



UNIVERSITAT DE
BARCELONA



MSc in Fundamental Principles of Data Science

Ethical Data Science

Information Disorder

Jordi Vitrià

2020-2021

3rd Assignment

What is the cause of the “post-truth” (infodemic) era? Propaganda or polarization?

ASSIGNMENT

3rd Assignment (Essay on "Polarization and Propaganda")

0 of 30 Submitted Due May 23, 2021

→

BOSTON REVIEW
A POLITICAL AND LITERARY FORUM

JOIN SH

April 22, 2021

Topics
DEMOCRACY
MEDIA
PHILOSOPHY
POLITICS

Share



PHILOSOPHY & RELIGION

Polarization or Propaganda?

Two rival theories paint very different pictures of the sources of our democratic dysfunction. The debate won't be settled by accusations of political convenience.

C. THI NGUYEN

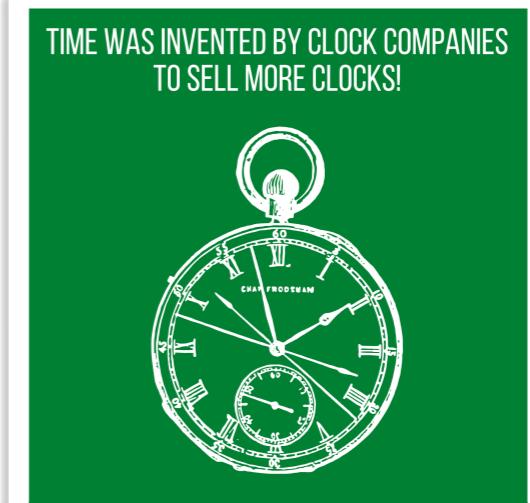


Motivation

Post-truth era is characterized by indifference to evidence.

Credit: Y.Harari

“A conspiracy theory is an explanation for an event or situation that invokes a conspiracy by sinister and powerful groups, often political in motivation, when other explanations are more probable”.



Part of the social contract, in effect, is that in return for political freedom -the right to move freely and to form associations and expect the security and stability needed to engage in projects, for instance- human agents are obliged to protect their autonomy, by keeping an eye out for would-be puppeteers, manipulators who would usurp control and destroy their responsibility.

Credit: D.Dennet in “Just Deserts. Debating Free Will”.

The problem: information disorder

Table 1. Definitions of Some Types of False Information

Term	Definition
Rumor	<i>“An item of circulating information whose veracity status is yet to be verified at the time of posting” [129].</i>
Fake News	<i>“A news article that is intentionally and verifiable false” [96].</i>
Hoax	<i>“A deliberately fabricated falsehood made to masquerade as truth” [51].</i>
Click-bait	<i>“A piece of low-quality journalism which is intended to attract traffic and monetize via advertising revenue” [106].</i>
Disinformation	<i>“Fake or inaccurate information which is intentionally false and deliberately spread” [116].</i>
Misinformation	<i>“Fake or inaccurate information which is unintentionally spread” [116].</i>

<https://arxiv.org/pdf/1909.03654.pdf>

As the role of AI in society grows, **algorithms** can enable malign actors to **steal** information, **compromise** individual privacy, **change individual behavior** and even distort elections without much of a trace.

We need to understand this **social phenomenon!**

Technology Threats

- 1) User profiling and microtargeting;
- 2) Deep fakes;
- 3) Humans finding themselves “out of the loop” of AI systems.

AI Threats

- 1) User profiling and microtargeting;
- 2) Deep fakes;
- 3) Humans finding themselves “out of the loop” of AI systems.

Adversaries will increasingly be able to identify individuals' unique characteristics, beliefs, needs, and vulnerabilities. They will then be able to deliver highly-personalised content, and thereby target with maximum effectiveness those who are most vulnerable to influence.

Example:

The use of psychometrics in 2016 US presidential election. Cambridge Analytica amassed large amounts of data, built **personality profiles** for more than 100 million registered US voters and then, allegedly, used these profiles for targeted advertising **to change voting behavior**.

AI Threats

- 1) User profiling and microtargeting;
- 2) Deep fakes; *'Deep fakes' – digitally manipulated audio or visual material that is highly realistic and virtually indistinguishable from real material.*
- 3) Humans finding themselves “out of the loop” of AI systems.

By 2030, however, deep fakes could become indistinguishable from genuine information and easier to produce. Telling the difference between the original and manipulated content may become close to impossible for news consumers, and progressively difficult for machines.

Deep fakes make it possible for malign actors to deny the truth in two ways: not only may fake videos be passed off as real to create doubt but authentic information can be passed off as fake.

Example:

In May 2019, US House Speaker Nancy Pelosi was herself the victim of a deceptive video, in which she appears to drunkenly slur her words.

AI Threats

- 1) User profiling and microtargeting;
- 2) Deep fakes; *'Deep fakes' – digitally manipulated audio or visual material that is highly realistic and virtually indistinguishable from real material.*
- 3) Humans finding themselves “out of the loop” of AI systems.



AI Threats

- 1) User profiling and microtargeting;
- 2) Deep fakes;
- 3) Humans finding themselves “out of the loop” of AI systems.

As the next wave of AI research focuses on creating tools that are better able to understand human language, context, and reasoning, AI-enabled bots could end up in the driver's seat, with an ability to generate content, persuade, and tailor content for different audiences.

The human factor

Information Disorder

Rumours, conspiracy theories and fabricated information are far from new.

THE GERMAN "KADAVER" FACTORIES.

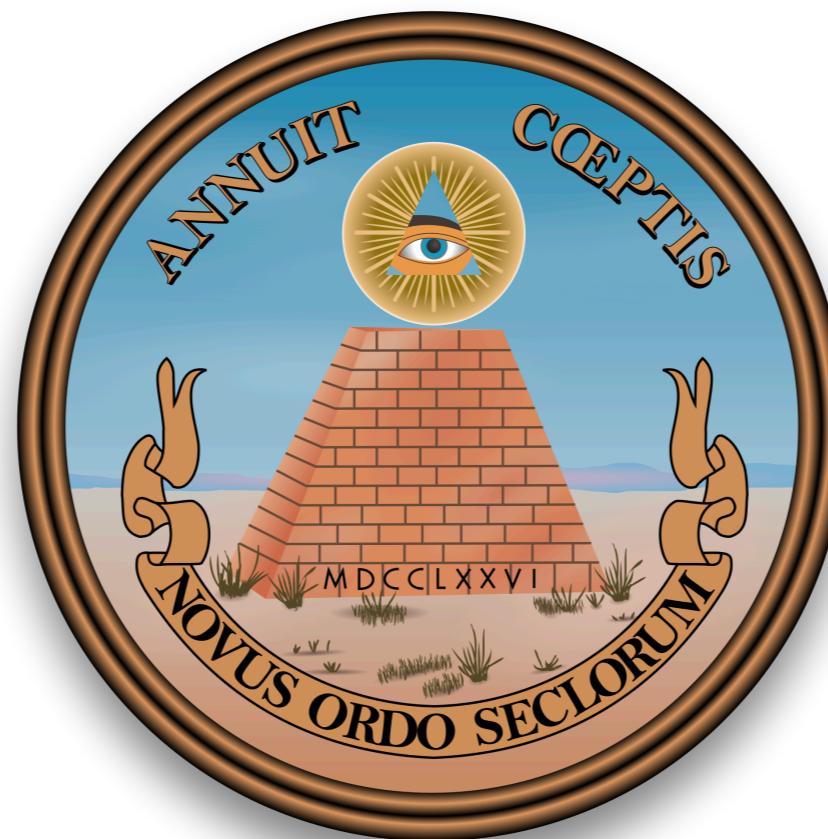
LORD R. CECIL (Hitchin, U.), replying to Mr. R. MCNEILL (St. Augustine's, U.), who asked whether the Government would take steps to make it known as widely as possible in Egypt, India, and the East generally, that the Germans use the dead bodies of their own soldiers, and of their enemies when they obtain possession of them, as food for swine, and to an inquiry by Mr. DILLON (Mayo, E., Nat.) whether the Government had any solid ground for believing to be well founded the statements, widely circulated in this country, that the German Government had set up factories for extracting fat from the bodies of soldiers killed in battle, said:—The Government have no information beyond that contained in extracts from the German Press which have been published in the Press here. In view of other actions taken by the German military authorities there is nothing incredible in the present charge against them. His Majesty's Government have allowed the circulation of the facts as they appeared through the usual channels.

In the spring of 1917, as World War One raged across Europe, both the Times and the Daily Mail in London published accounts from 'anonymous sources' that claimed they had visited a 'Kadaver' factory called Kadaververwertungsanstalt in Germany. This factory was said to extract glycerine from the corpses of the fallen to make soap and margarine.

<https://www.thesocialhistorian.com/fake-news/>

Information Disorder

Rumours, conspiracy theories and fabricated information are far from new.



The **New World Order theory** states that a group of international elites controls governments, industry, and media organizations, with the goal of establishing global hegemony. They are alleged to be implicated in most of the major wars of the last two centuries, to carry out secretly staged events, and to deliberately manipulate economies. Organizations alleged to be part of the plot include the [Federal Reserve System](#), the [Council on Foreign Relations](#), [Trilateral Commission](#), the [Bilderberg Group](#), the [European Union](#), the [United Nations](#), the [World Bank](#), the [International Monetary Fund](#), [Bohemian Grove](#),^[67] [Le Cercle](#)^[68] and Yale University society [Skull and Bones](#).

https://en.wikipedia.org/wiki/List_of_conspiracy_theories#New_World_Order



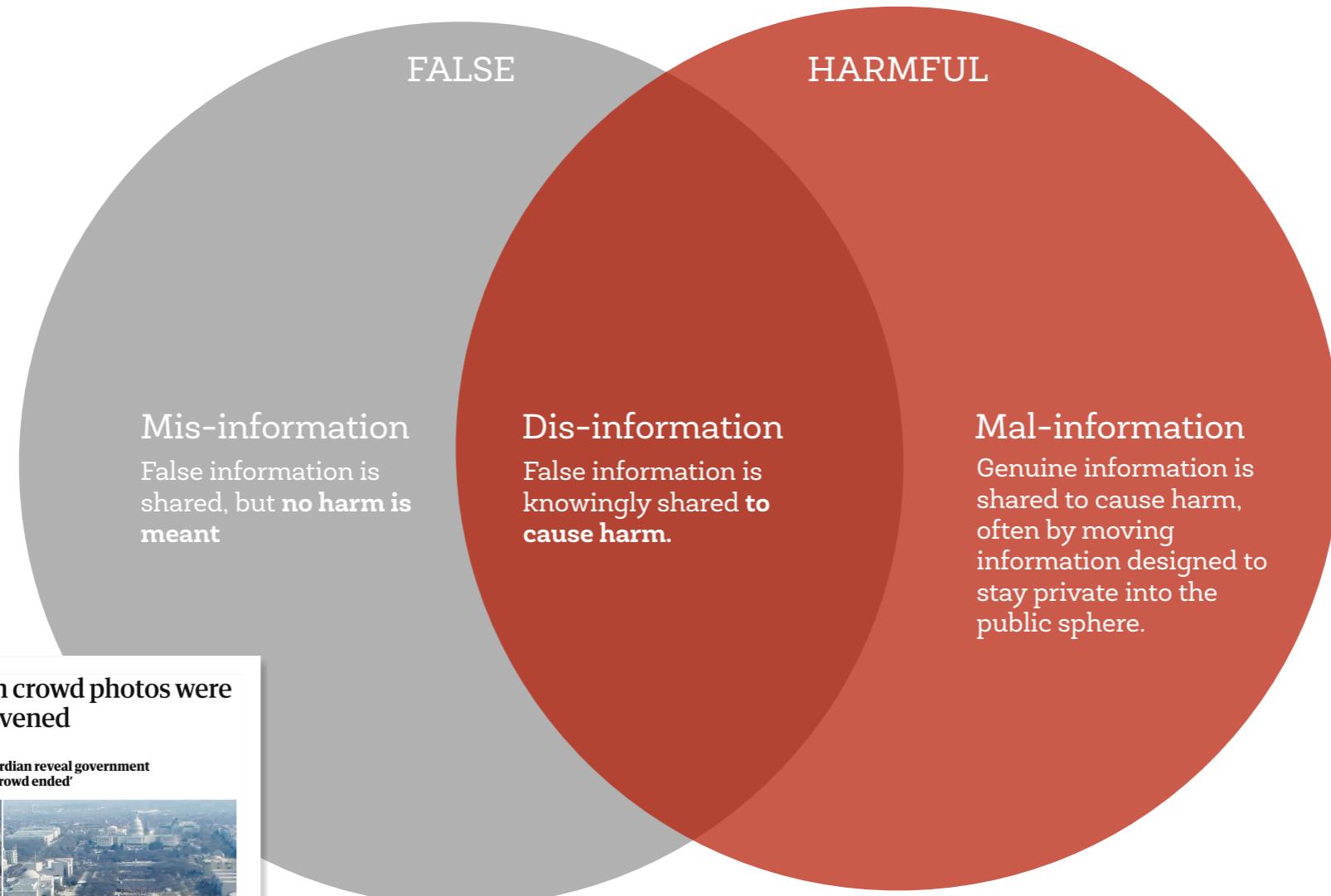
What we know

Information Disorder

There are three types of Information Disorder:

- **Dis-information.** Information that is **false** and deliberately created to **harm** a person, social group, organization or country.
- **Mis-information.** Information that is **false**, but **not created with the intention of causing harm**.
- **Mal-information.** Information that is based on reality, used to inflict **harm** on a person, organization or country.

Information Disorder



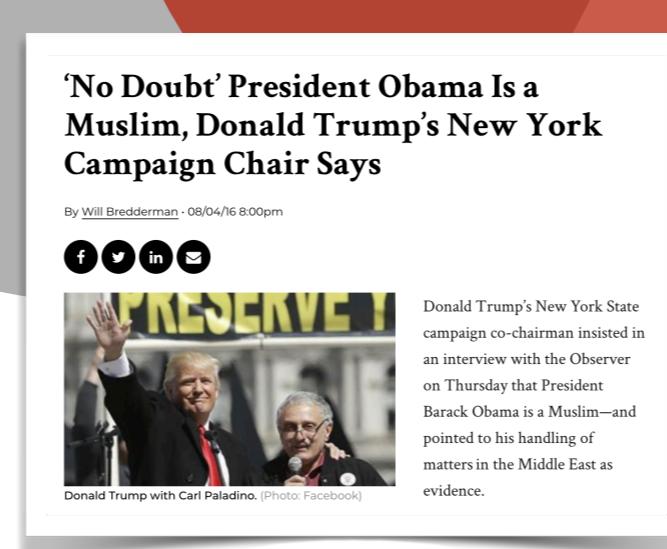
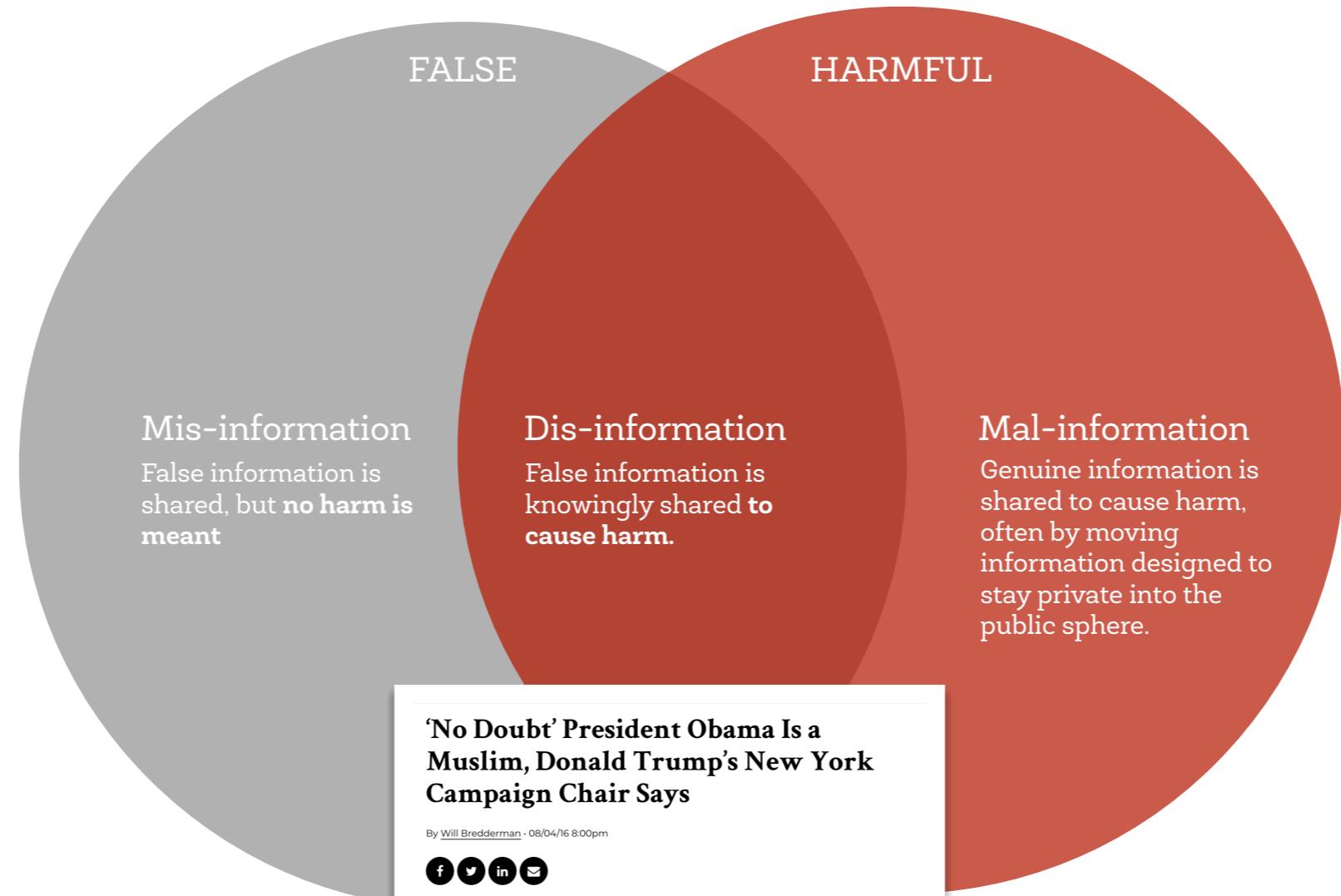
Trump inauguration crowd photos were edited after he intervened

Exclusive: documents released to Guardian reveal government photographer cropped space 'where crowd ended'



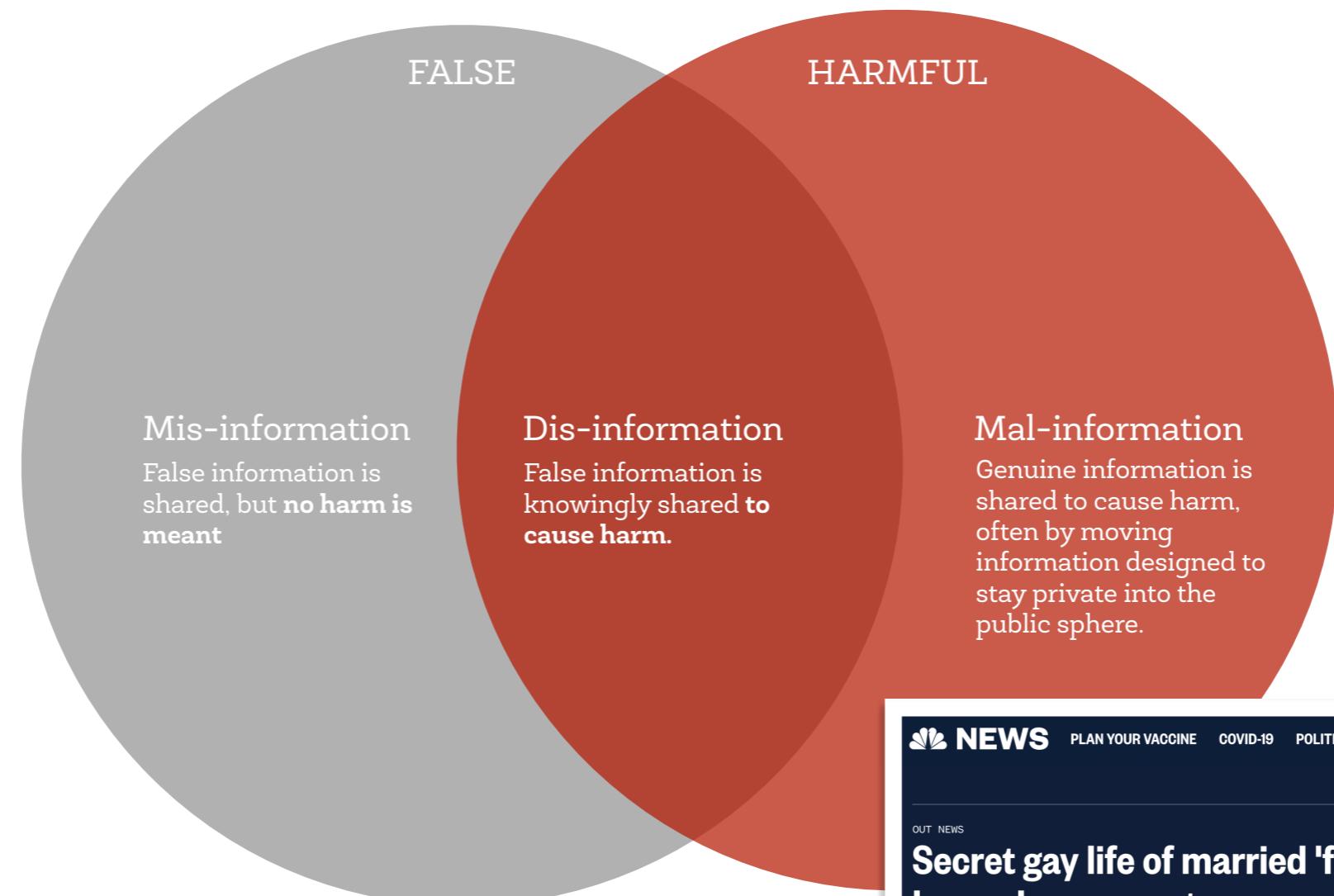
A combination of photos shows the crowds attending the inauguration ceremonies of Donald Trump, left, and Barack Obama. These pictures were taken by Reuters, and were not the edited NPS images. Photograph: Staff/Reuters

Information Disorder



<https://www.coe.int/en/web/freedom-expression/information-disorder>

Information Disorder



Information Disorder

How did we get to this point?

- The 2016 US Presidential election led to an immediate search for answers from those who had not considered the possibility of a Trump victory.

“In the final three months of the US presidential campaign, 20 top-performing false election stories from hoax sites and hyper-partisan blogs generated 8,711,000 shares, reactions, and comments on Facebook. Within the same time period, the 20 best-performing election stories from 19 major news websites generated a total of 7,367,000 shares, reactions, and comments on Facebook.

Information Disorder

Why?

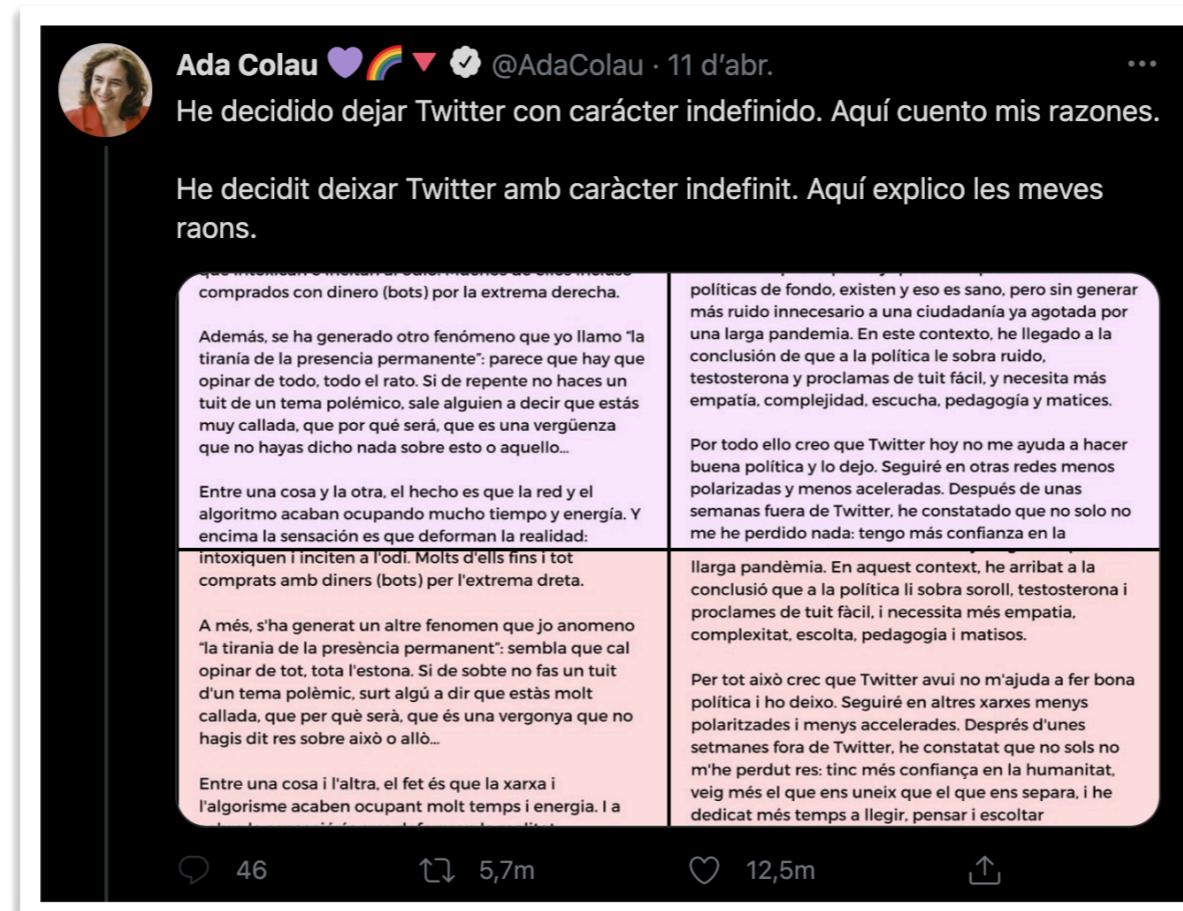
- Modern information environment is widely **accessible, cheap** and sophisticated editing and publishing technology has made it easier than ever for anyone to create and distribute content.
- The **speed** at which information is disseminated has been supercharged by an accelerated news cycle and mobile handsets.
- Popular **social networks** make it difficult for people to judge the credibility of any message, because posts from publications as unlike as the New York Times and a conspiracy site look nearly identical.

“What we see unfolding right before our eyes is nothing less than Moore’s Law applied to the distribution of mis-information:
an exponential growth of available technology coupled with a rapid collapse of costs.”

Frederic Filloux

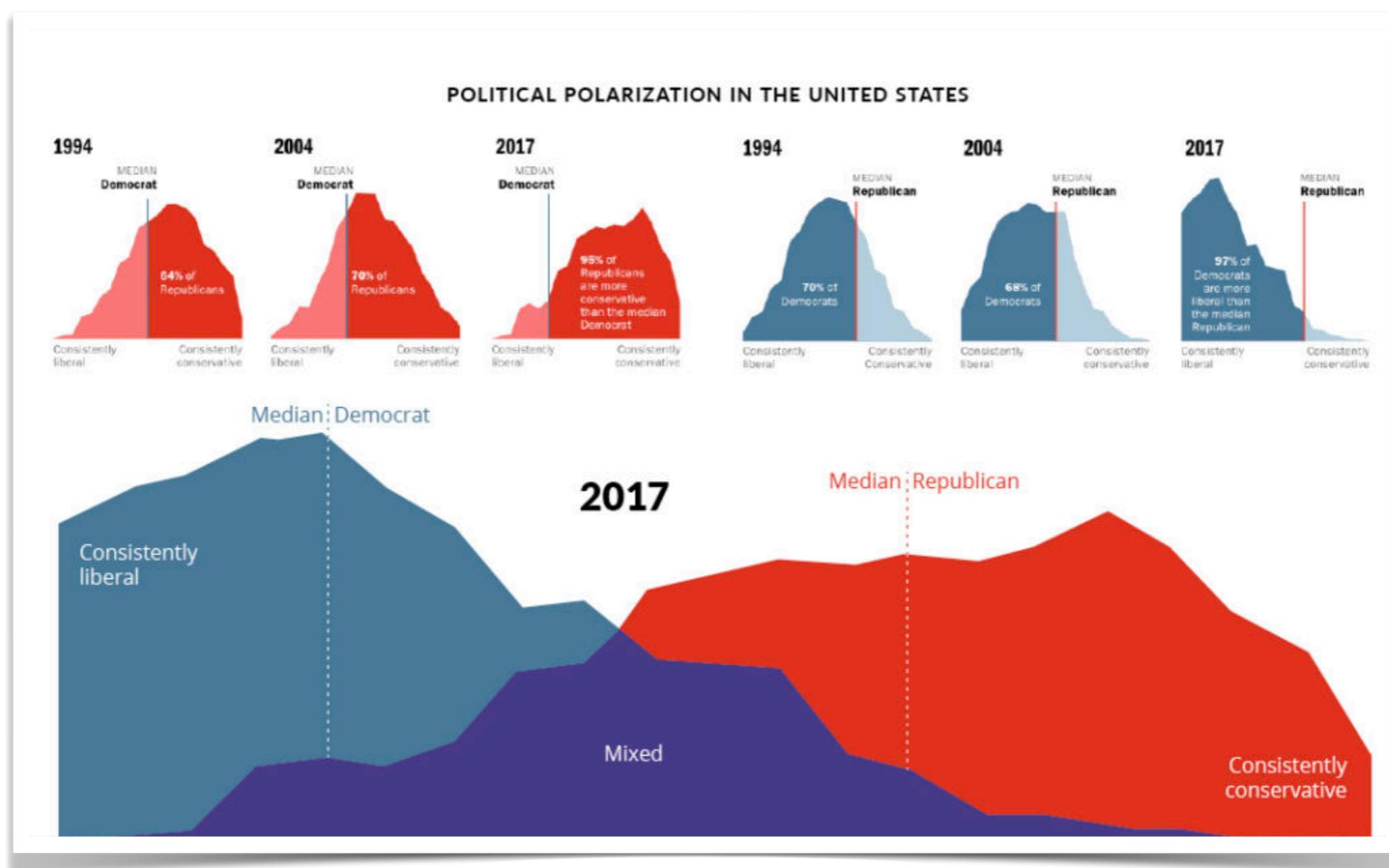
Information Disorder

We have to look for explanations for how **societies**, particularly in the West, have become so **segregated** in terms of terms of age, race, religion, class and politics.



Information Disorder

We have to look for explanations for how **societies**, particularly in the West, have become so **segregated** in terms of terms of age, race, religion, class and politics.



Communication Theory

The **transmission view of communication** is the commonest in our culture – perhaps in all industrial cultures... It is defined by terms such as ‘**imparting**,’ ‘**sending**,’ ‘**transmitting**,’ or ‘**giving information to others**’.

The **ritual view of communication**, by contrast, is not about the act of imparting information but the representation of shared beliefs.

Under a transmission view of communication, one sees the newspaper as an instrument for disseminating knowledge. However, a ritual view of communication does not consider the act of reading a newspaper to be driven by the need for new information. Rather, it likens it to attending a church service.

Considering information consumption and dissemination from merely the transmission view is unhelpful as we try and understand information disorder.

Conceptual Framework

In trying to understand any example of information disorder, it is useful to consider it in three elements:

- 1) **Agent.** Who were the ‘agents’ that created, produced and distributed the example, and what was their motivation?
- 2) **Message.** What type of message was it? What format did it take? What were the characteristics?
- 3) **Interpreter.** When the message was received by someone, how did they interpret the message? What action, if any, did they take?

Conceptual Framework

It is also productive to consider the life of an example of information disorder as having three phases:

- 1. Creation.** The message is created.
- 2. Production.** The message is turned into a media product.
- 3. Distribution.** The message is distributed or made public.

The agent that creates the content is often fundamentally different from the agent who produces it.

For example, the motivations of the mastermind who ‘creates’ a state-sponsored dis-information campaign are very different from those of the low-paid ‘trolls’ tasked with turning the campaign’s themes into specific posts.

And once a message has been distributed, it can be reproduced and redistributed endlessly, by many different agents, all with different motivations.

Example

WTOE 5 NEWS
YOUR LOCAL NEWS NOW

[TOP STORIES](#) [COMMUNITY](#) [ENTERTAINMENT](#) [SPORTS](#) [LIFE](#) [ABOUT](#) [LATEST NEWS](#) [VIDEOS](#) [ARCHIVE](#)

Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement

TOPICS: Pope Francis Endorses Donald Trump



Example

Agent						
Type of Actor	Motivation	Level of organization	Use of automation	Intended audience	Intent to harm	Intent to mislead
<i>Unofficial actor</i>	<i>Financial</i>	<i>Connected to a network of sites</i>	<i>No</i>	<i>US voters</i>	<i>Yes</i>	<i>Yes</i>
Message						
Format	Durability	Level of accuracy	Imposter	Legality	Target of message	
<i>Text</i>	<i>Election campaign</i>	<i>Fabricated</i>	<i>No</i>	<i>Yes</i>	<i>Hillary Clinton</i>	
Interpreter						
How is the message interpreted by different people?	What action is taken? Is the message re-produced?					

<https://www.coe.int/en/web/freedom-expression/information-disorder>

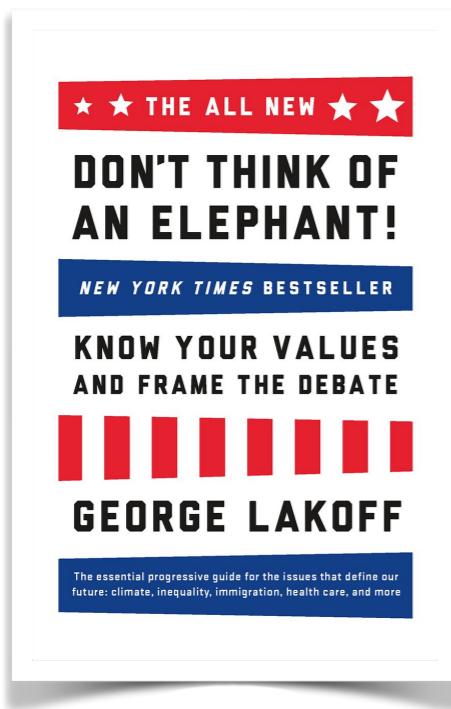
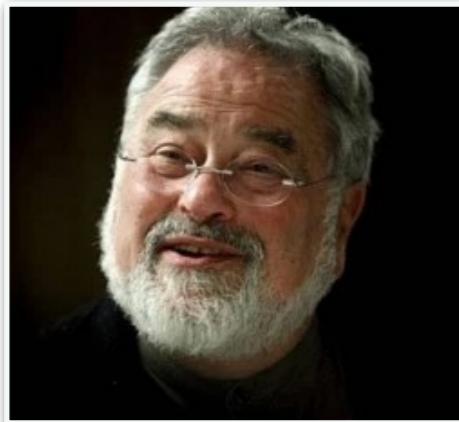
About the Interpreters

How do they make sense of the messages?

Messages are decoded by individual audience members in one of three ways:

1. **Hegemonic.** Accepting the message as it was encoded.
2. **Negotiated.** Accepting aspects of the message, but not all of it.
3. **Oppositional.** Declining the way the message was encoded.

About the Interpreters



George Lakoff sees rationality and emotions as being tied together to the extent that, as human beings, **we cannot think without emotions**.

The emotions in our brains are structured around certain metaphors, narratives and frames. They help us make sense of things, and, without them, we would become disoriented. We would not know what or how to think.

Lakoff distinguishes **two different kinds of reason**: 'False reason' and 'real reason.'

False reason sees reason as fully conscious, as literal, **disembodied**, yet somehow fitting the world directly, and working not via frame-based, metaphorical, narrative and emotional logic, but via the **logic of logicians** alone.

Real reason, alternatively, is an unconscious thought that '**arises from embodied metaphors**'.

He argues that **false reason does not work in contemporary politics**, as we've become increasingly **emotional** about our political affiliations.

About the Interpreters

Similarly, Maffesoli argues that to understand someone's behavior, one must **consider** the sociological implications of the many different, **small and temporary groups that he or she is a member of at any given time of day.**

This tribal mentality partly explains why **many social media users distribute dis-information when they don't necessarily trust the veracity of the information** they are sharing: they would like to conform and belong to a group, and they 'perform' accordingly.

About the Interpreters

Human memory is not a recording device, but rather a **process of (re)construction** that is vulnerable to both internal and external influences.

The challenge for the human brain today is how these (high-frequency) influences work in the context of social networks that are bombarding us with information.

“Our minds have always been a battleground for various social forces, but the sheer number of agents and institutions vying for control of our thoughts and feelings today is so large that it is confusing and destabilising for many.”

Wikimedia testimony at the UK Parliament

About the Interpreters

People use a set of key heuristics, or mental shortcuts, when evaluating the **credibility** of a source or message:

1. **Reputation.** Based on recognition and familiarity
2. **Endorsement.** Whether others find it credible
3. **Consistency.** Whether the message is echoed by multiple sites
4. **Expectancy violation.** Whether a site looks and behaves in the expected manner.
5. **Self-confirmation.** Whether a message confirms one's beliefs
6. **Persuasive intent.** The intent of the source in creating the message.

Filter Bubbles and Echo Chambers

‘The “**public sphere**” is the of the shared spaces – real, virtual or imagined – whereby social issues are discussed and public opinion is formed .

A healthy public sphere is essential for democracy and must be inclusive, representative and characterised by respect for rational argument.

The most significant challenge to any theory of a shared public sphere is that **humans**, when we have a choice about who to connect with or not, **tend to establish and continue relationships with people who have views similar to our own.**

We are programmed to enjoy spending time in ‘echo chambers,’ as it requires less cognitive work.

Filter Bubbles and Echo Chambers

Agents who are creating dis-information understand that, when people consume and share these messages, they will be doing so increasingly from inside these echo chambers, with no one to challenge the ideas.

This means the people who will interpret their messages are much less likely to have an ‘oppositional’ (rejecting the way the message was encoded) or ‘negotiated’ (accepting only some aspects of the message) reading.

The fundamental problem is that “filter bubbles” worsen polarization by allowing us to live in our own online echo chambers and leaving us with only opinions that validate, rather than challenge, our own ideas.

Filter Bubbles and Echo Chambers

Blue Feed, Red Feed

See Liberal Facebook and Conservative Facebook, Side by Side

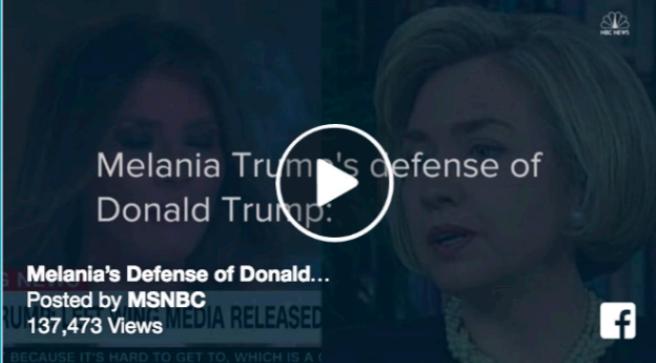
[Twitter](#) [Facebook](#)

By **Jon Keegan**
Published May 18, 2016 at 8:00 a.m. ET | Updated hourly

FILTER FEEDS BY TOPIC:

HILLARY CLINTON DONALD TRUMP BARACK OBAMA GUNS ABORTION ISIS PRESIDENTIAL DEBATE

LIBERAL ⓘ



Melania Trump's defense of Donald Trump:
Melania's Defense of Donald...
Posted by **MSNBC**
137,473 Views

BECUSE IT'S HARD TO GET TO, WHICH IS A

Melania Trump recently defended Donald J. Trump's lewd comments in an interview with CNN.
Watch how similar they are to Hillary Clinton's past defenses of her husband: <http://on.msnbc.com/2dly9Ki>

477 640 481

SHOWING POSTS ABOUT:
"HILLARY CLINTON"



Western Journalism ⓘ
about an hour ago

The latest email leak is an absolute nightmare for Hillary Clinton. Right after she was complaining about the so-called 'gender pay gap,' she was slammed with a cold hard truth that exposes one of her darkest secrets.



As Hillary Rails Against 'Gender Pay Gap' Leake...
"There are huge discrepancies ..."

Wall Street Journal's "Blue Feed, Red Feed" shows just how different the echo chambers can be.

Filter Bubbles and Echo Chambers

AllSides *Don't be fooled by media bias and fake news.*
Unbiased news does not exist; we provide balanced news and civil discourse.

Donate Join

NEWS MEDIA BIAS TOPICS Q BALANCED SEARCH TALK DICTIONARY SCHOOLS ABOUT

Balanced News Facts & Fact Checking Perspectives Blog Story of the Week Editorial Philosophy Apps & Extensions

TOP STORIES: COVID-19 | Vaccines | Recovery | Immigration | George Floyd | President Biden | Election Laws | School Reopenings

 **Perspectives: Earth Day 2021**
April 22 is Earth Day, which is a day intended to demonstrate support for environmental preservation. Earth Day was founded in 1970 by nonprofit group Environmental Teach-In, Inc., and...
From the Center U.S., other countries deepen climate goals at Earth Day summit Reuters L L E R R
From the Left Four steps this Earth Day to avert environmental catastrophe The Guardian L L E R R
From the Right On Earth Day, youth activists can change the narrative Washington Examiner L L C R R

Get the AllSides Mobile App
Balanced news on the go. Available for download on [iOS](#) and [Android](#).  

Story of the Week
Email Address [Sign Up](#)
Get balanced news in your inbox weekly.

Follow AllSides
[f](#) [Twitter](#) [Instagram](#) [RSS](#) @AllSidesNow

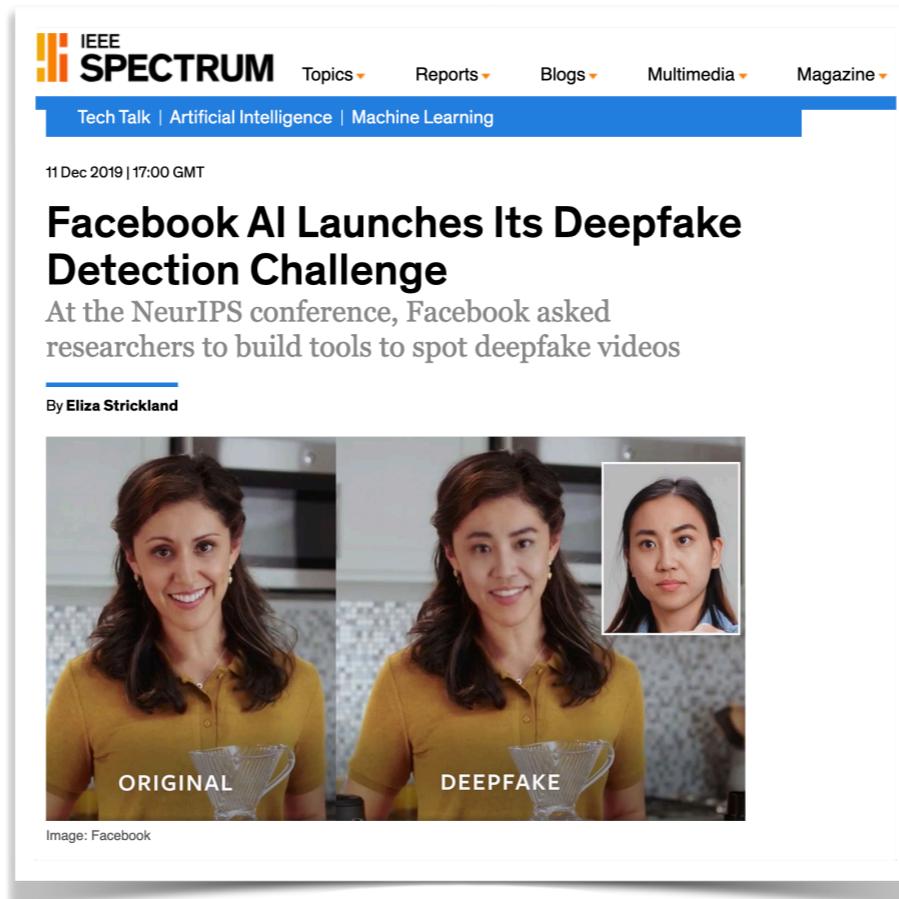
 **U.S. Senate Votes to Pass Asian American Hate Crimes Bill**
On Thursday, the U.S. Senate voted 94-1 to pass the COVID-19 Hate Crimes Act, aimed at combatting racism toward Asian American and Pacific Islander communities. The bill would boost hate...
From the Left The Senate strongly condemns anti-Asian hate crimes by passing new bill Vox L L E R R
From the Center In Rare Moment Of Bipartisan Unity, Senate Approves Asian American Hate Crimes Bill NPR (Online News) L L E R R
From the Right Senate Passes Sweeping Bill Aiming To Crack Down On Hate Crimes The Daily Caller L L E R R

 **Police Killing of Ma'Khia Bryant, Who Attacked 2 With Knife, Sparks Outcry, Media Criticism**
Just twenty minutes before Derek Chauvin's guilty verdict was announced on Tuesday, a police officer in Columbus, Ohio fatally shot Ma'Khia Bryant, a 16-year-old black girl in...
From the Center Ma'Khia Bryant: Columbus police release body camera footage in shooting of 16-ye... The Columbus Dispatch L L E R R
From the Left Ohio officials release more body cam video of fatal police shooting of Black teen and... CNN (Online News) L L E R R
From the Right 'NBC Nightly News' under fire for editing out key part of 911 call before Ma'Khia Bryant... Fox News (Online News) L L C R R

Attempts at solutions

Much of the debate about solutions presupposes communication as information transmission. But this cannot explain or solve the problem of information disorder.

From this point of view, **technological solutions** can address only a part of the problem.



Attempts at solutions

Blacklisting, flagging and credibility scores: attempting to be the ‘arbiter of truth’ is not very effective.

Stronger Media: trust in mainstream media has been falling for decades, as has trust in other public institutions.



https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FINAL.pdf

Attempts at solutions

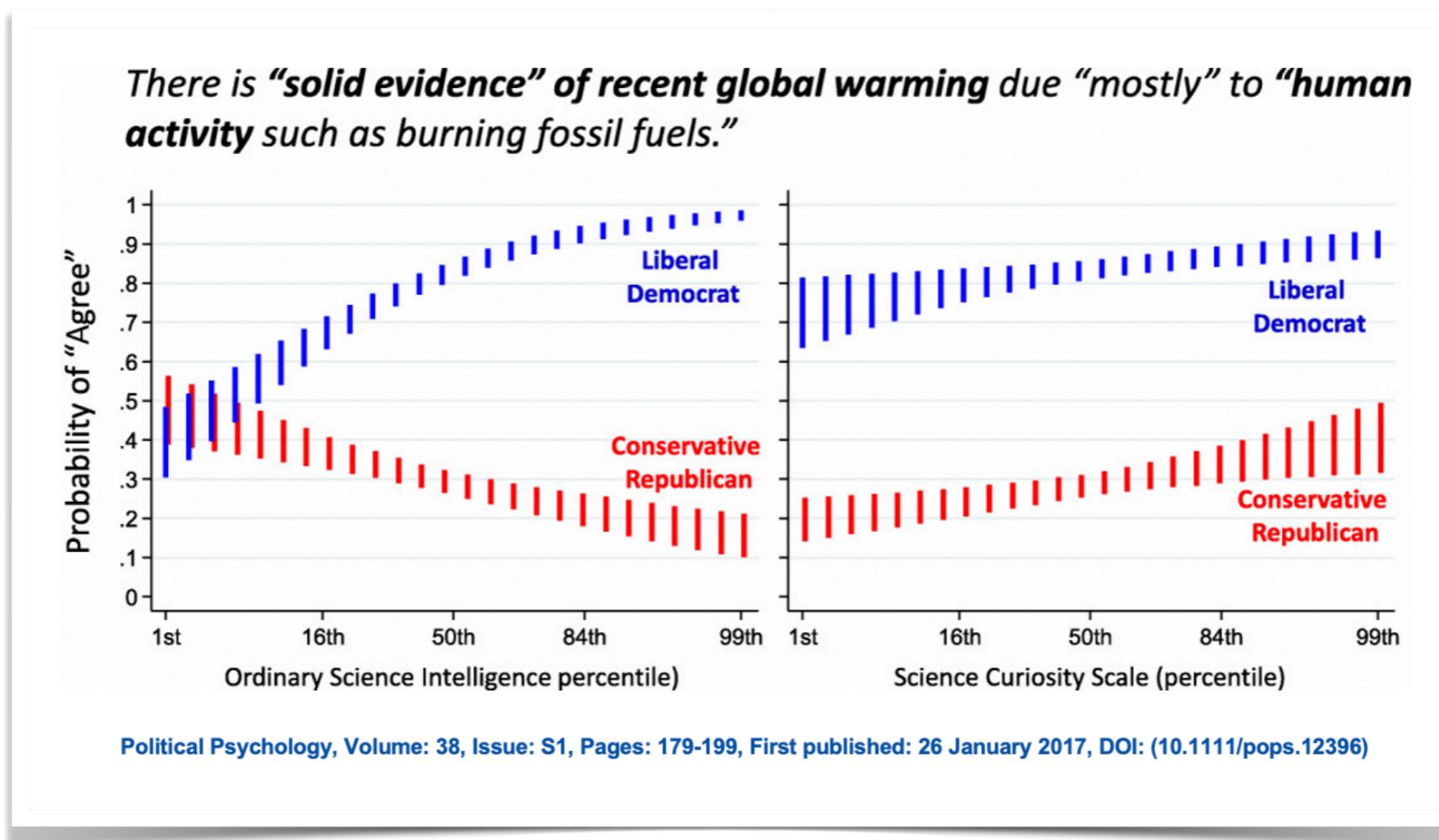
Strategic Silence. For manipulators, it doesn't matter if the media is reporting on a story in order to debunk or dismiss it; the important thing is getting it covered in the first place.

Regulation. We are starting to see legislation directed at information disorder.

Attempts at solutions

Education. There is a need to educate people on the power of images to manipulate and persuade. Programs that have focused on critical thinking, source evaluation and emotional manipulation have seen success.

The graph of despair: political views are correlated with your views on climate change





Additional Resources

NEW YORK TIMES BESTSELLER

THE FILTER BUBBLE

"Well-timed ... a powerful indictment."—*The Wall Street Journal*

HOW THE [REDACTED] NEW
PERSONALIZED [REDACTED]
WEB IS [REDACTED] 
[REDACTED] CHANGING WHAT
WE READ [REDACTED] AND
[REDACTED] HOW WE THINK
ELI PARISER

"Vital."—*Time.com*