

Social Impact of Artificial Intelligence: research challenges

Virgilio Almeida

virgilio@dcc.ufmg.br

@virgilioalmeida

9/08/2022


Escola FAPESP 60 anos

U F *m* G



Six Transformations to achieve the Sustainable Development Goals

Jeffrey D. Sachs¹, Guido Schmidt-Traub^{2*}, Mariana Mazzucato³, Dirk Messner⁴,
Nebojsa Nakicenovic⁵ and Johan Rockström⁶



Transformation 6. Artificial intelligence and other digital technologies—sometimes referred to as the Fourth Industrial Revolution— are disrupting nearly every sector of the economy, including agriculture (precision agriculture), mining (autonomous vehicles), manufacturing (robotics), retail (e-commerce), finance (e-payments and trading strategies), media (social networks), health (diagnostics and telemedicine), education (online learning), public administration (e-governance and e-voting) and science and technology.



A.I. Is Not Sentient. Why Do People Say It Is?

Robots can't think or feel, despite what the researchers who build them want to believe.



Facebook Approved Pro-Genocide Ads in Kenya After Claiming to Foster 'Safe and Secure' Elections

Kenya's national cohesion watchdog has threatened to suspend the social network from the country in a week if it doesn't mitigate hate

By Dell Cameron | Friday 2:03PM | Comments (15) | Alerts



Meta CEO Mark Zuckerberg speaks at Georgetown University in Washington, Thursday, Oct. 17, 2019.



How to act?



Science

+ FOLLOW

China's court AI reaches every corner of justice system, advising judges and streamlining punishment

- Smart court's electronic reach allows the system to access police, prosecutor and government databases and integrate with China's social credit system
- Chinese law professor warns, 'We must be alert to the erosion of judicial power by technology companies and capital'



Stephen Chen in Beijing

+ FOLLOW

Published: 8:30pm, 13 Jul 2022

Why you can trust SCMP

TOP PICKS

Private-market investment solutions for HNW individuals

Paid Post: HSBC



News

Record-breaking Chinese bridge clears hurdle to faster travel in border province

18 Jul 2022



News

China's famous

512



Zhang Linghan, professor of law at the China University of Political Science and Law in Beijing, warned that the **rapid rise of AI could create a world where humans were ruled by machines.**

Its electronic reach allows the smart court to work with China's powerful "social credit system" to ban a person refusing to pay a debt from using a plane, high speed train, hotels or other social services.



Dutch scandal serves as a warning for Europe over risks of using algorithms

The Dutch tax authority ruined thousands of lives after using an algorithm to spot suspected benefits fraud – and critics say there is little stopping it from happening again.

BY [MELISSA HEIKKILÄ](#)



Most Read Articles

1. Sweden wins the NATO lottery

APRIL 13, 2022 | 4:04 AM

2. Ukraine announces capture of top Putin pal Viktor Medvedchuk

APRIL 13, 2022 | 5:56 AM

3. How Germany Inc. played Russian roulette – and lost

APRIL 13, 2022 | 4:00 AM

4. Ukraine rejects visit by German President Steinmeier

Predictive Policing through AI

An Algorithm That Grants Freedom, or Takes It Away

Across the United States and Europe, software is making probation decisions and predicting whether teens will commit crime. Opponents want more human oversight.

Virgilio Almeida, 2020



Brazilian Favelas need racial Justice, not Facial Recognition

By

RELATED POSTS

Latin America: Reinforcing democratic institutions to achieve peace, justice and inclusion

This policy brief calls for a rethink of the social contract in Latin America to tackle the redistribution of power...

COVID-19 and the Crisis of Governance: The impact of the Pandemic on Peace, Justice and Inclusion (SDG 16)

As countries adopt COVID-19 response-and-recovery plans, there is an opportunity to achieve more peaceful and just societies with more inclusive...

SDGs as Citizens' Agenda: Stress Peace, Justice and Strong Institutions

The Sustainable Development Goals (SDGs) for the 2030 Agenda for Sustainable Development are not only an agenda belonging to either...

Yet, in Brazil and other contexts marked by extreme racial, ethnic and gender inequalities, such technologies can backfire.

Historically marginalized groups are more exposed to mistakes made by such mechanisms.

Additionally, these technologies can be used as a tool for community control – increasing violence, hypervigilance by state actors and judicial bias in criminal proceedings.

← Tweet



Frank Pasquale
@FrankPasquale

...

“The global far right is using YouTube and other social media to radicalize people who follow algorithmic recommendations to hate speech and misinformation while countering grassroots efforts to deplatform such dangerous language.”



These Machines Won't Kill Fascism: Toward a Militant Progressive Vision for Te...
The left must vie for control over the algorithms, data, and infrastructure that shape our lives.
thenation.com

“A **consolidated transdisciplinary approach** to understanding and managing human collective behavior will be a monumental challenge, yet it is a necessary one. Given that algorithms and companies are already altering our global patterns of behavior for financial reasons, there is no safe hands-off approach..”

“Misinformation, algorithms, dark patterns, and other issues arise at a rate far greater than they can be adequately characterized, much less addressed.”

“With the rise of the digital age, this social network is increasingly coupled to algorithms that create unprecedented feedback effects.”



Facebook Services Are Used to Spread Religious Hatred in India, Internal Documents Show

Company researchers say anti-Muslim material is rife and calls to violence coincided with deadly riots last year. Facebook says it's working to improve enforcement.



EXTREMISM ONLINE

The Human Infrastructure of Fake News in Brazil

The role of algorithms in promoting disinformation has received a great deal of attention in recent years, due in large part to the centrality of Facebook in the 2016 US presidential election and the UK Brexit campaign. However, David Nemer argues that in countries such as Brazil, where peer-to-peer messaging apps like WhatsApp are popular, more attention needs to be paid to the "human infrastructure" of coordinated disinformation campaigns.

by David Nemer
20 hours ago



By Li Yuan – New York Times

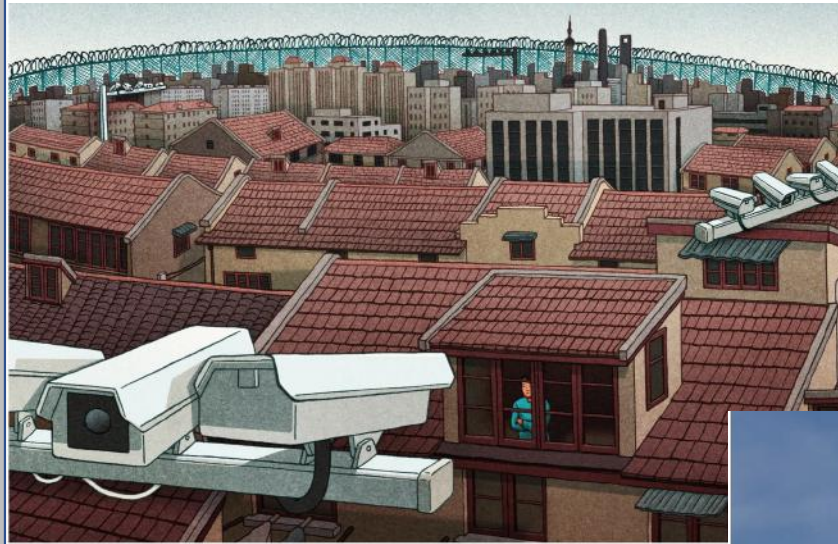
Published May 6, 2022

He fears an even more dystopian China than what it is today: a digital totalitarian regime that surveils everyone, makes each neighborhood an on-site concentration camp and controls the society with the same iron fist in a future crisis, be it war, famine, climate disaster or economic meltdown.

THE NEW NEW WORLD

Has Shanghai Been Xinjianged?

Shanghai used to be the glamorous China, while Xinjiang was the dark China. Now both are casualties of authoritarian excess.



Xinmei Liu



Surveillance cameras in Shanghai this month. Alex Plavevski/EPA, via Shutterstock

Surveillance Society

The stage, which we are living through now, a chaos caused by the profit-driven algorithmic amplification, dissemination and microtargeting of corrupt information, much of it produced by coordinated schemes of disinformation.

Its effects are felt in the real world, where they splinter shared reality, poison social discourse, paralyze democratic politics and sometimes instigate violence and death.

The New York Times

Opinion

The Coup We Are Not Talking About

We can have democracy, or we can have a surveillance society, but we cannot have both.

By Shoshana Zuboff

Dr. Zuboff, a professor emeritus at Harvard Business School, is the author of "The Age of Surveillance Capitalism."

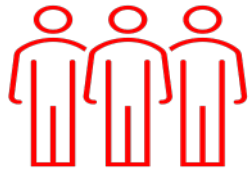
Jan. 29, 2021



DECIPHER ME OR I'LL DEVOUR YOU: The Sphinx Enigma

- Researchers
- Journalists
- Academics
- Policymakers
- Legislators

....



SOCIOTECHNICAL systems:
social media platforms

- Powerful complex and opaque sociotechnical systems
- algorithms and data

AI suggested 40,000 new possible chemical weapons in just six hours

'For me, the concern was just how easy it was to do'

By Justine Calma | @justcalma | Mar 17, 2022, 3:06pm EDT

f t SHARE



1Zrb9bo5WBoll.ZNI" in a new tab



Você precisa de um plano completo para implantar um cobot?

Descubra os caminhos para uma implementação de sucesso.

COMEÇAR AGORA

“We just want more researchers to acknowledge and be aware of potential misuse. When you start working in **the chemistry space**, you do get informed about misuse of **chemistry**...In machine learning, there’s nothing of the sort.”

Amazon Alexa

Amazon's Alexa could turn dead loved ones' voices into digital assistant

Technology promises ability to 'make the memories last' by mimicking the voice of anyone it hears

Alex Hern *Global technology editor*

@alexhern

Thu 23 Jun 2022 11.19 BST



Amazon previously demonstrated how the reanimated voice of an older woman was used to read her grandson a bedtime story. Photograph: Samuel Gibbs/The Guardian

Amazon plans to let people turn their dead loved ones' voices into digital assistants, with the company promising the ability to "make the memories last".

Imminent
Dystopia:
Black Mirror?

If Amazon's Alexa thinks you sound sad, should it suggest that you buy a gallon of ice cream?

He came away convinced that companies should be barred from analyzing what we say and how we sound to recommend products or personalize advertising messages.

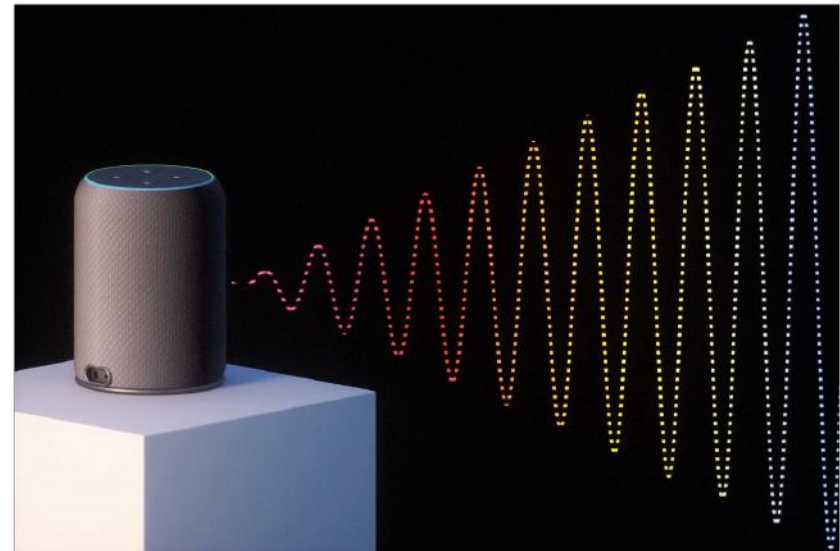
But I got on board with his point that it's worthwhile to start a public conversation about what could go wrong with voice technology, and decide together where our collective red lines are — before they are crossed.

“Using the human body for discriminating among people is something that we should not do,” he said.

The New York Times

Should Alexa Read Our Moods?

The time to debate the power of voice technology is now. (Before Alexa starts suggesting comfort foods.)



Maria Chimishkyan

By [Shira Ovide](#)

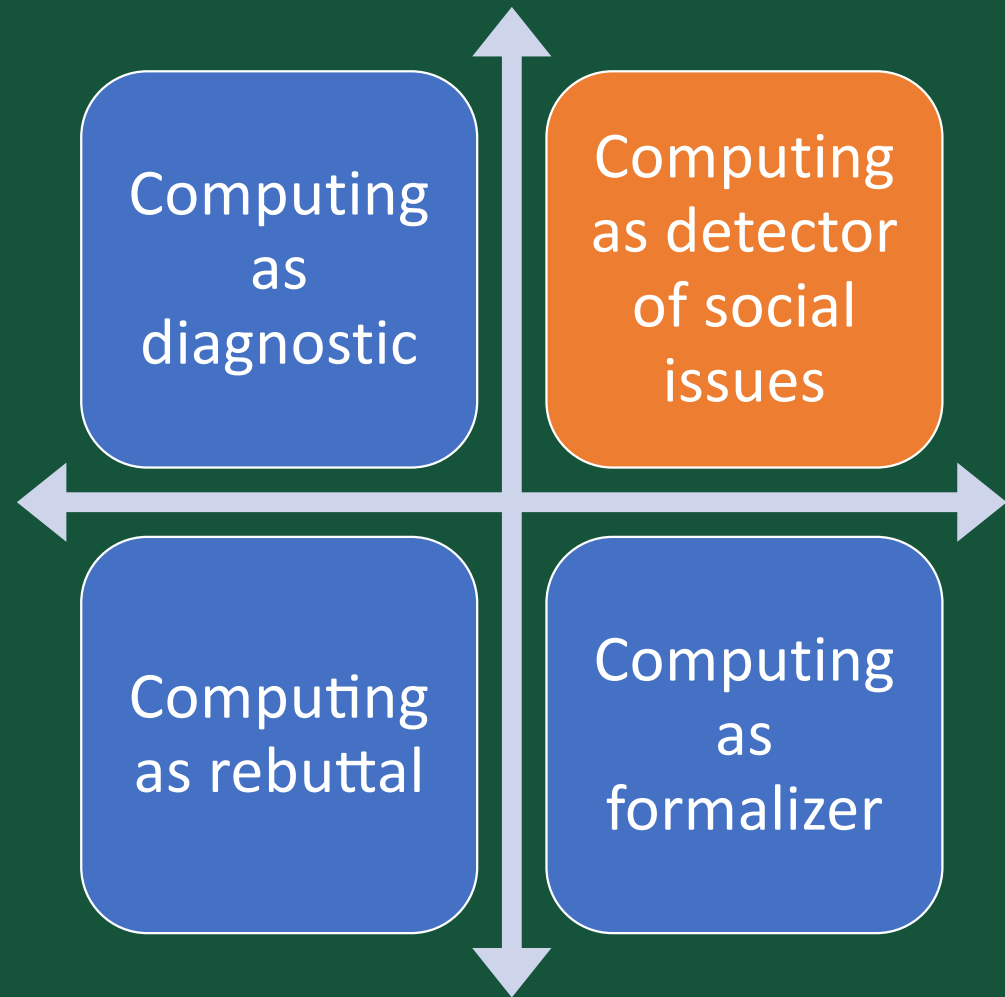
May 19, 2021

Sociotechnical systems



New roles for computing
research
in times of
complex and powerful
sociotechnical systems

New roles for computer science in these times of sociotechnical systems.



(*) Abebe, R. et al, Roles for computing in social change, Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency²⁰, 2020

Computing as a detector of social issues

Its ad-delivery system is excluding women from opportunities without regard to their qualifications. That would be illegal under US employment law.



MS TECH | PEXELS

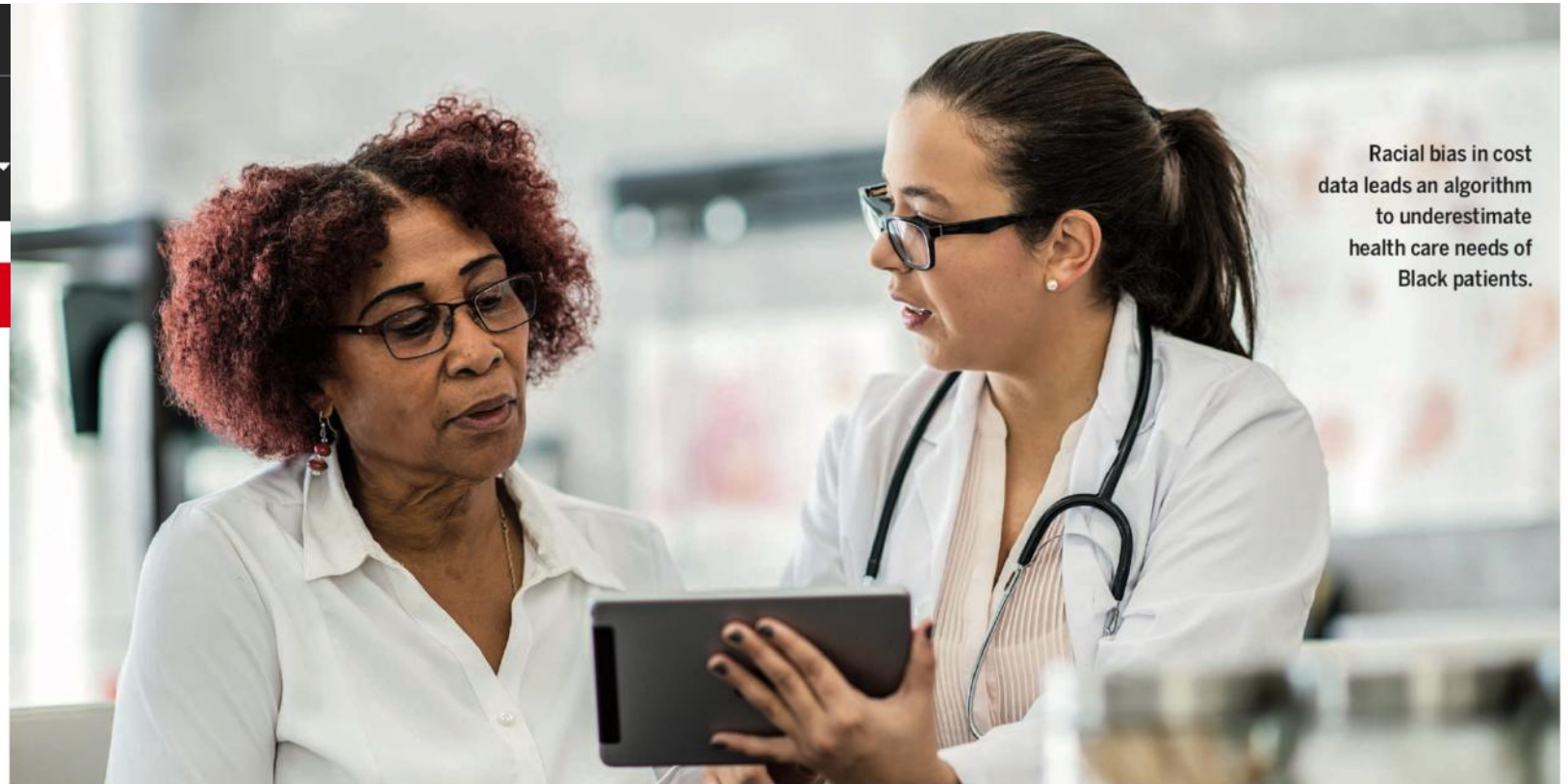
Tech policy / AI Ethics

Facebook's ad algorithms are still excluding women from seeing jobs

Its ad-delivery system is excluding women from opportunities without regard to their qualifications. That would be illegal under US employment law.

[Auditing for Discrimination in Algorithms Delivering Job Ads](#), [Basileal Imana](#), [Aleksandra Korolova](#), and [John Heidemann](#), In Proceedings of [The Web Conference 2021 \(WWW '21\)](#), April 19–23, 2021, Ljubljana, Slovenia.

Computing as diagnostic



Racial bias in cost data leads an algorithm to underestimate health care needs of Black patients.

SOCIAL SCIENCE

Assessing risk, automating racism

A health care algorithm reflects underlying racial bias in society

By **Ruha Benjamin**

As more organizations and indus-

era, the intention to deepen racial inequities was more explicit, today coded inequity is perpetuated precisely because those

beyond the algorithm developers by constructing a more fine-grained measure of health outcomes, by extracting and clean-

Automating care
US news

Supported by

OPEN SOCIETY
FOUNDATIONS

About this content

Virginia Eubanks and
Alexandra Mateescu

Wed 28 Jul 2021 11.00 BST



Computing
as rebuttal

'We don't deserve this': new app places US caregivers under digital surveillance

Tech companies and lawmakers promise that EVV will increase efficiency and accountability in home care and will reduce fraud, waste and abuse in government-funded programs.



▲ Melissa Harville checks the electronic visit verification app for her partner Kevin Hooper outside their home in Greenbrier, Arkansas. Photograph: Rory Doyle/The Guardian

Electronic visit verification systems make homecare more difficult, trap people with disabilities, and cost more than paper timesheets. So why are they rolling out across the country?

Advertisement

Subscribe to The

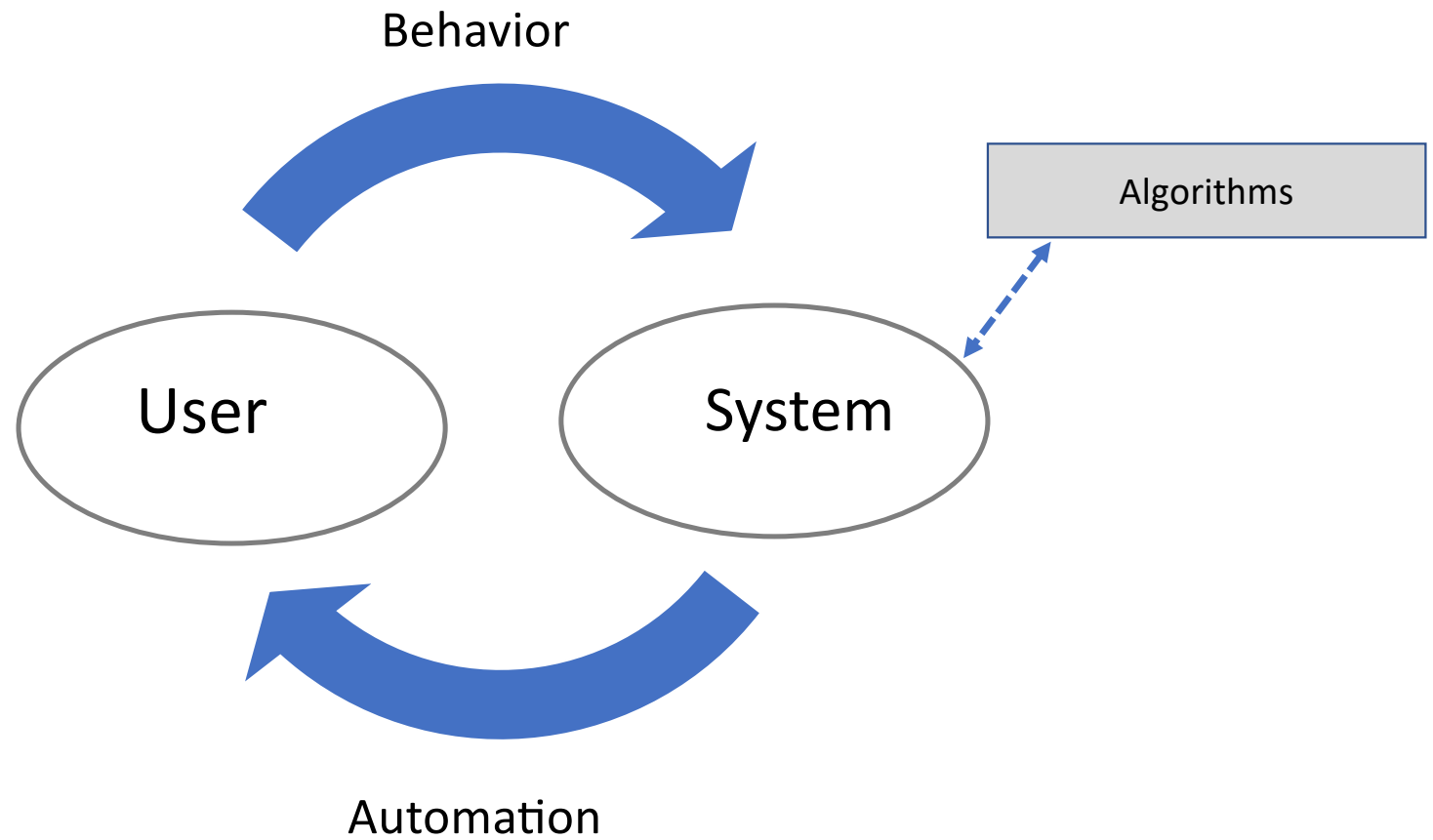
Peeking Beneath the Hood of Uber

- First systematic audit of Uber's surge pricing algorithm
 - Prices update every 5 minutes
 - 70% of surges last \leq 10 minutes
 - Uber divides cities into surge areas
 - May save significant money by walking into an adjacent area
 - Supply and demand do correlate with surge pricing

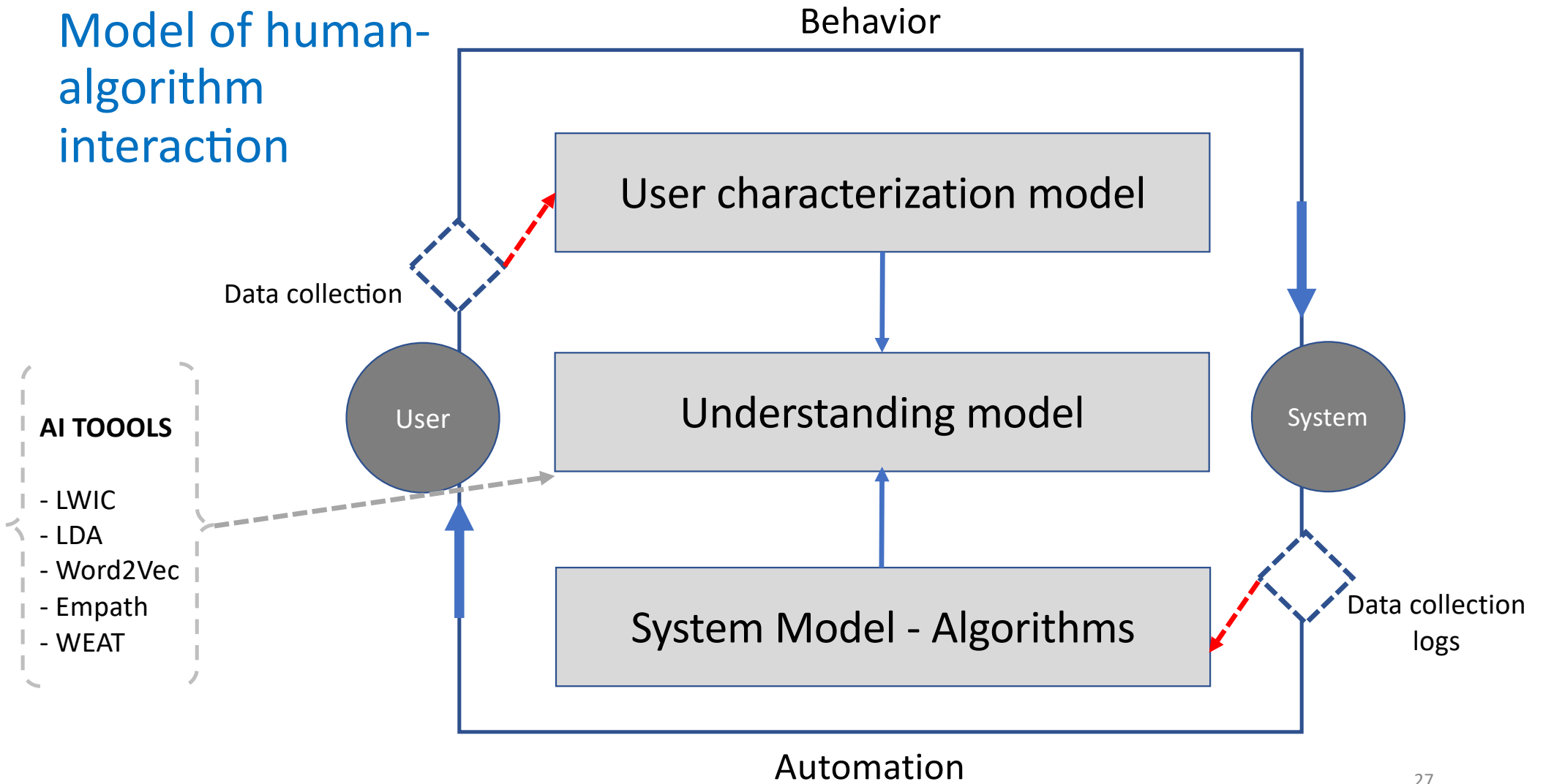


Experimental Research

Simple Model of Human-Algorithm Interaction

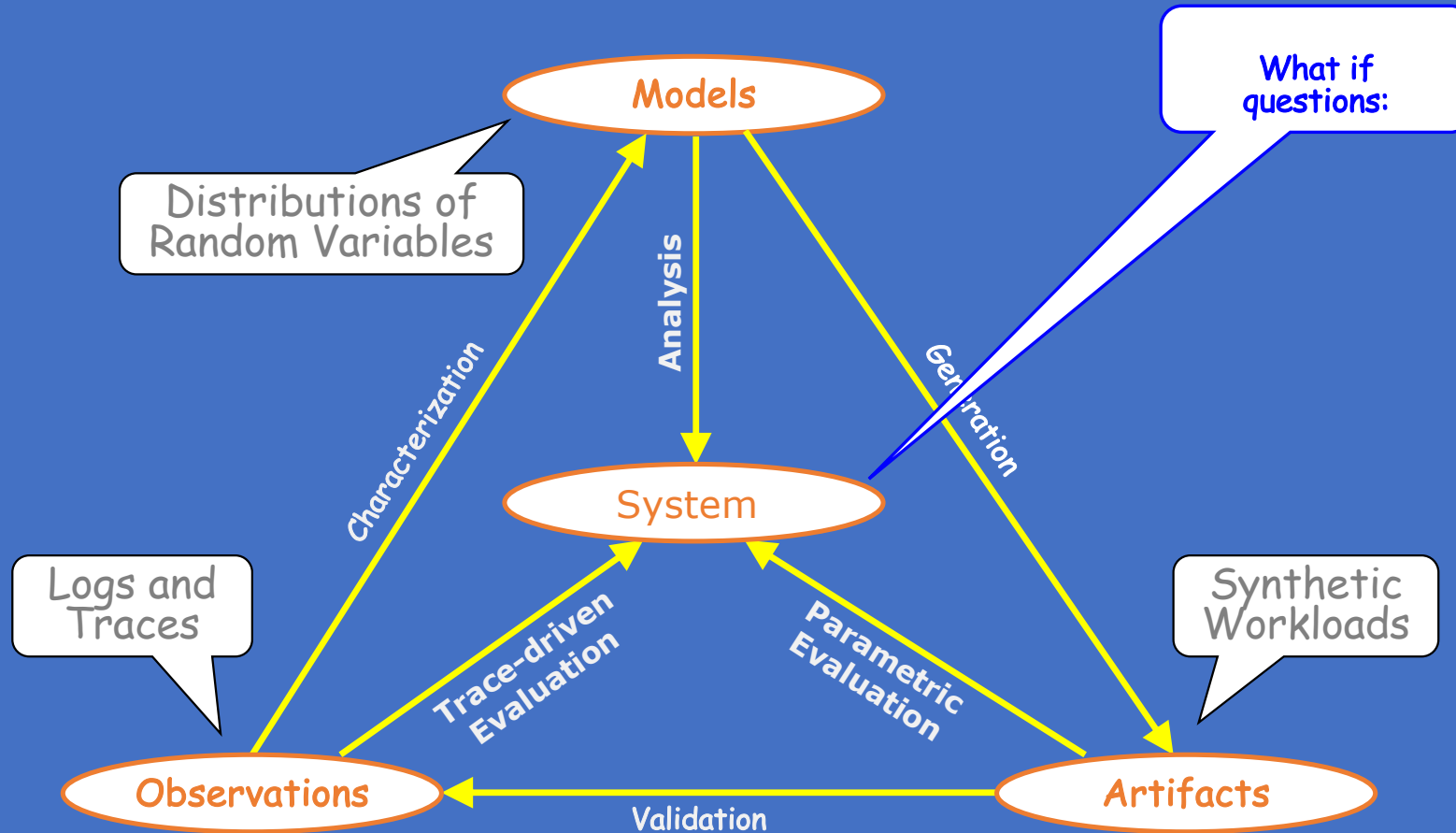


Model of human-algorithm interaction



System Model

Measure → Analyze → Model → Synthesize → Computational Models

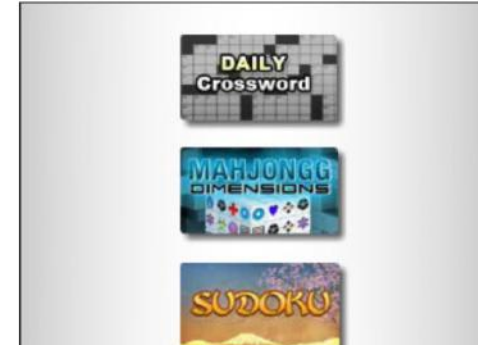
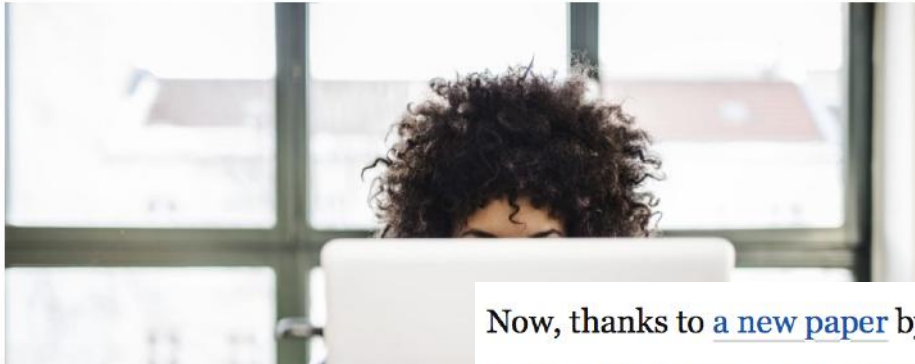




The Intersect

Study: Image results for the Google search ‘ugly woman’ are disproportionately black

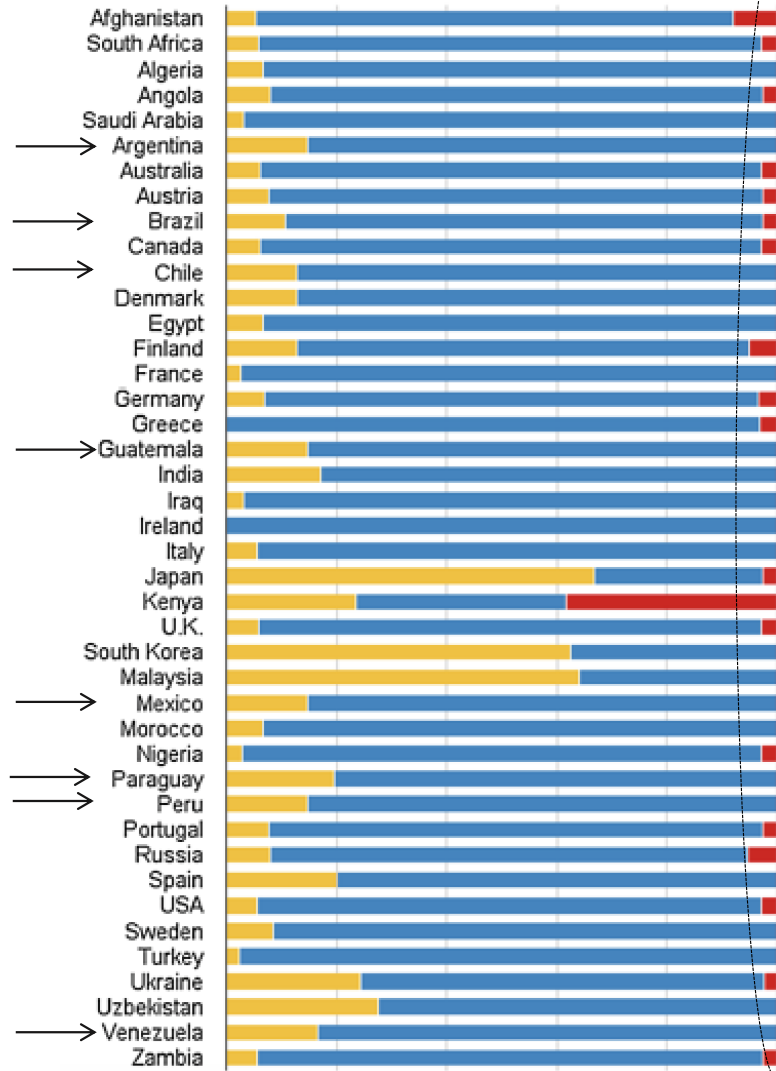
By [Caitlin Dewey](#) August 10 at 1:36 PM



Now, thanks to [a new paper](#) by a trio of researchers at Brazil’s Universidade Federal de Minas Gerais, you can add a new prejudice to the list: Google’s depictions of female beauty, they found, are both ageist and racist.

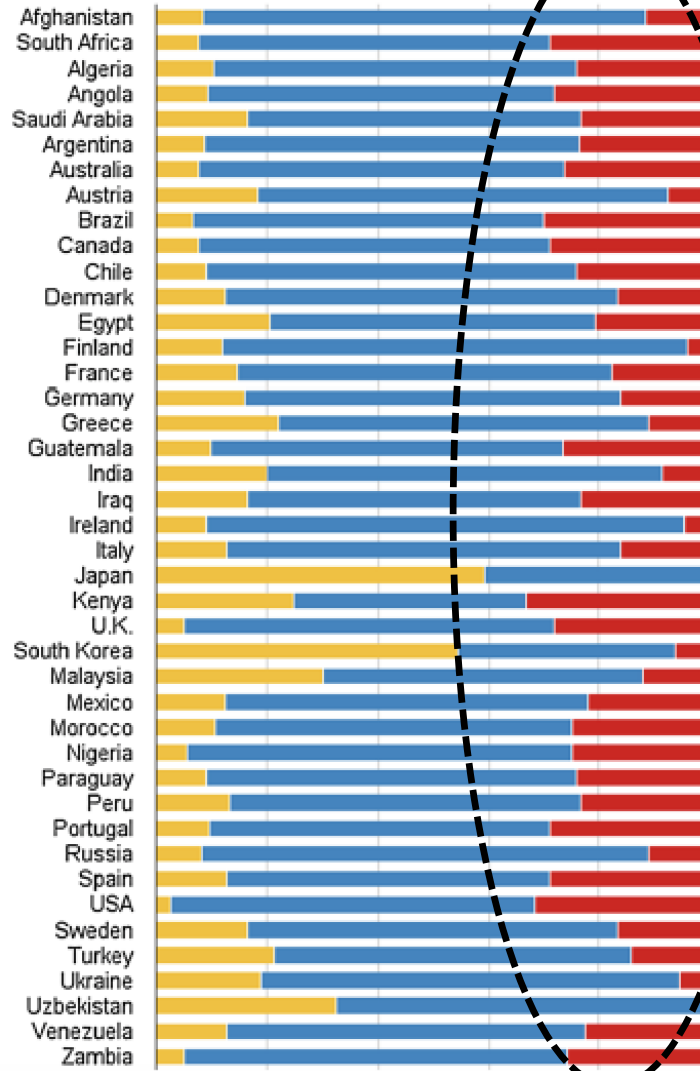
Race Fractions for the query: 'beautiful woman'

Asian White Black



Race Fractions for the query: 'ugly woman'

Asian White Black



Analyzing right-wing conservative YouTube channels: hate, violence and discrimination



Raphael Ottoni
<rapha@dcc.ufmg.br>

Evandro Cunha
<evandrocunha@dcc.ufmg.br>

Gabriel Magno
<magno@dcc.ufmg.br>

Pedro Bernadina
<pedronascimento@dcc.ufmg.br>

Wagner Meira Jr.
<meira@dcc.ufmg.br>

Virgilio Almeida
<virgilio@dcc.ufmg.br>

Research Questions

- . Is the presence of hateful vocabulary, violent content and discriminatory biases more, less or equally accentuated in right-wing conservative channels?**
- . Are, in general, commentators more, less or equally exacerbated than video hosts in an effort to express hate and discrimination?**

. YouTube community seems to **amplify** a discriminatory bias against **Muslims** and **weaken** the bias against **LGBT people**.

. The implicit bias analysis shows **no differences** between **conservative** and **baseline** captions regarding immigrants and LGBT people, but it does show **against Muslims**.

Some Findings

Characterizing Attention Cascades in WhatsApp Groups



Josemar Alves Caetano, Gabriel Magno, Marcos Gonçalves, Jussara Almeida, Humberto T. Marques-Neto, Virgílio Almeida

In WebSci '19: Proceedings of the 10th ACM Conference on Web Science, June 2019 Pages 27–36

WhatsApp, politics and misinformation



May 2018

The Americas

WhatsApp is upending the role of unions in Brazil. Next, it may transform politics.

Facebook's WhatsApp flooded with fake news in Brazil election

Anthony Boadle

5 MIN READ

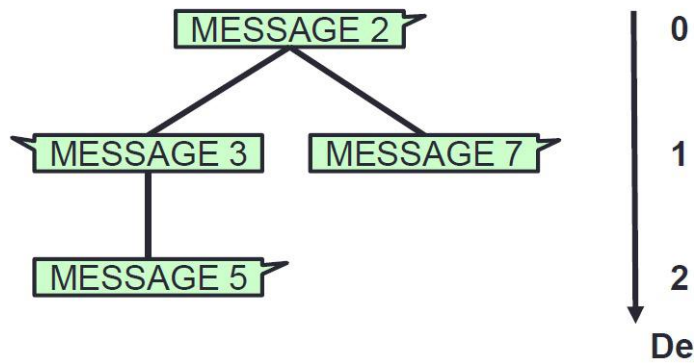


October 2018



Research questions

- How different are attention cascades in political and nonpolitical groups on WhatsApp?
- What is the impact of false information on the characteristics of attention cascades?



Attention
cascade
example

Studying attention cascades on WhatsApp using 3 dimensions

Structural

- Number of nodes
- Depth
- Maximum breadth

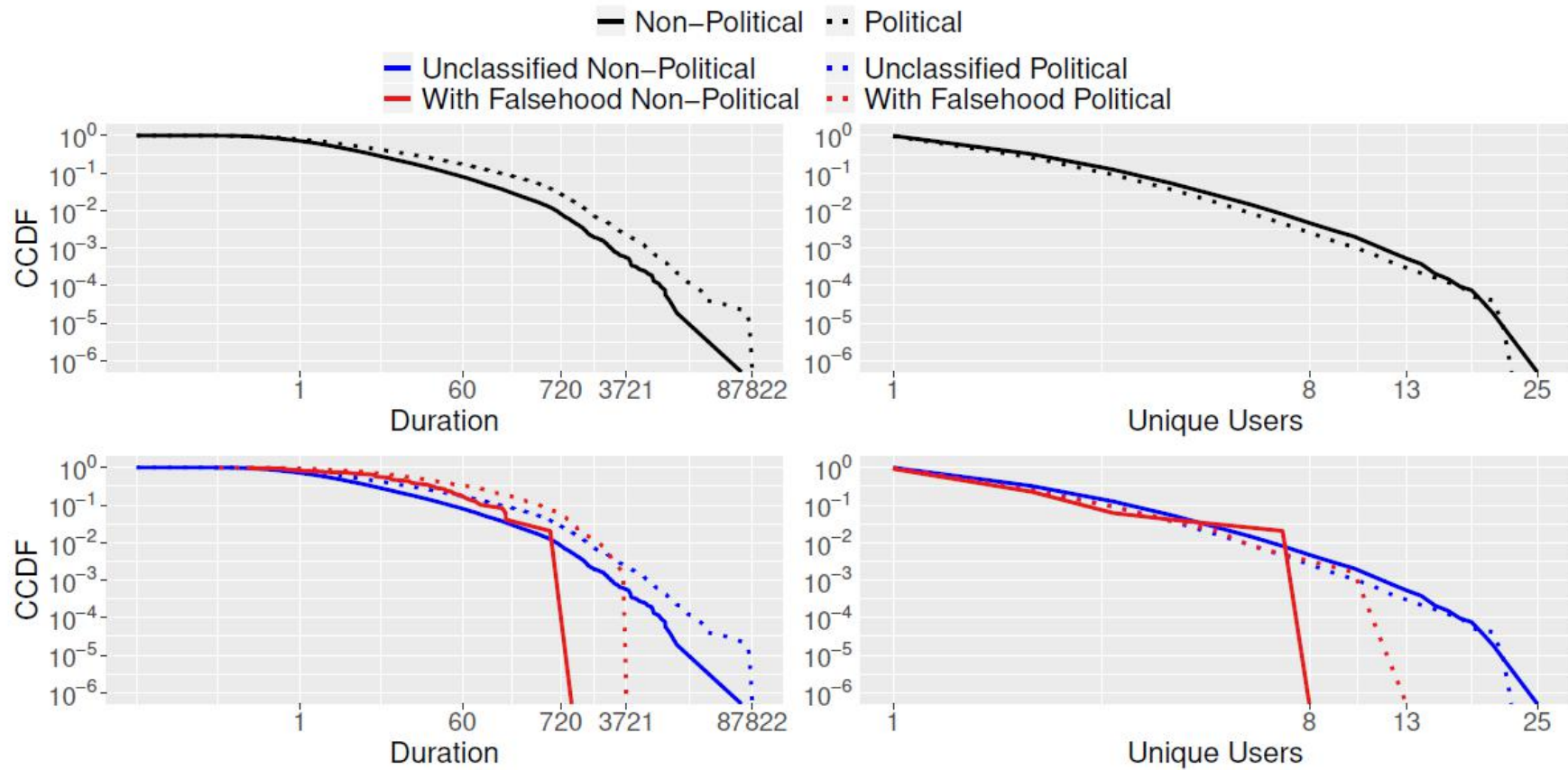
Temporal

- Duration
- Structural characteristics over time

User participation

- Unique users
- Relationships motifs

Cascade duration and unique users



Findings

- Differences in the structural and temporal characteristics of political groups when compared to non-political groups.
- Cascades with false information in political groups are deeper, wider and reach more users than cascades with falsehood in non-political groups.
- Political cascades last longer than non-political ones.
- Attention cascades in WhatsApp groups reflect real-world events that capture public attention

- E.g.: computing research as detector of social issues: hate speech, discrimination, etc.



Auditing Radicalization Pathways on YouTube

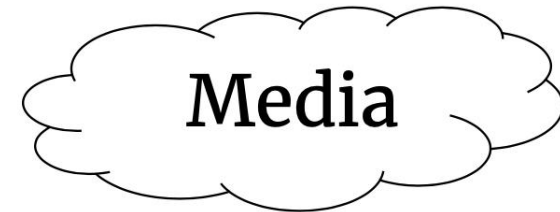


Manoel Horta Ribeiro
Raphael Ottoni
Robert West
Virgílio A. F. Almeida
Wagner Meira Jr.

FAT* '20: Proceedings of the ACM 2020 Conference on Fairness,
Accountability, and Transparency.



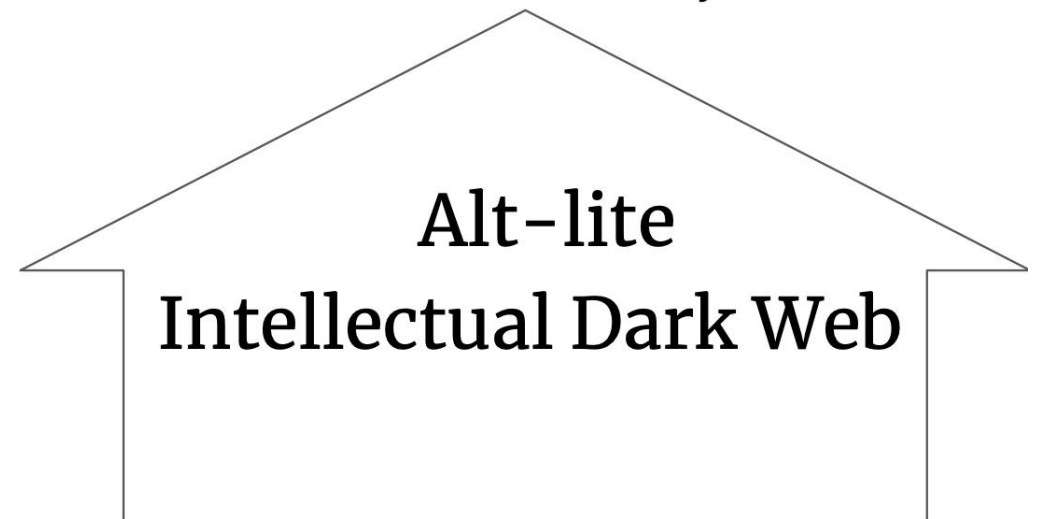
**349
channels
330k
videos
72M+
comments**



(for comparison)

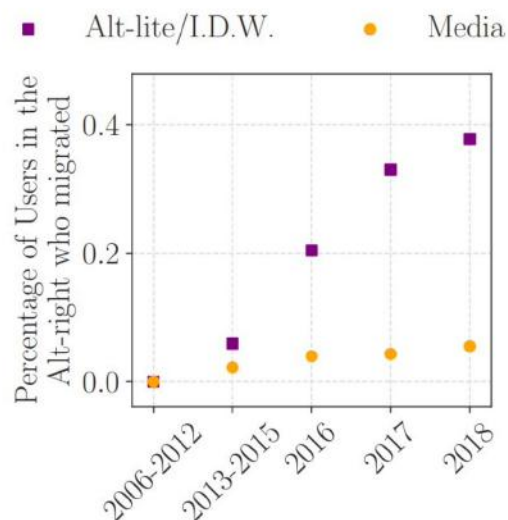
Alt-right

Radical community



Alleged "Gateways"
(to the Alt-right)

But how significant is this pipeline?



~40% of users who commented in the Alt-right in 2018 went through the pipeline (for all levels of exposure).

Research questions:

To which extent do users systematically gravitate towards more extreme content?

Do algorithmic recommendations steer users towards more extreme content?

Findings

- Hypothesis: radicalization happens on YouTube
- By analyzing 349 channels, 330k videos and 72M comments, we found:
 - compelling evidence of significant and consistent migration from milder to more extreme communities
 - recommender system does not seem to be systematically benefiting the more extreme videos, but still binds communities together.
- We did not find why this phenomenon happens on YouTube.

There Is a Racial Divide in Speech-Recognition Systems, Researchers Say

Technology from Amazon, Apple, Google, IBM and Microsoft misidentified 35 percent of words from people who were black. White people fared much better.



Amazon's Echo device is one of many similar gadgets on the market. Researchers say there is a racial divide in the usefulness of speech recognition systems. Grant Hindsley for The New York Times



By Cade Metz

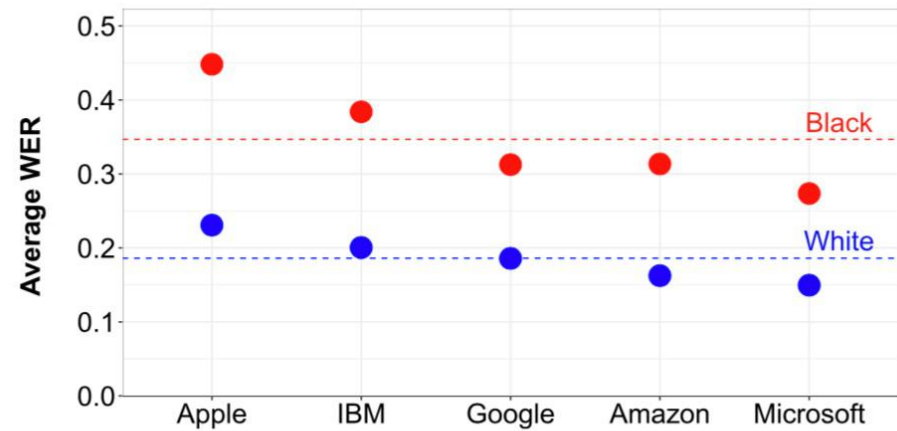
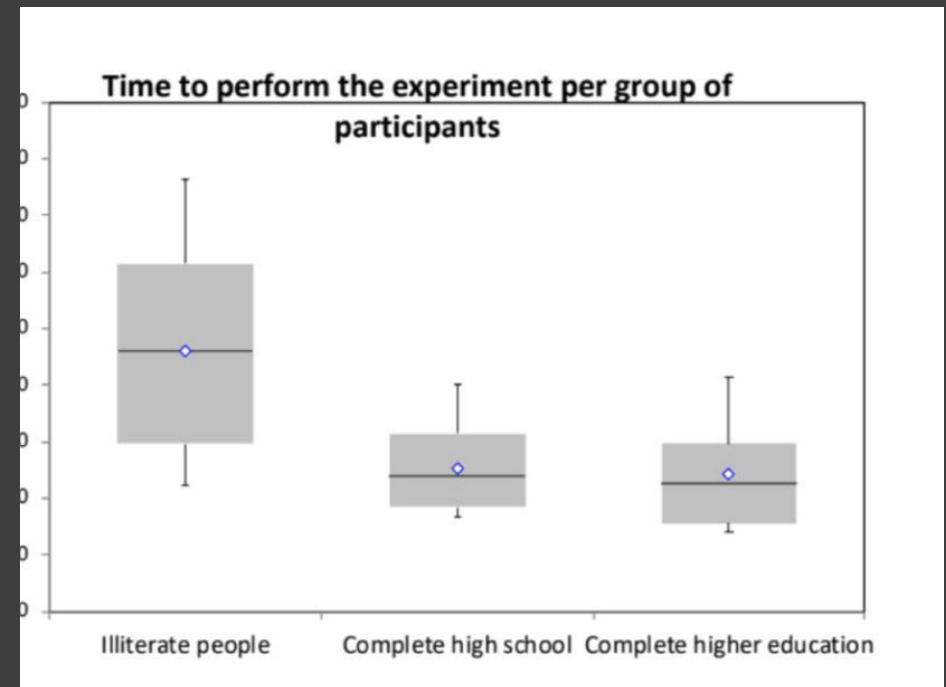
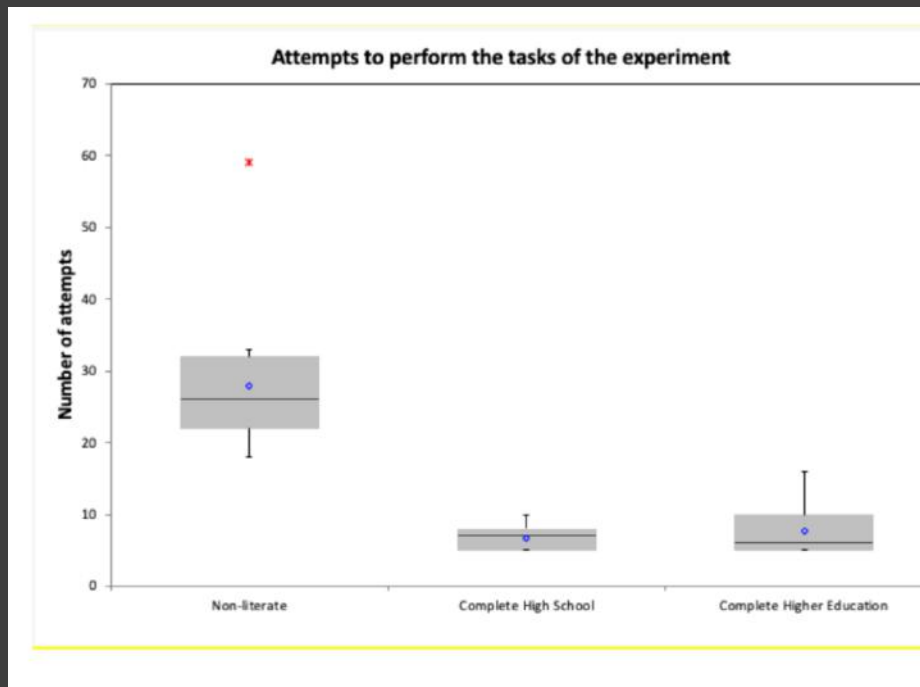


Fig. 1. The average WER across ASR services is 0.35 for audio snippets of black speakers, as opposed to 0.19 for snippets of white speakers. The

- Source: A. Koenecke et al., "Racial disparities in automated speech recognition," Proc. Nat. Acad. Sci., vol. 117, 354 no. 14, pp. 7684–7689, Apr 2020.



How do illiterate people interact with Intelligent Voice Assistants?

- Helen T., Furtado F., Almeida V. et al, "How do illiterate people interact with an Intelligent Voice Assistant? (2021) submitted.

Characterizing Vaccination Movements on YouTube in the United States and Brazil

Marcelo Sartori Locatelli
Josemar Caetano
Wagner Meira Jr.
Virgilio Almeida

June 2022



Research Questions

- Q1: What are the characteristics that differentiate pro-vaccine and anti-vaccine online comments?
- Q2: What kind of language drives engagement in online comments?
- Q3: What are similarities and differences between pro and anti-vaccination campaigns in the United States and Brazil?
- Q4: What are characteristics of the pro and anti-vaccine videos removed from YouTube?

Channel Category	Channel Country	Accounts Number	Total Comments	Comments per Account	% incl URL
Anti-Vaccine	BR	539.678	2.308.712	4,28	21,54%
Pro-Vaccine	BR	350.774	649.644	1,85	22,33%
Anti-Vaccine	USA	280.979	584.844	2,08	13,50%
Pro-Vaccine	USA	1.233.565	3.017.758	2,45	12,99%
Anti-Vaccine	OTHER	248.463	709.947	2,86	14,57%
Pro-Vaccine	OTHER	437.707	1.261.304	2,88	18,14%

Table 1. Basic Statistics by Channel Category and Country for the YouTube Dataset. % **incl URL** represents the percentage of the total number of comments that contain URLs. While building this table, we discarded 8.025 comments that have neither country nor category.

Dataset Data API v3

- Data on 142 channels (pro and anti-vaccine);
- Vaccine related videos;
- Comments on those videos;
- Detection of video takedowns;
- Videos tracked from 11th July 2021 to 3rd October 2021.

Results - Tone of Comments – Valence (Brazil)

Pro-Vaccine

Anti-Vaccine

$-1 \leq v \leq -0.75$

$-0.75 < v < -0.5$

$-0.5 \leq v \leq 0.5$

$0.5 < v < 0.75$

$0.75 \leq v \leq 1$

vaccine
report
fiocruz
researcher
healthy
to_prevent
respiratory
workers
to_cough
poverty

health
disease
isolation
case
avoid
confrontation
sus
mask
science
solidarity

brazil
bolsonaro
president
coronavirus
brazilian
family
population
pandemic
covid
combat

minister
god
freedom
left
media
nation
lord
chaos
success
venezuela

hydroxychloroquine
alliance
communist
conservative
leftist
terrorist
cheer
christian
patriot
pastor

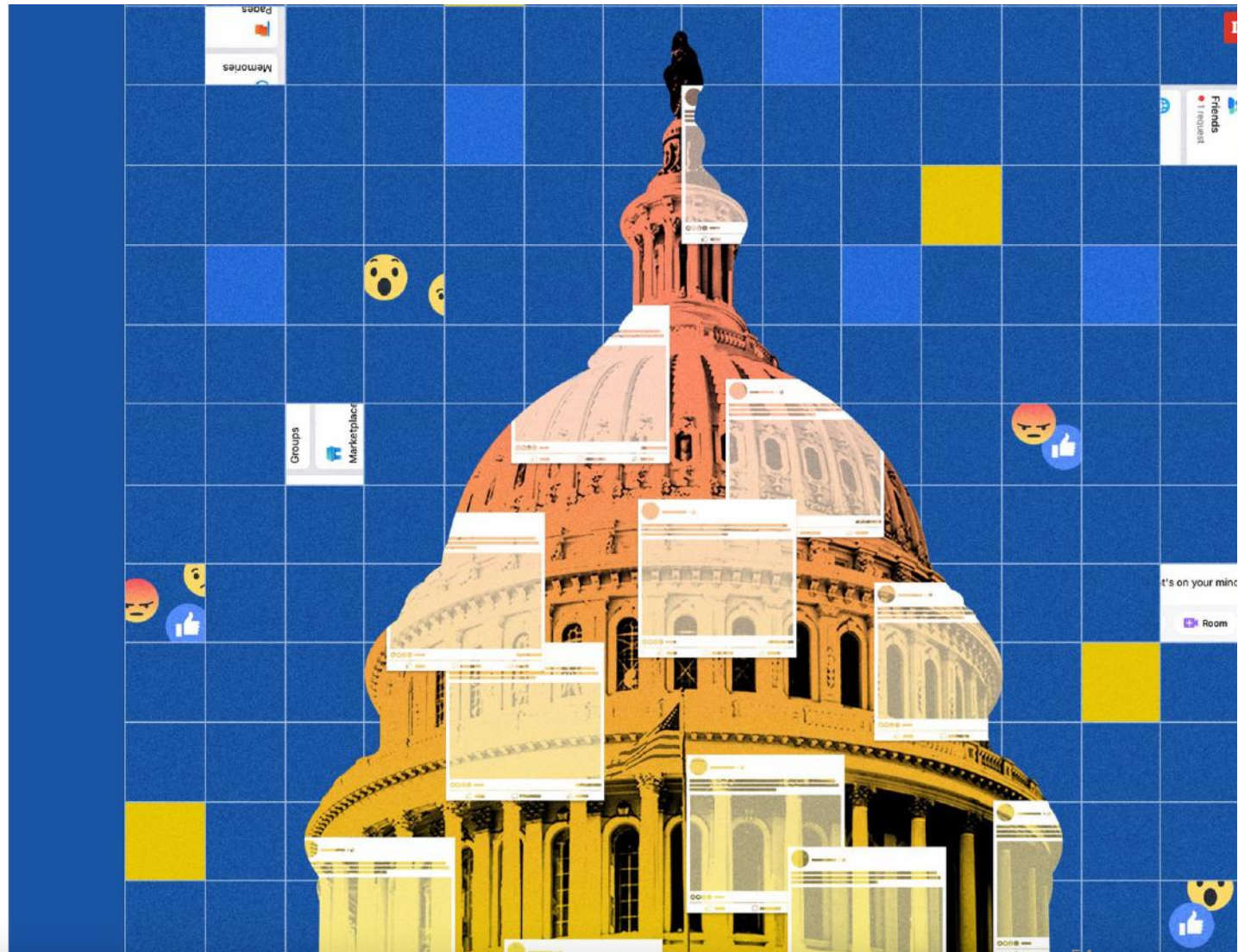
Conclusion

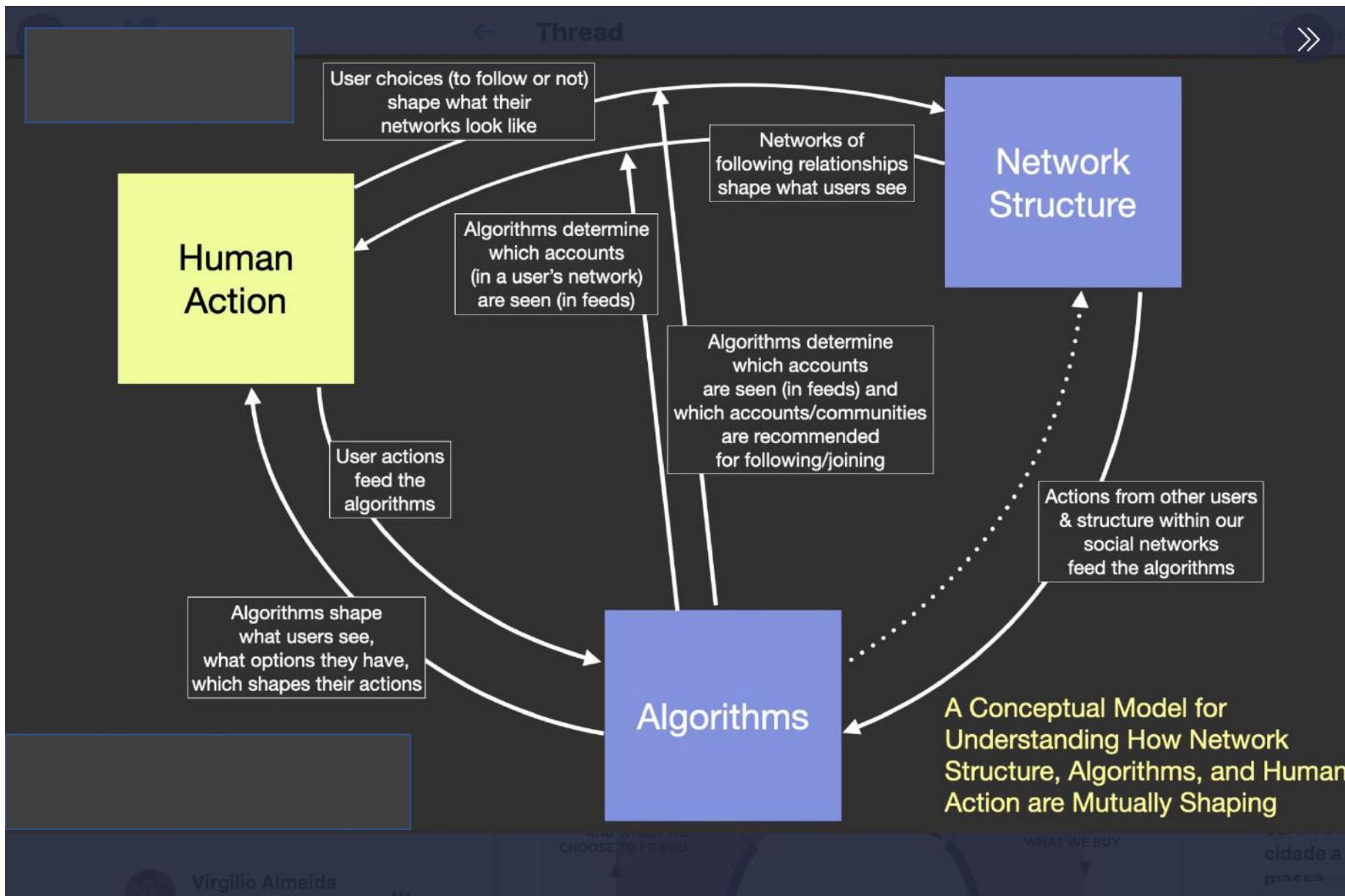
- The results provide insights to understanding communication techniques of pro- and anti-vaccine groups;
- The results can be useful for designing interventions to minimize dissemination of anti-vaccine content;
- Moderation policies on social networks vary depending on the language and the perceived importance of the country.

Facebook Hosted Surge of Misinformation and Insurrection Threats in Months Leading Up to Jan. 6 Attack, Records Show

by [Craig Silverman, ProPublica](#), [Craig Timberg, The Washington Post](#), [Jeff Kao, ProPublica](#), and [Jeremy B. Merrill, The Washington Post](#)

Jan. 4, 8 a.m. EST





Onde atuar para minimizar a disseminação de desinformação em campanhas políticas?

Source: Starbird, 55 K. 2021



Second Revolution. Such a development, says Norbert Wiener, is certain. When it does come, he argues, it will usher in "the second industrial revolution," which will devalue the human brain as the first industrial revolution devalued the human arm. He points out that only a few hand workers can now compete with power-driven machines.

Soon, he warns, there will be wholly automatic factories with artificial brains keeping track of every process. They will order raw materials, inspect them, store them, route them through the plant. They will pay bills, blow the factory whistle and pay the help (if any).

January 23, 1950 | Vol. LV No. 4

**Artificial intelligence
(AI)**

'Risks posed by AI are real': EU moves to beat the algorithms that ruin lives



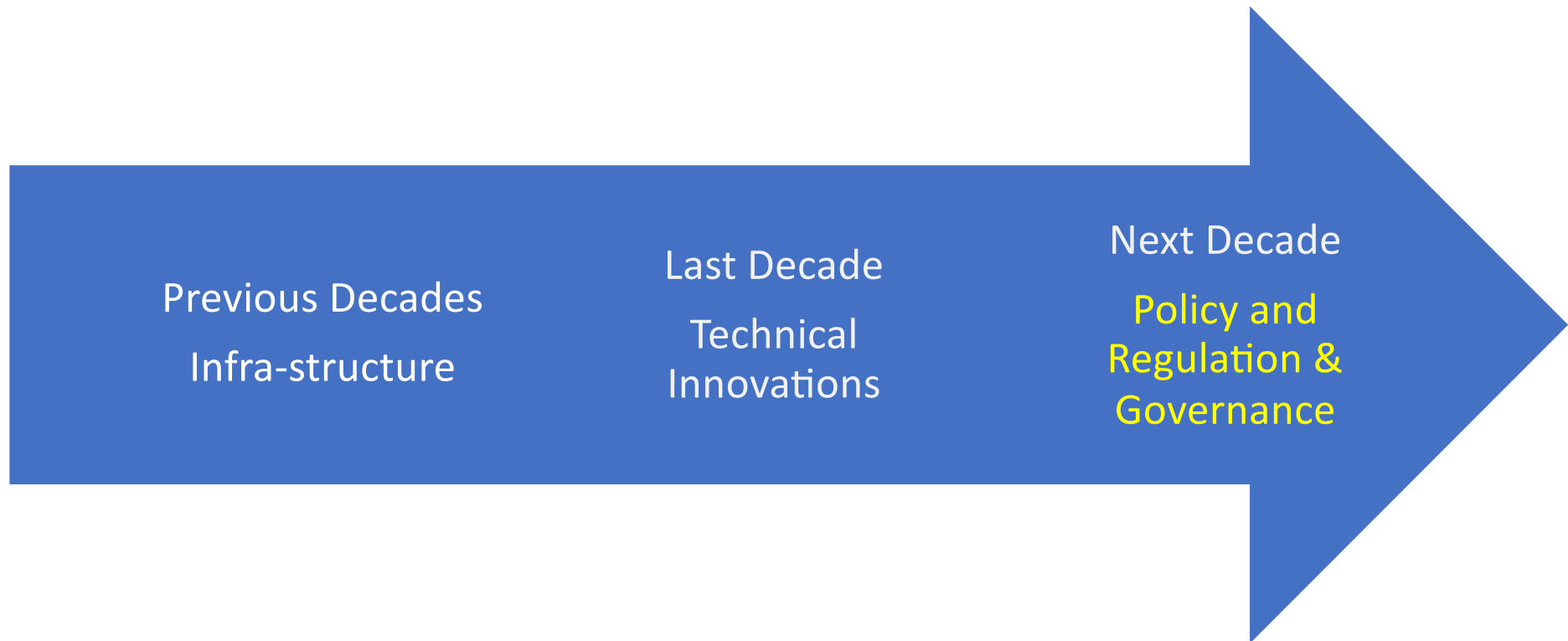
Kalyeena Makortoff
Banking correspondent

🐦 @kalyeena

Sun 7 Aug 2022 13.26 BST



Evolution of the Digital World: narratives

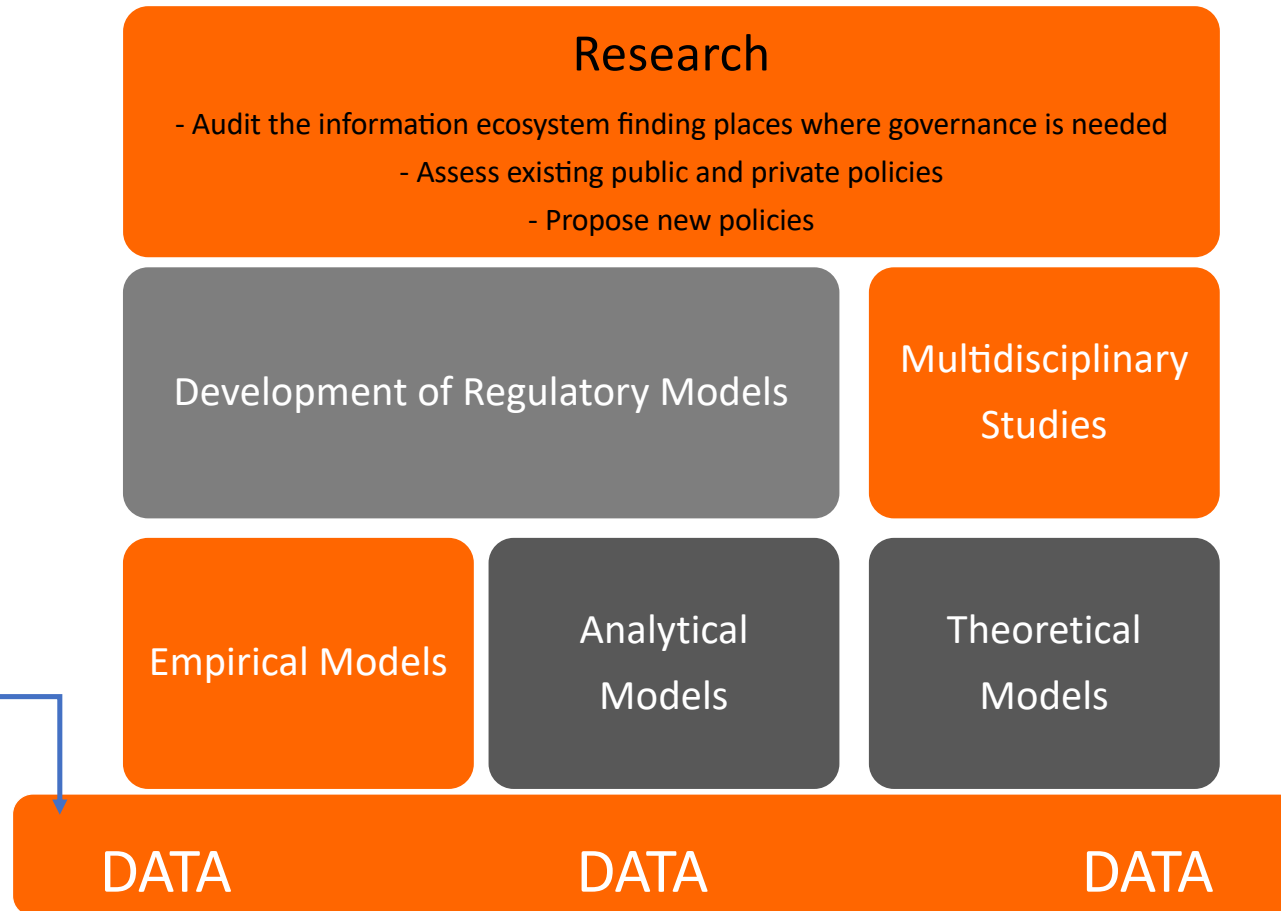


Digital Governance

- “Digital governance is the practice of establishing and implementing policies, procedures, and standards for the proper development, use, and management of the infosphere” (Floridi 2018).
- Digital governance is the capacity of multiple institutions in the digital world to govern (in a legitimate, inclusive, and secure manner) the use of digital commons to produce sustainable services and public policies implemented by governments and firms in a non-territorial and results-based manner (Filgueiras and Almeida, 2020).
- Three essential and common elements of governance structures: i) Decision-making; ii) Accountability and transparency; iii) Coordination

Computing research community can shed light on key problems of digital governance – evidence-based policies

data from social media platforms are a key element.



obrigado!

virgilio@dcc.ufmg.br

publication list available at:

<https://scholar.google.com/citations?user=sPKpIPwAAAAJ&hl=en&oi=ao>

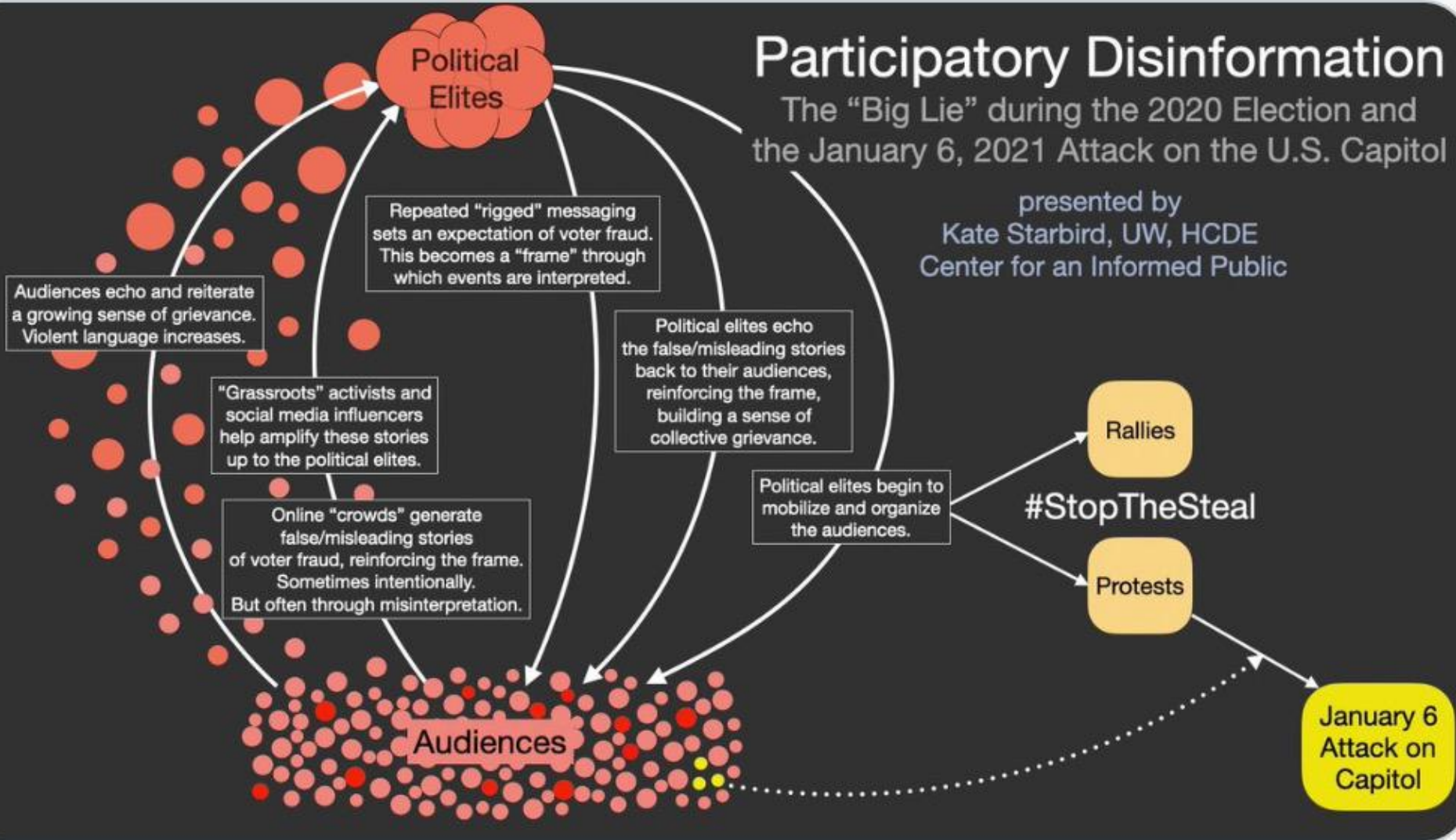
Complexity of technological societies

- According to Yochai Benkler (Harvard), one should not assume that technology is exogenous to markets, nor that it is deterministic.
- Yochai Benkler (Harvard University) has challenged several current theories that suggest phenomena like the automation of work are either inevitable or will follow a predetermined path.
- Using automation as an example, he argued that it is not predetermined but will be affected by a complex web of background choices and institutions whether a society will move towards a replacement of the workforce with robots or complement human labor with machines.

Participatory Disinformation

The "Big Lie" during the 2020 Election and the January 6, 2021 Attack on the U.S. Capitol

presented by
Kate Starbird, UW, HCDE
Center for an Informed Public



The Right to Be Forgotten: A Data-Driven Study

Jason Xue, ECNU and NYU-Shanghai

Gabriel Magno, UFMG

Evandro Cunha, UFMG

Virgilio Almeida, UFMG

Keith W. Ross, NYU and NYU-Shanghai





Researchers Uncover a
Flaw in Europe's Tough
Privacy Rules

TECH TIP
Taking OS X Security Seriously



BITS
Huawei Draws Attention
of U.S. Regulators

TECHNOLOGY

Researchers Uncover a Flaw in Europe's Tough Privacy Rules

By MARK SCOTT JUNE 3, 2016

Europe likes to think it [leads the world](#) in protecting people's privacy, and that is particularly true for the region's so-called [right to be forgotten](#). That legal right allows people connected to the Continent to ask the likes of Google to remove links about themselves from online search results, under certain conditions.

Yet that right — one of the world's most widespread efforts to protect people's privacy online — may not be as effective as many European policy



The RTBF Law: Motivation

- Suppose when people google your name, the first link points to an article:
 - about you going bankrupt 20 years ago.
 - or a crime for which you were acquitted.
 - or a minor crime committed as a child.
 - or personal private information, such as your home address or sexual orientation.
- What can you do besides cry?



Proof of Concept: El Mundo

- Query El Mundo with 37 Spanish terms related to crime and download 85K articles.
- Extract names from articles.
- To reduce query effort, use heuristic filters to reduce title+name combinations to 6,410 queries. (4,164 articles)
- Sequentially query google.es from a single machine.
- Discover 2 previously unknown RTBF delisted links.
- Double check by querying google.es with title only and google.com with name only.

- We took a data- driven approach to study the RTBF in the traditional media outlets, its consequences, and its susceptibility to inference attacks.
- With moderate resources and hacking skills, a transparency activist can determine delisted URLs and the names of the people who requested the de-listings.
- The activist can then republish on URLs and names on his own well-known website.
- Can lead to the Streisand effect, putting the efficacy of the RTBF law into question.