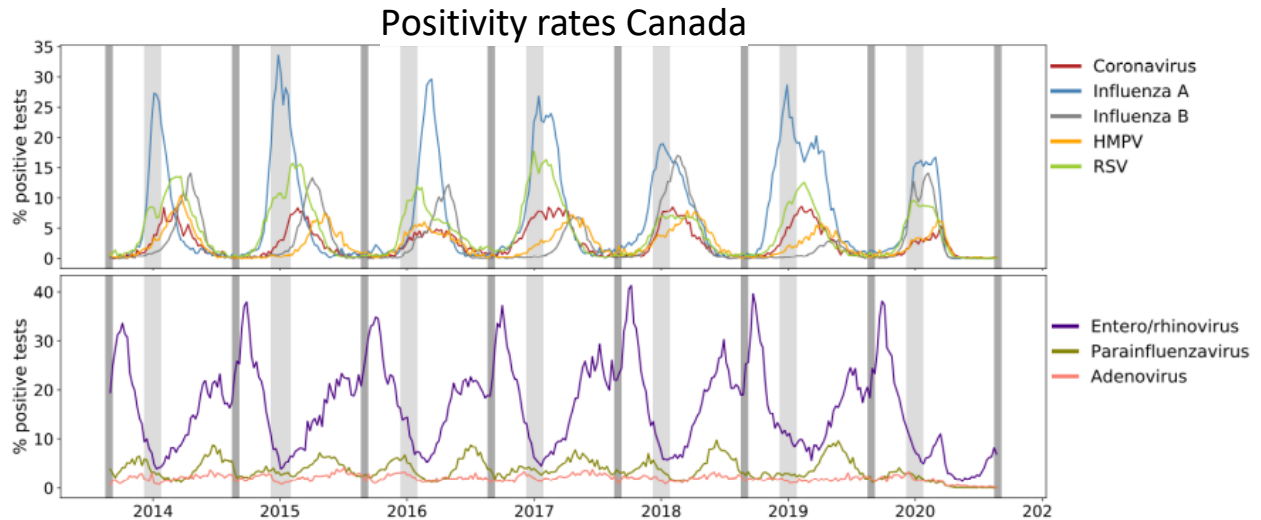
- a. **Identify** different epidemic dynamics.
- b. **Model** dynamics:
 - **Estimate** the effect of weather or vaccination.
- c. **Test** whether different dynamics lead to differences in reported symptoms.
- d. **Improve** nowcasting models for the flu-NIRVs season

Data:

- Epidemiological data: NIRVs positivity rates.
- Climate data (weighted for population).
- Symptoms data: Twitter, Google, Influenza Net.

Country	Epidemiological data					
	Frequency	From	Seasons	National data	Regional data	Format
USA	Weekly	16/17	4	hCoVs	hCoVs	% Positive test
				hMPV	hMPV	
				RSV	RSV	
				Adenovirus		
				Parainfluenza		
Canada	Weekly	13/14	7	hCoV	hCoV	% Positive test
				hMPV	hMPV	
				RSV	RSV	
				Adenovirus	Adenovirus	
				Parainfluenza	Parainfluenza	
				Enterorhino virus	Enterorhino virus	

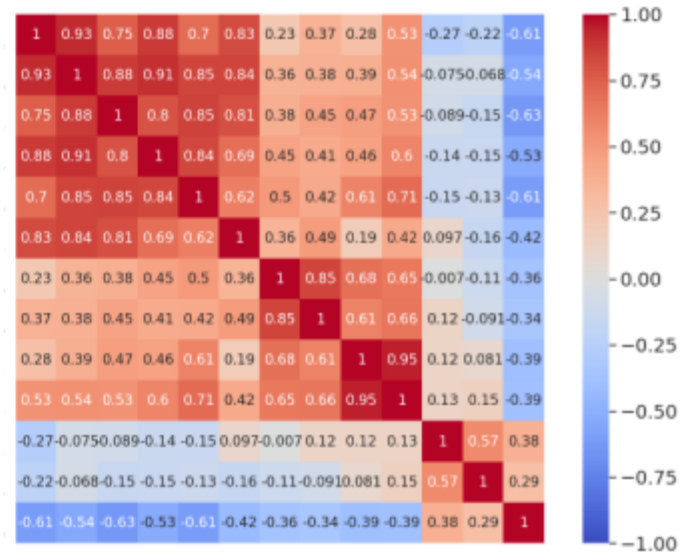
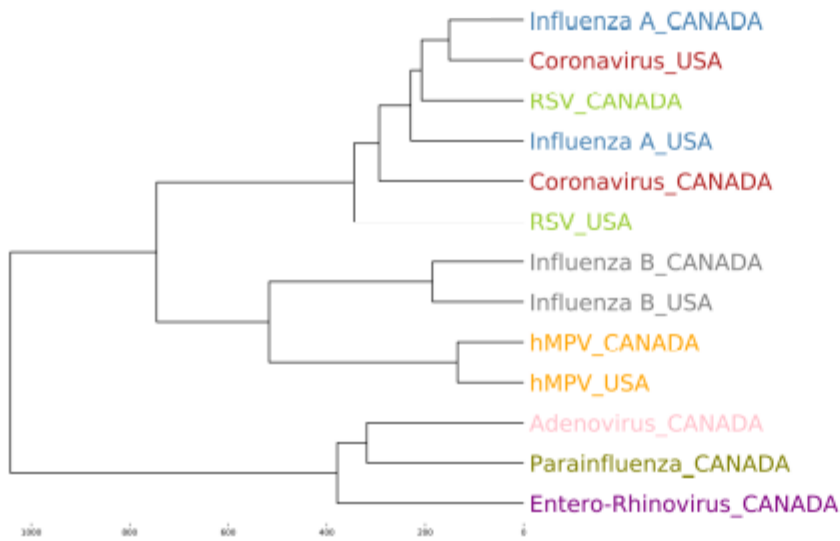
Influenza-like



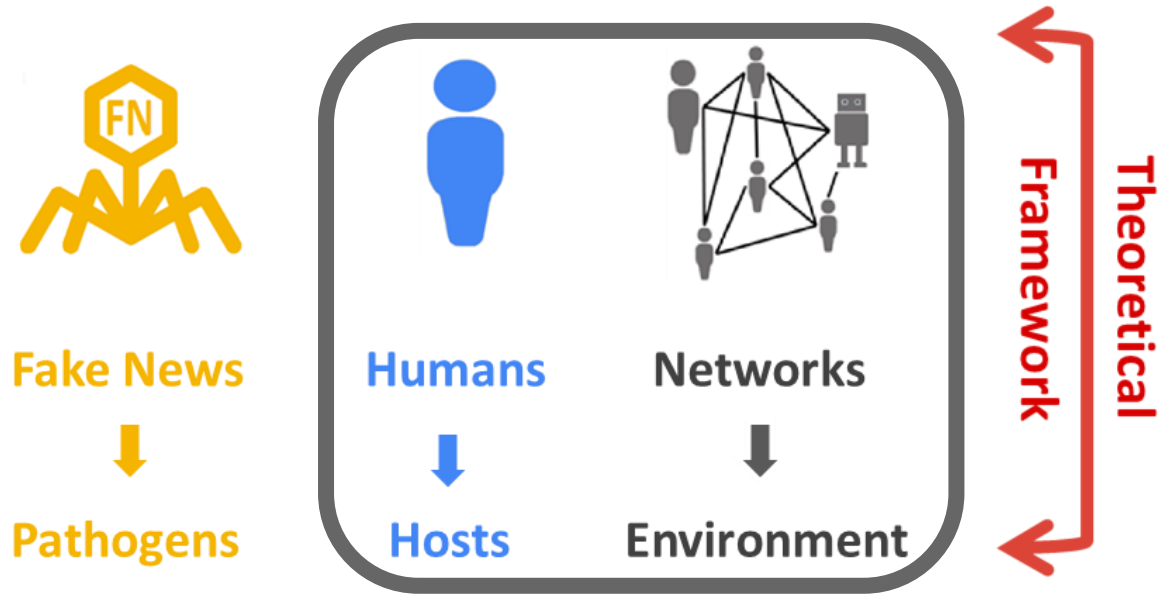
Non influenza-like

1. Pattern identification: NIRVs dynamics

Positivity rates normalized Canada & USA



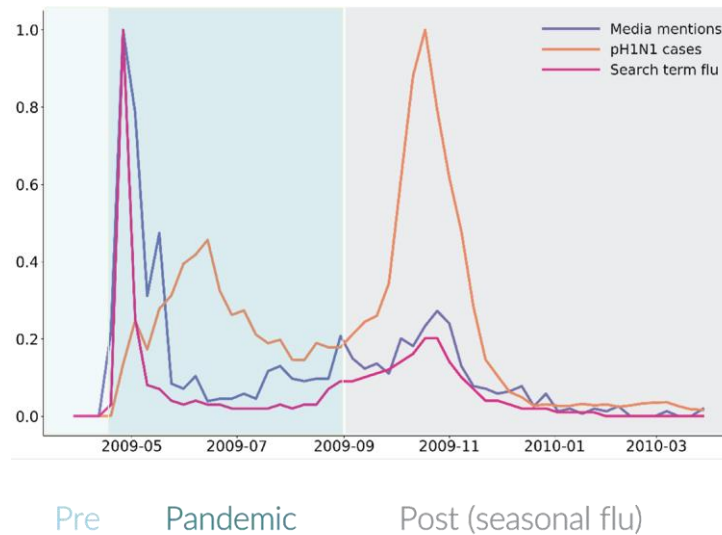
1. SIR epidemiological models



USING ONLINE BEHAVIOUR TO TRACK DISEASES

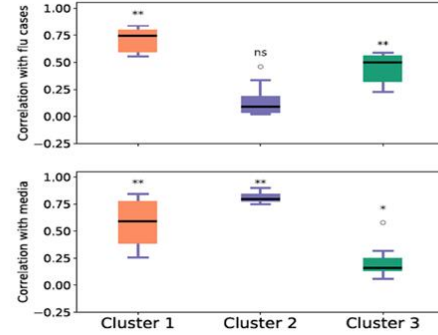
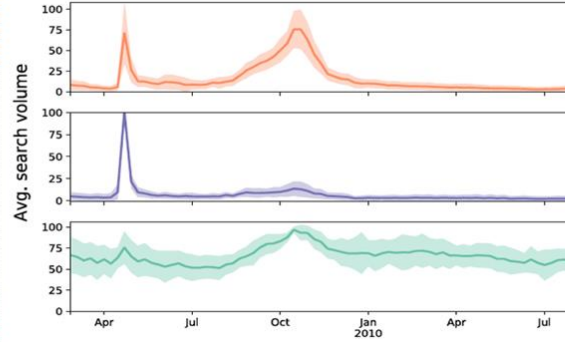
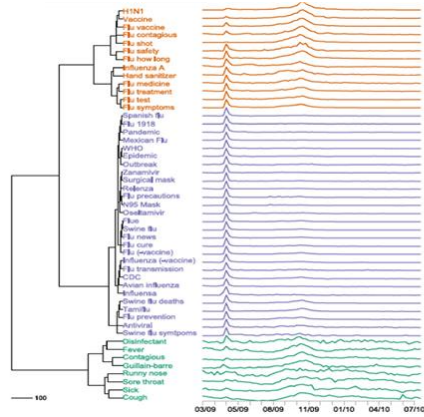
David Almeida, Joao Loureiro, Sara Mesquita, Lília Perfeito, Cláudio Haupt Vieira

2009 Flu Pandemic - USA

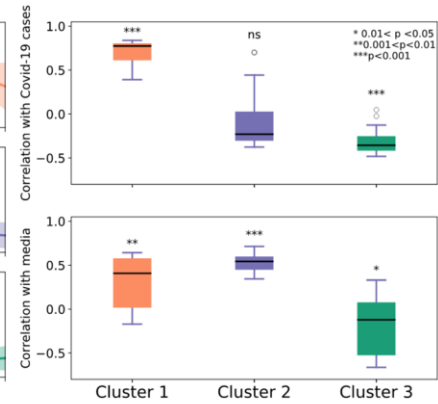
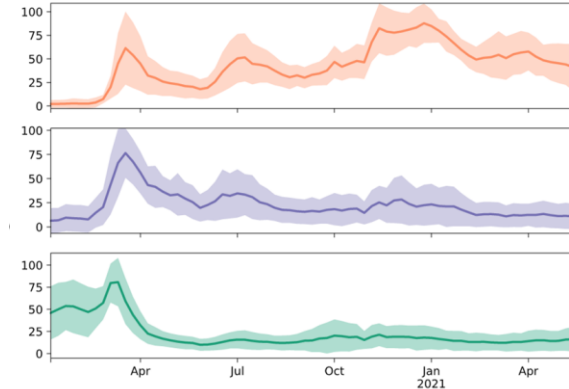
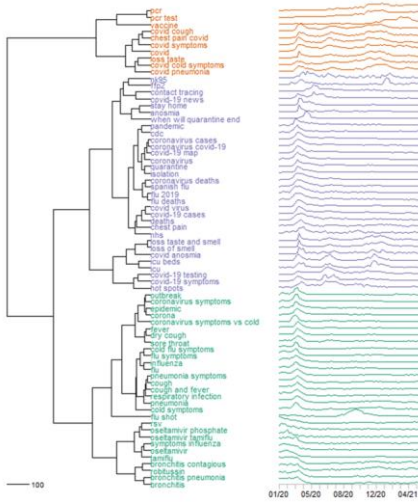


Google searches for “flu” might be driven more by the media than by the disease.

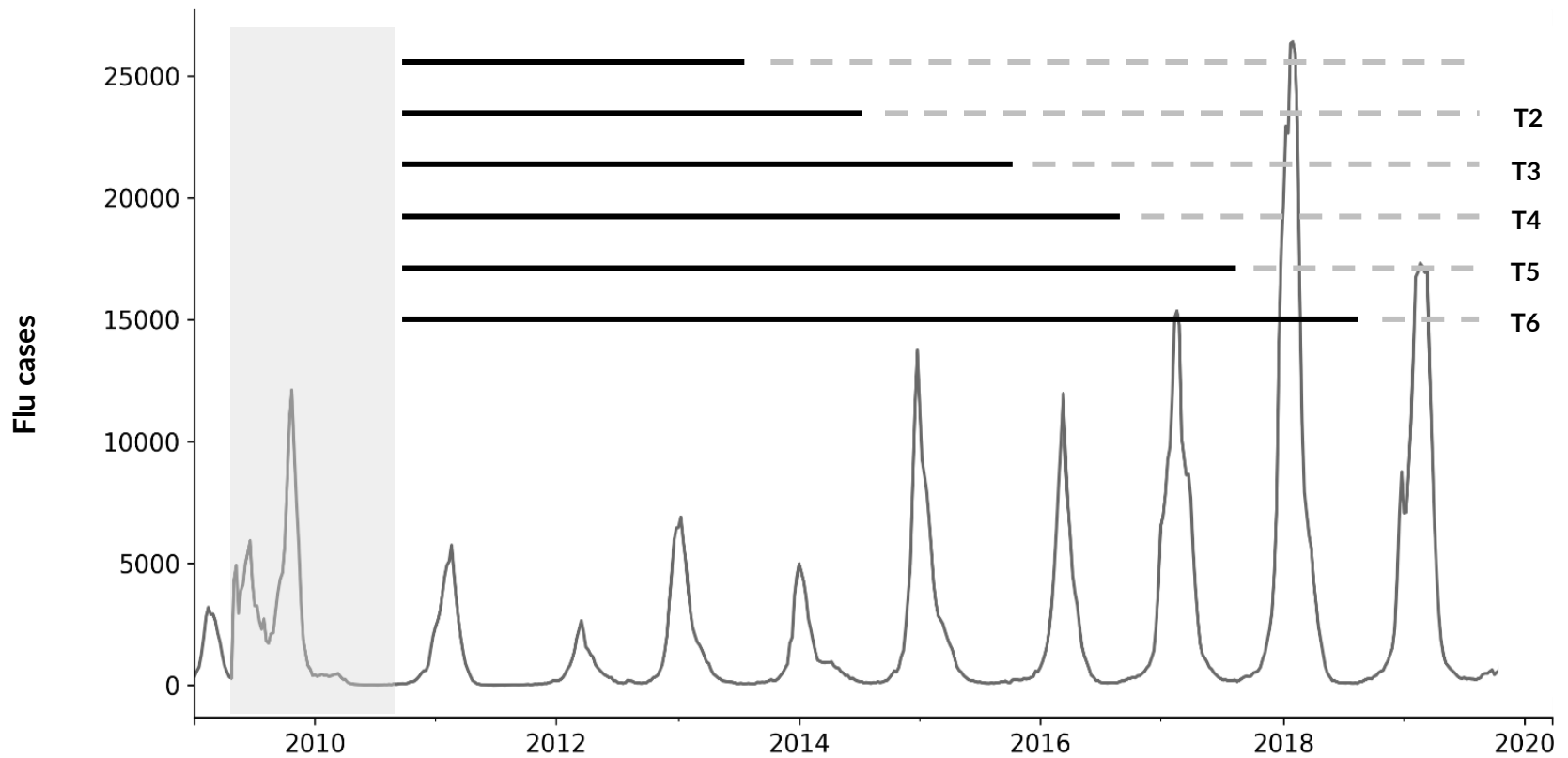
2009 Flu Pandemic



Covid-19 Pandemic

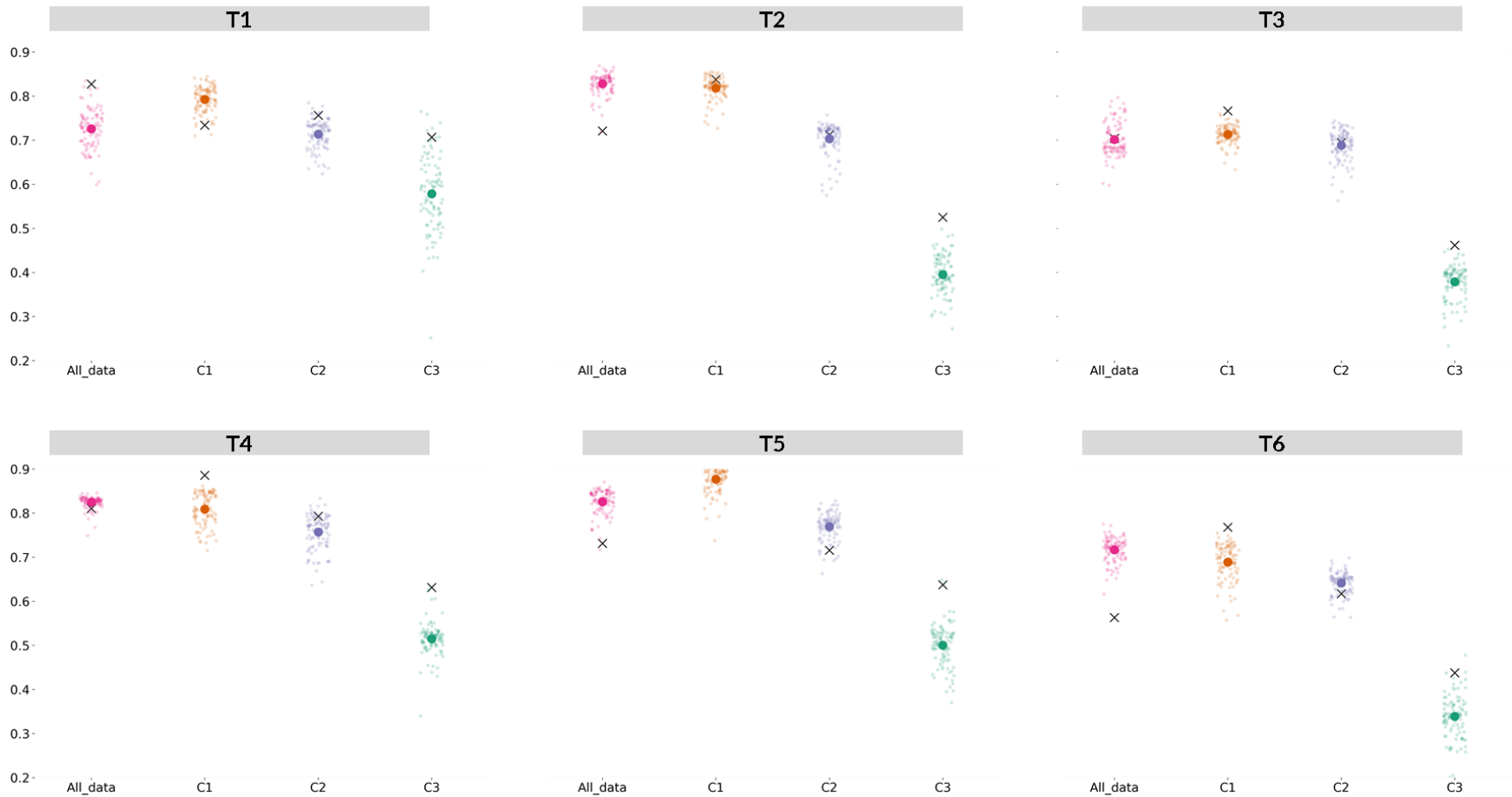


Can we use it to improve prediction?

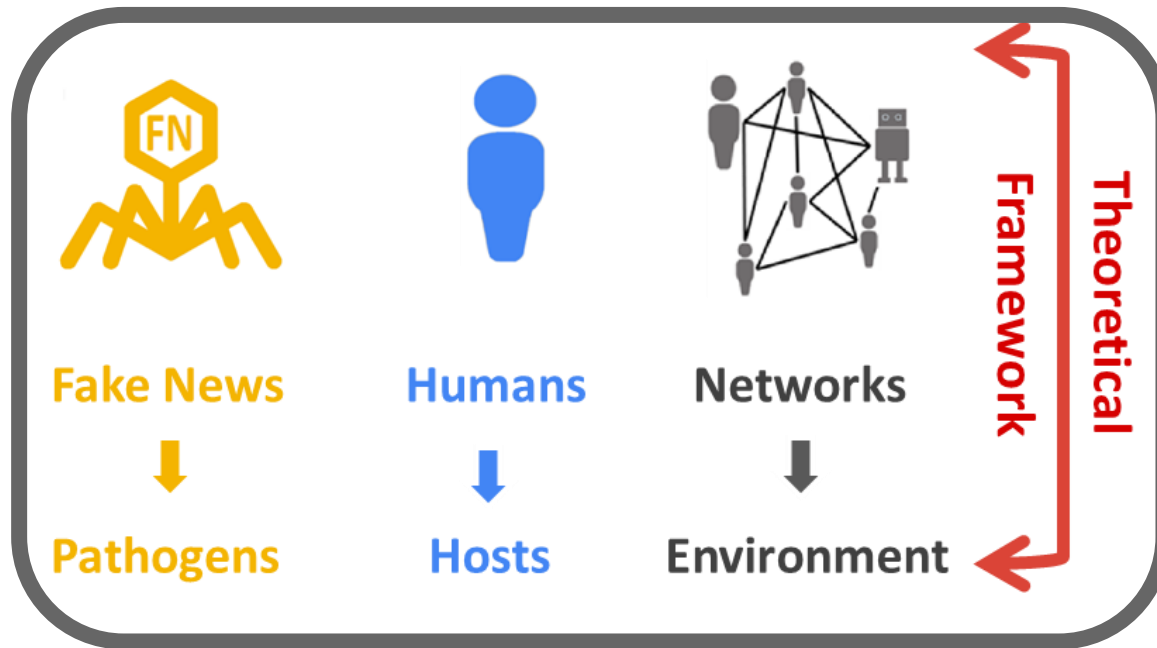


X = Linear Regression results
● = Random Forest

2009 Flu pandemic



Curating the data improves forecasting, over using all data



PROTECTING PERSONAL DATA WITH MULTI-PARTY COMPUTATION

Paulo Almeida, Alex Davidson (former member)

Research with Personal Data

- Privacy
 - GDPR
- Ethics
 - Profiling
 - Societal harm
- Confidentiality
 - Economic interests

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

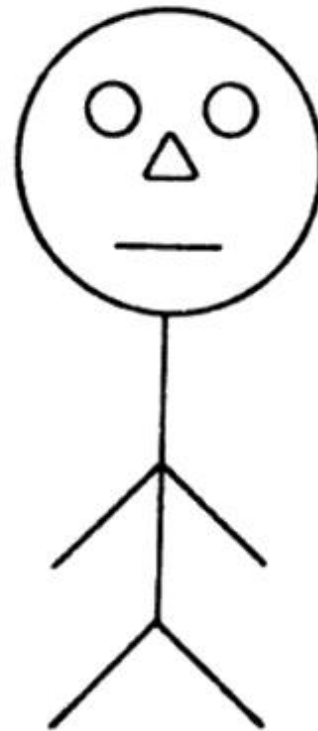
como é que se **beija**
como é que se **diz eu te amo**
como é que se **beija de lingua**
como é que se **engravidar**
como é que se **beija na boca**
como é que se **escreve**
como é que se **beija pela primeira vez**
como é que se **faz um facebook**
como é que se **faz um relatório**
como é que se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

pourquoi je suis

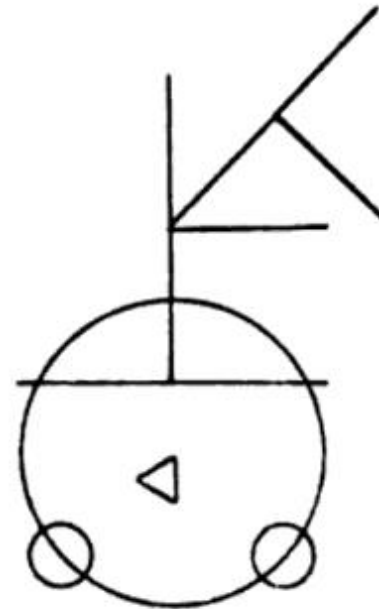
pourquoi je suis **moche**
pourquoi je suis **triste**
pourquoi je suis **toujours fatigué**
pourquoi je suis **célibataire**
pourquoi je suis **toujours célibataire**
pourquoi je suis **devenu rebelle pdf**
pourquoi je suis **seule**
pourquoi je suis **toujours fatiguée**
pourquoi je suis **jalouse**
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 gâteau**
comment faire **du caramel**
comment faire **de la glue**
comment faire **du pain**



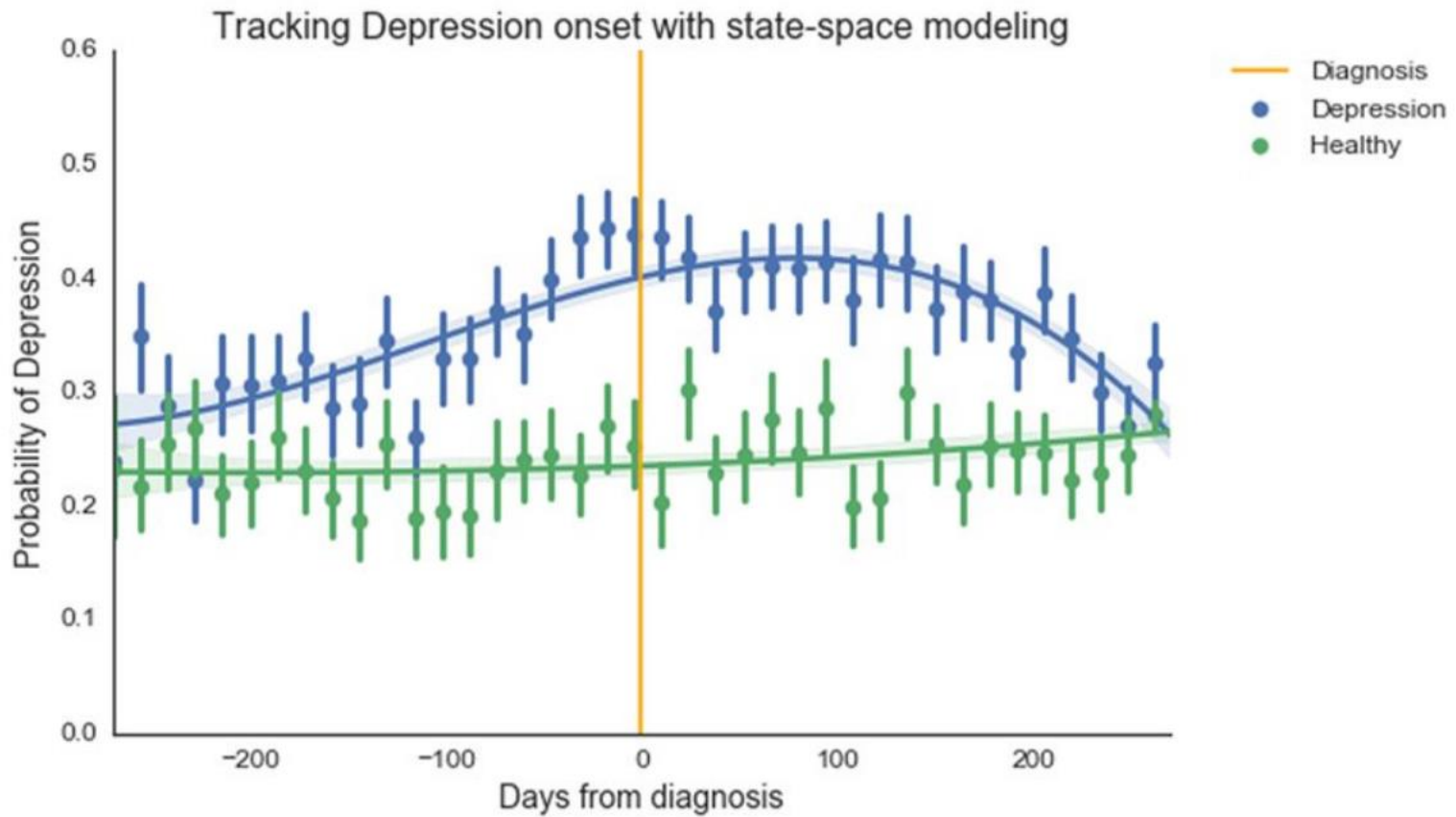
Intended
FACEMAN



Picture drawn by bugged
FACEMAN program

Figure 9. Stick men drawn by LOGO programs (from Sussman, 1973)

Thinking: Readings in Cognitive Science, (1978) edited by P. N. Johnson-Laird, P. C. Wason, page 26.

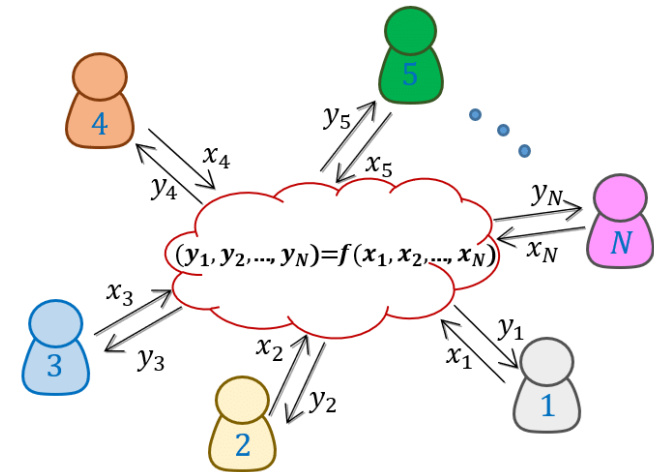


Solutions

- Homomorphic encryption
- Federated learning
- Differential Privacy
- Synthetic Data
- Multi-party Computation

Multi-party Computation

- Computation of combined data
 - Inputs remain private
 - Processing can be local
 - Different Algorithms (Shamir, Yao)
- Use cases
 - Commodity price setting trusted by all parties
 - Gender gap wage study without revealing salaries
- MPC at SPAC
 - MP-SPDZ library
 - Proof of concept: Docker containers
 - Outreach



From the Industrial Revolution to the Digital Revolution



First

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



1800

1784: First mechanical loom



Second

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



1900

1870: First assembly line

Third

IT systems automate production lines further.



2000

1969: First programmable logic controller

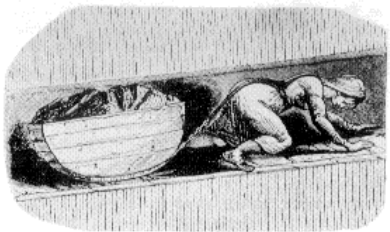
Fourth

IoT and cloud technology automate complex tasks.



Today

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

1833 UK
Factory Act

1938 US Federal Fair
Labor Standards Act

1973: ILO Conference

1800

1900

2000

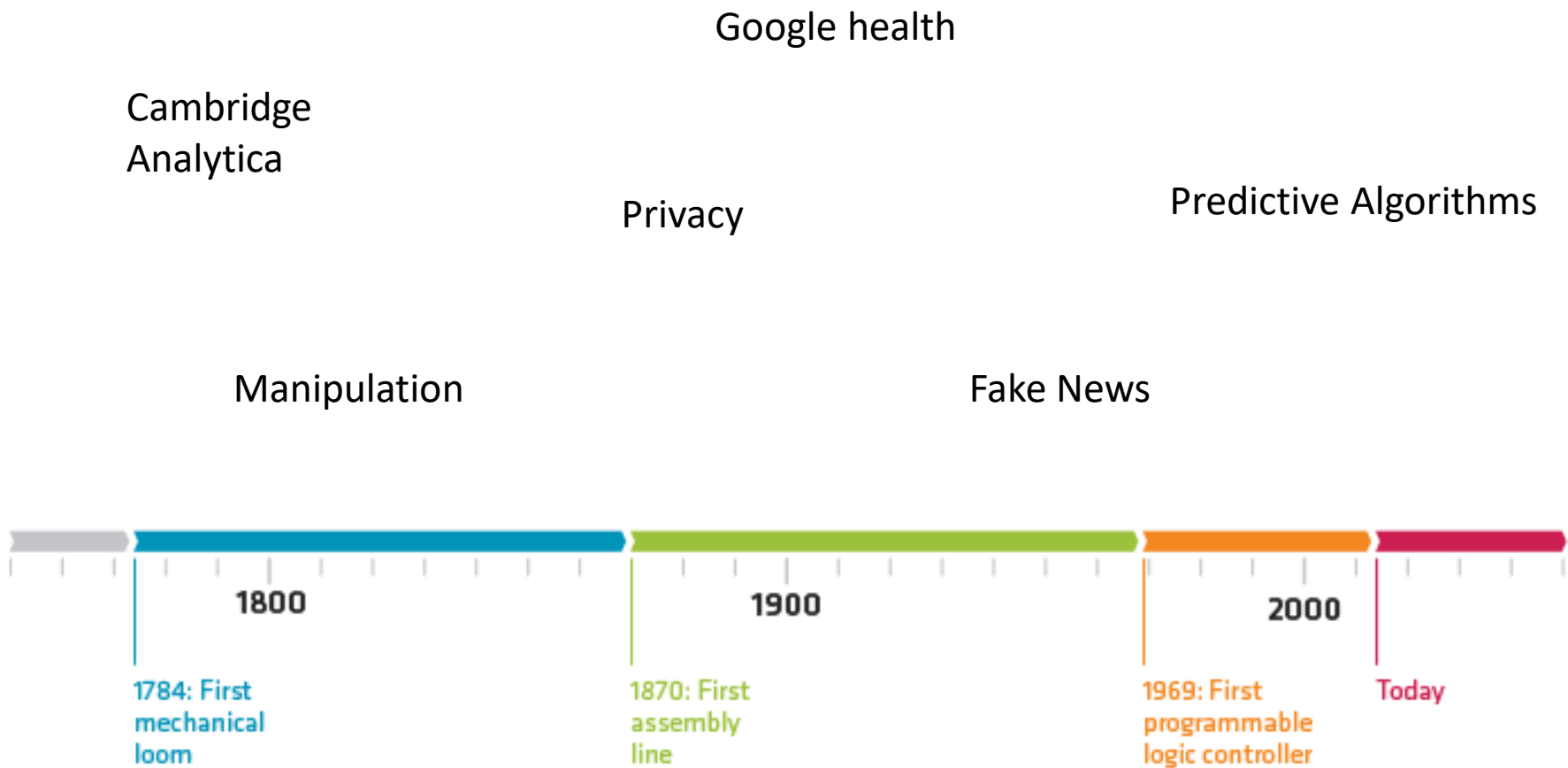
1784: First
mechanical
loom

1870: First
assembly
line

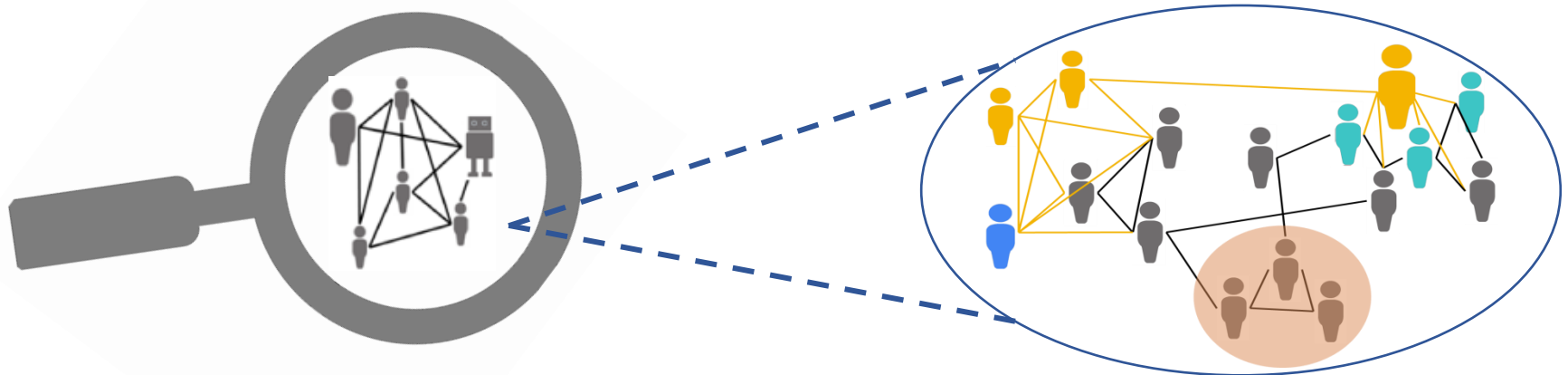
1969: First
programmable
logic controller

Today

From the Industrial Revolution to the Digital Revolution



WE NEED A BROAD SOCIETAL DISCUSSION



There are rich and essential parts of human life which are alogical, which are immaterial and non-quantitative in character, and which cannot be seen under the microscope, weighed with the balance, nor caught by the most sensitive microphone.

If science deals with quantitative problems of a purely logical character, if science has no recognition of or concern for value or purpose, how can modern scientific man achieve a balanced good life, in which logic is the companion of beauty, and efficiency is the partner of virtue?

In one sense the answer is very simple: our morals must catch up with our machinery. To state the necessity, however, is not to achieve it.

Thank you!

socialcomplexity.eu

@mjoanasa

@SPAClab

<https://www.facebook.com/SPACLab>