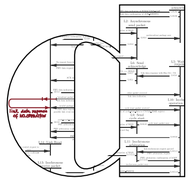
Ethics for Nerds

An Advanced Course in Computer Science
Summer Semester 2020

Current Topics 2.1



Prof. Holger Hermanns,
Kevin Baum, Sarah Sterz

UNIVERSITÄT
DES
SAARLANDES

A FILTERED EXPIERIENCE

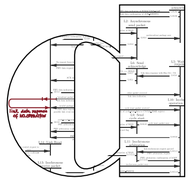
A FILTERED EXPIERIENCE

Ethics for Nerds

An Advanced Course in Computer Science
Summer Semester 2020

Current Topics 2.1
Filter Bubbles

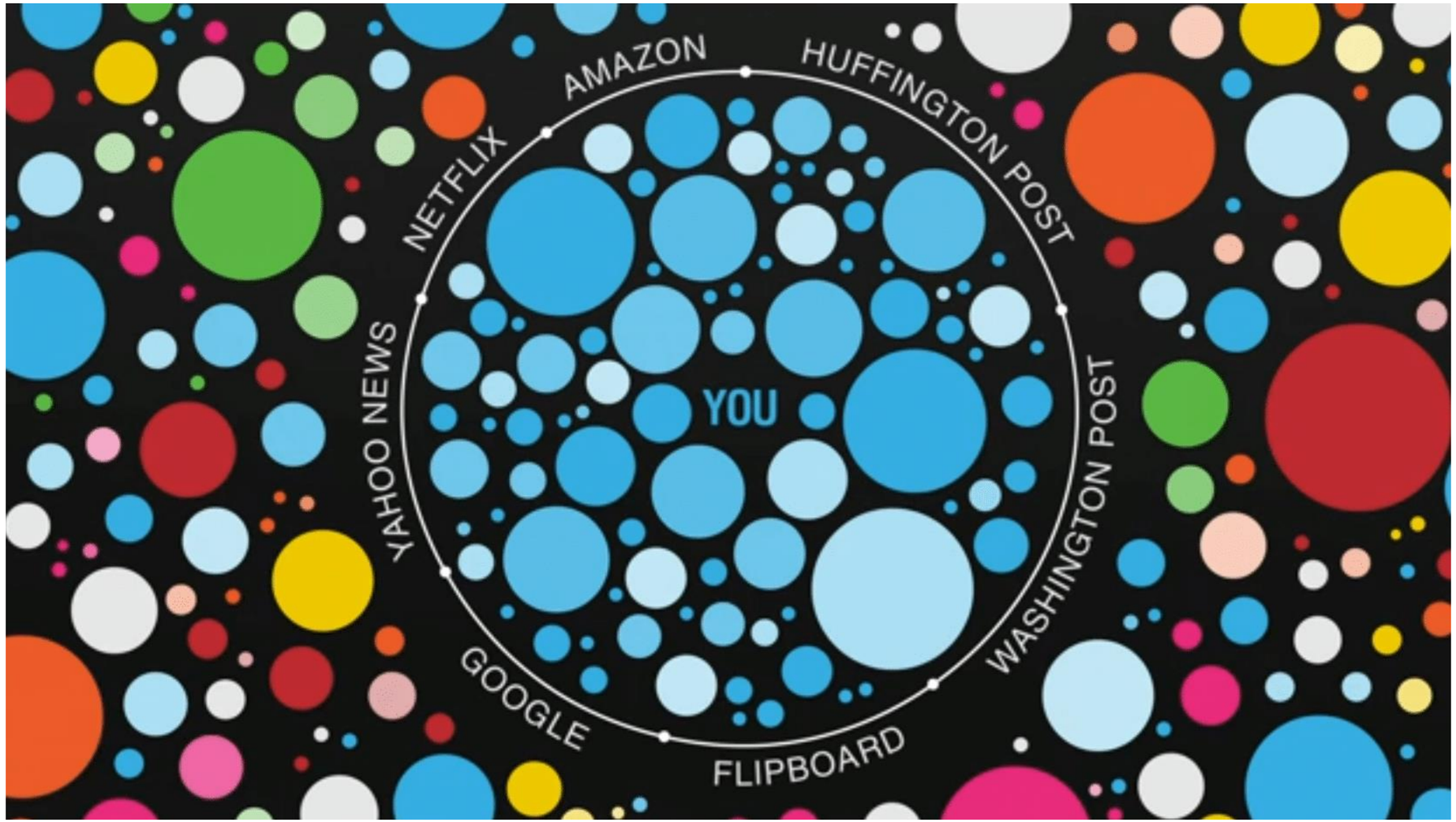
What are they (not)? And are they morally problematic?



Prof. Holger Hermanns,
Kevin Baum, Sarah Sterz

UNIVERSITÄT
DES
SAARLANDES

FILTER BUBBLE



From Eli Pariser's 2011 TED Talk:
https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?language=de

FILTER BUBBLE VS ECHO CHAMBERS

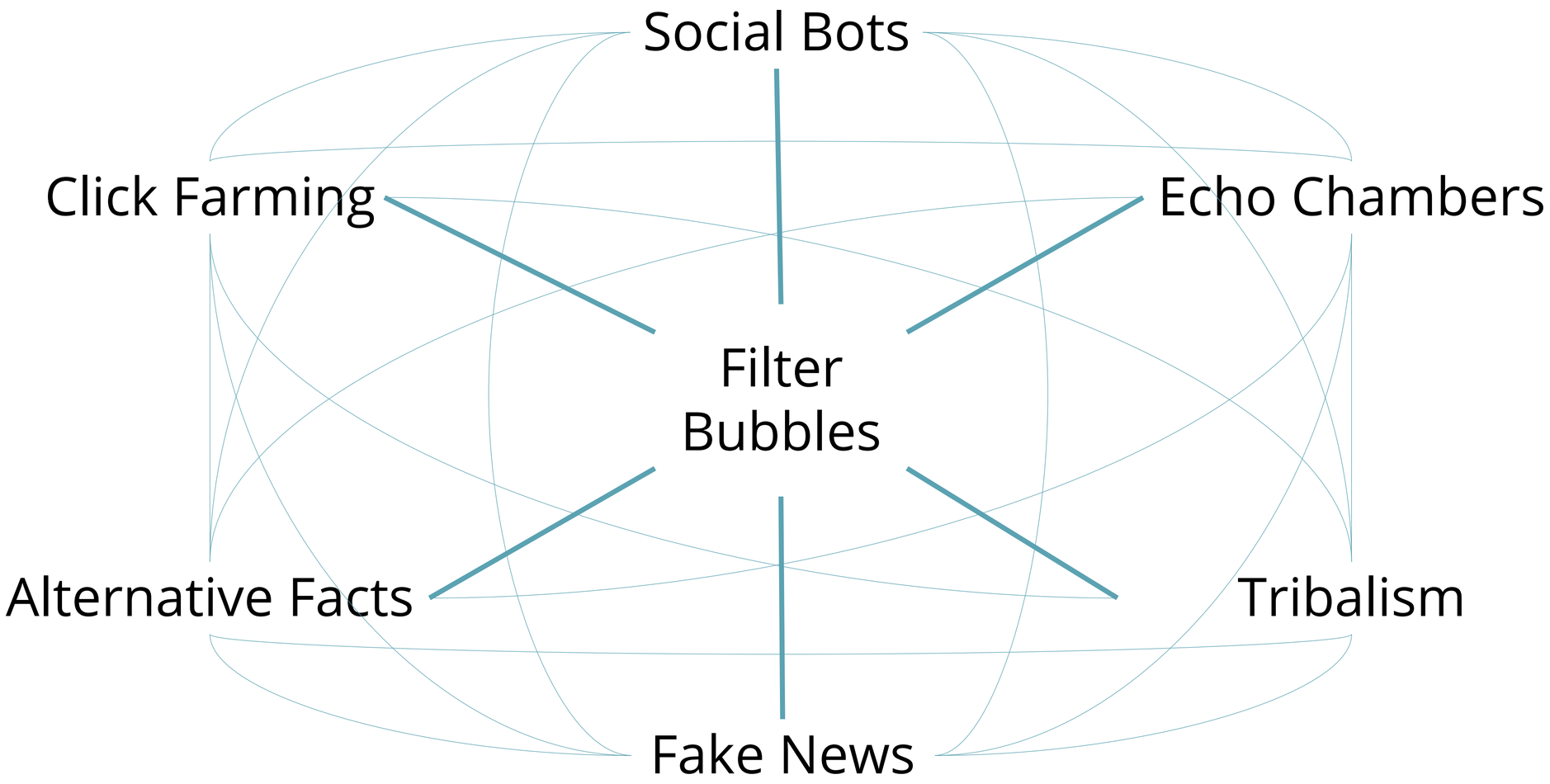
(not necessarily intentionally)

An 'epistemic bubble' [or 'filter bubble'] is an informational network from which relevant voices have been excluded by omission. […] An 'echo chamber' is a social structure from which other relevant voices have been actively discredited.

C Thi Nguyen in *Escape the echo chamber*

https://aeon.co/essays/why-its-as-hard-to-escape-an-echo-chamber-as-it-is-to-flee-a-cult

NEXUS


Social Bots
Click Farming
Echo Chambers
Filter Bubbles
Alternative Facts
Tribalism
Fake News

POSSIBLE PROBLEMS

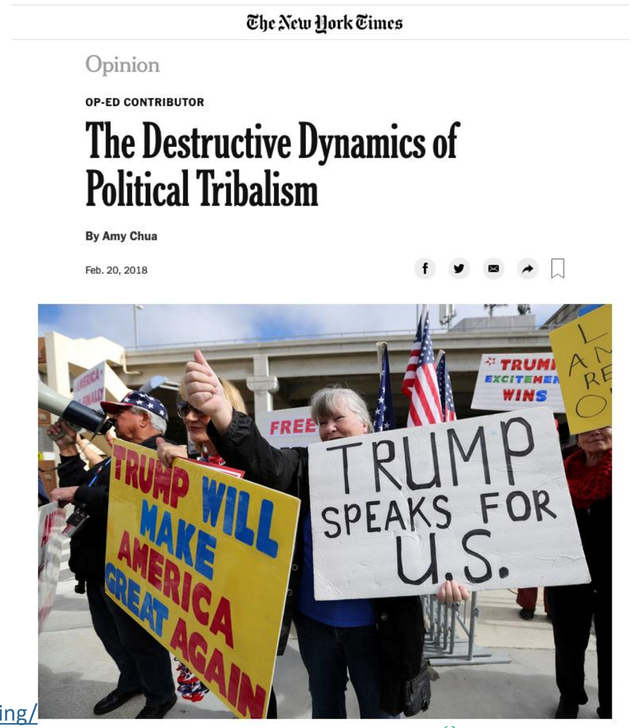
Filter bubbles & echo chambers…

- support wrong beliefs and believes in conspiracy theories/myths;
- undermine democracy and public discourse;
- strengthen polarization;
- allow to manipulate behaviour on large scale;
- undermining autonomy
- …

https://commons.wikimedia.org/wiki/File:30C3_TinFoil-Hat_02.jpg

Illustration: Sam Whitney; Getty Images

https://www.washingtonpost.com/graphics/2017/business/russian-ads-facebook-targeting/



https://www.nytimes.com/2018/02/20/opinion/destructive-political-tribalism.html

REASONABLE DOUBT

https://www.wired.com/story/polarization-politics-misinformation-social-media/

WIRED BUSINESS CULTURE GEAR IDEAS SCIENCE SECURITY TRANSPORTATION

GIDEON LEWIS-KRAUS BUSINESS 01.15.2020 06:00 AM

Bad Algorithms Didn't Break Democracy

And better ones won't save it. To get past misinformation and tribal rancor online, we need to face why people really want misinformation and rancor.

OVER THE PAST **five decades,** America's war on drugs has been motivated and organized by the fantasy that the proliferation of substance abuse is fundamentally a supply problem. The remedy, accordingly, has been to restrict the production and distribution of narcotics: Smash the cartels, cauterize the trafficking routes, arrest the dealers. This approach has, predictably enough, devolved into a self-sustaining game of whack-a-mole.

Since 2016, the panic about misinformation online has been driven by a similar fantasy. The arguments predicated on this view have become familiar, almost boilerplate. One recent example was a November speech given by the comedian Sacha Baron Cohen.

"Today around the world, demagogues appeal to our worst instincts. Conspiracy theories once confined to the fringe are going mainstream," said the actor, in a rare performance in character as himself. "It's as if the Age of Reason—the era of evidential argument—is ending, and now knowledge is increasingly delegitimized and scientific consensus is dismissed. Democracy, which depends on shared truths, is in retreat, and autocracy, which depends on shared lies, is on the march." As Baron Cohen put it, it's "pretty clear" what's behind these trends: "All this hate and violence is being facilitated by a handful of internet companies that amount to the greatest propaganda machine in history."

Most Popular

SECURITY

Trump's New Intelligence Chief Spells Trouble

GARRETT M. GRAFF

SCIENCE

How to Watch SpaceX Launch Astronauts to the ISS

DANIEL OBERHAUS

BUSINESS

Twitter Finally Fact-Checked Trump. It's a Bit of a Mess

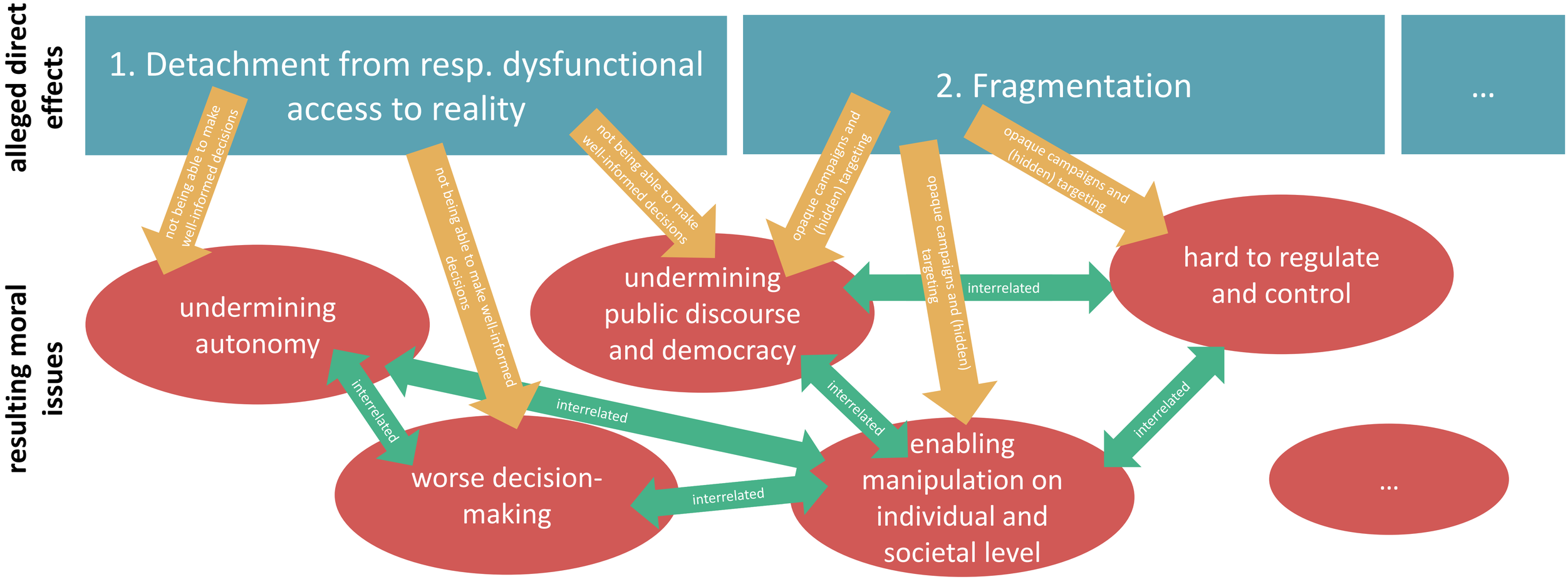
GILAD EDELMAN

SCIENCE

Trump's New Space Force Missile Might Be Too 'Super-Duper'

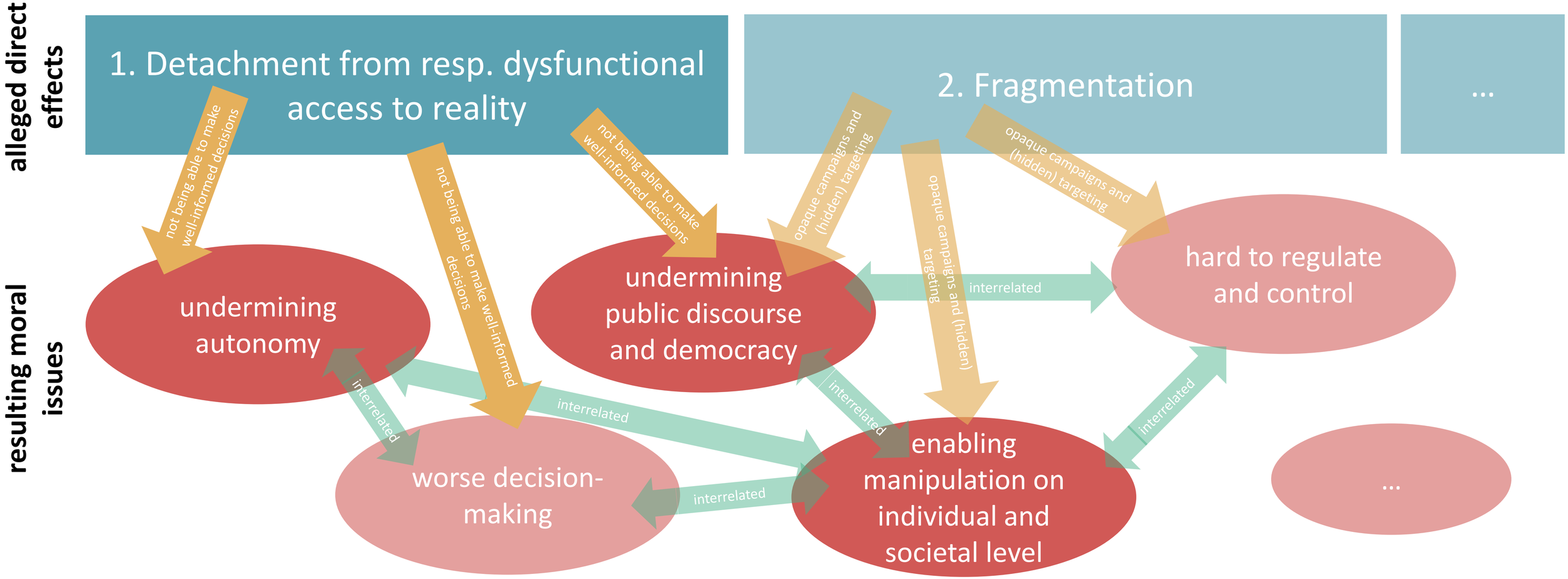
POSSIBLE PROBLEMS: MORE STRUCTURE

A complex landscape…



POSSIBLE PROBLEMS: MORE STRUCTURE

A complex landscape…



PLAN

- Today:

1. Detachment from resp. dysfunctional access to reality → not being able to make well-informed decisions

- Later:

undermining autonomy

enabling manipulation on individual and societal level

CENTRAL QUESTION

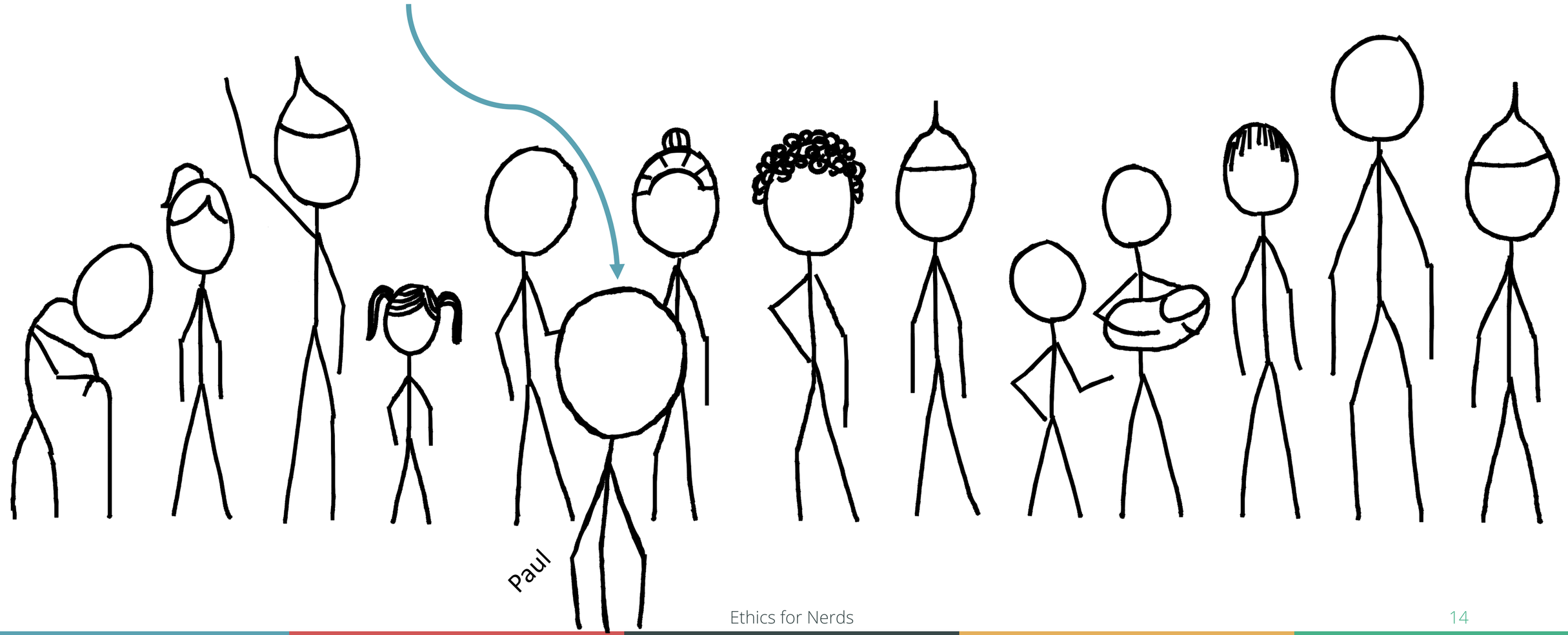
1. Detachment from resp. dysfunctional access to reality → not being able to make well-informed decisions

How is it that filter bubbles

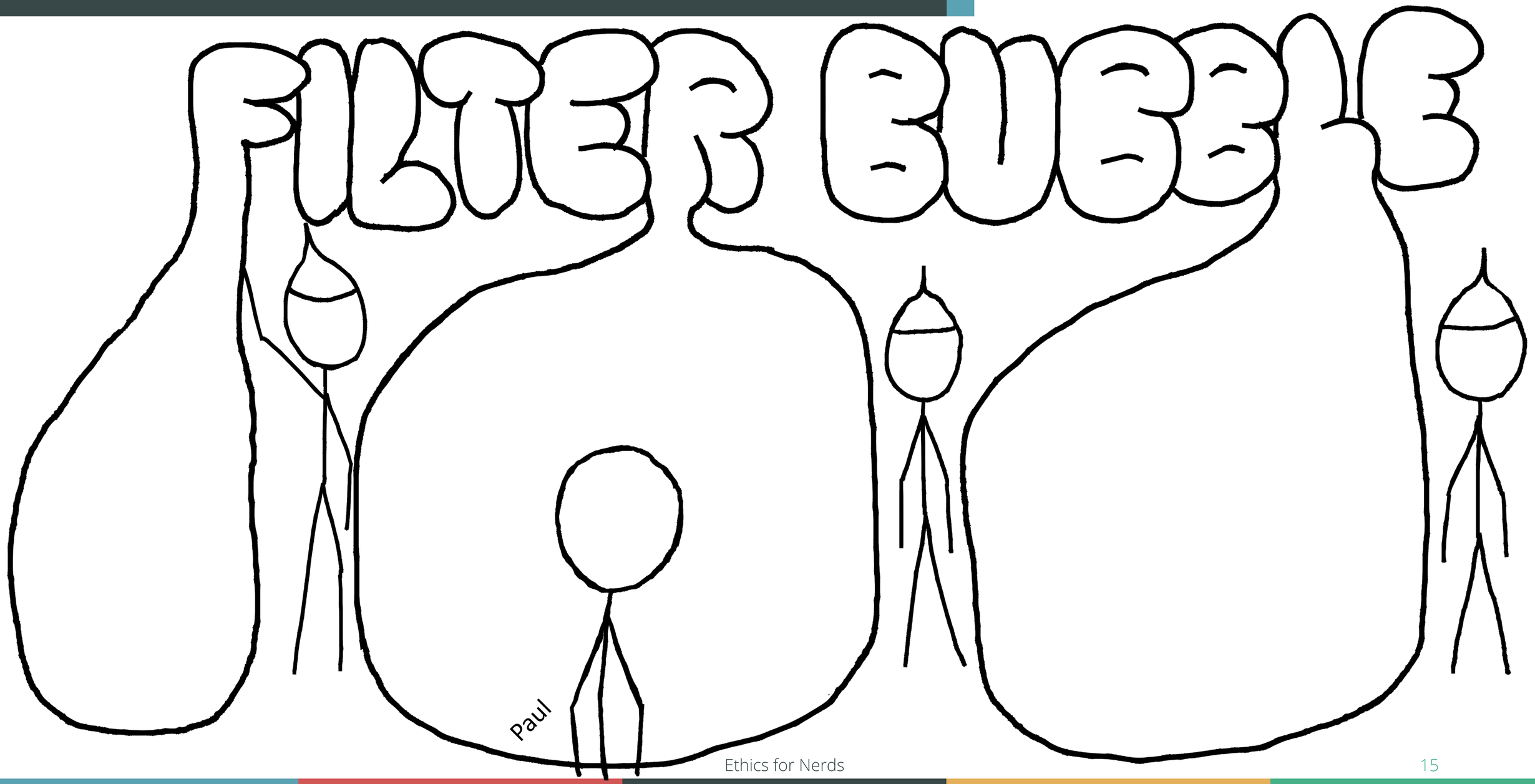
- detach people from reality,
- render their access to that reality dysfunctional and, through this,
- deprive people from making well-informed decisions?

GENERAL IDEA

Putting a tin foil hat on his head (like a minority does)
seems to be pretty unreasonable to Paul, but…



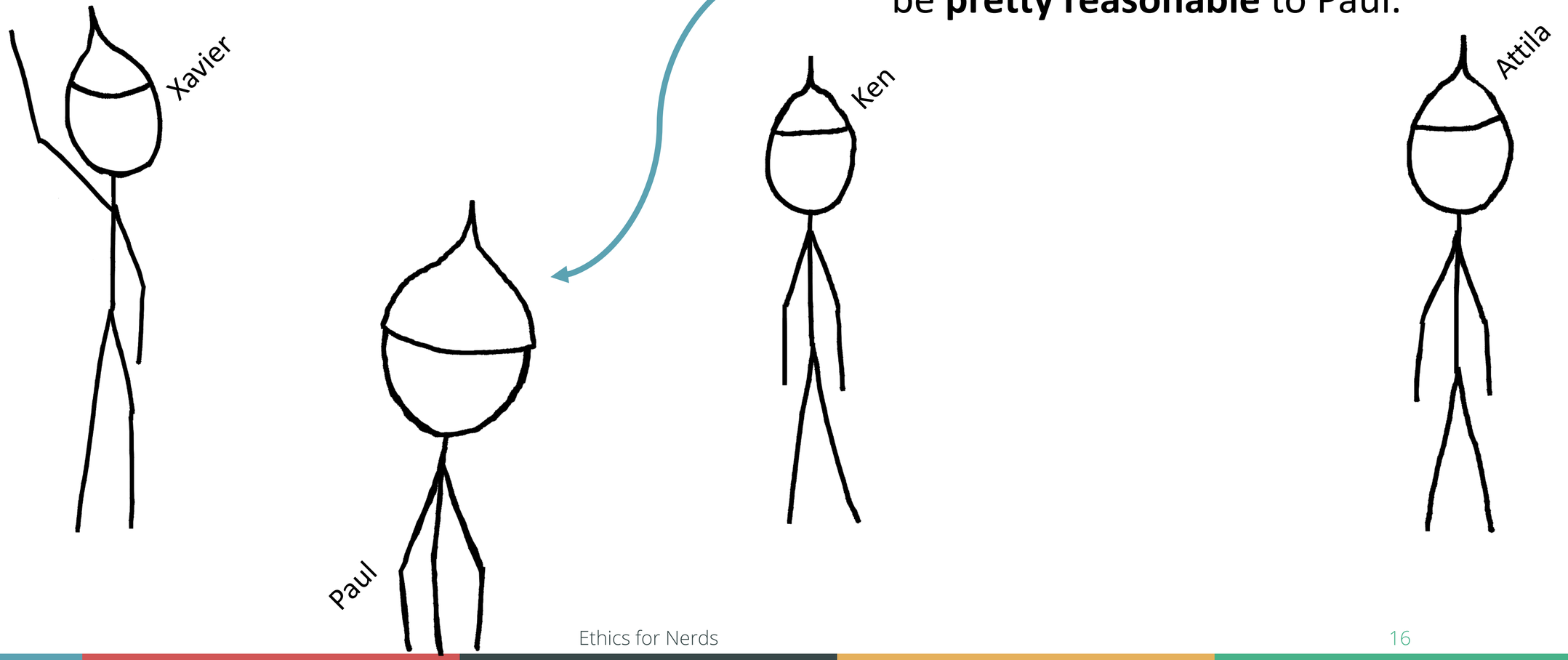
GENERAL IDEA



GENERAL IDEA

Before we present an argument for that claim, we need some preparations…

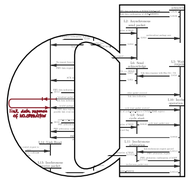
Putting a tin foil hat on his head **seems** to be **pretty reasonable** to Paul.



Ethics for Nerds

An Advanced Course in Computer Science
Summer Semester 2020

Current Topics 2.2
Filter Bubbles: Where do they come from?
Philosophical Groundwork: Epistemology in a nutshell



Prof. Holger Hermanns,
Kevin Baum, Sarah Sterz

UNIVERSITÄT
DES
SAARLANDES

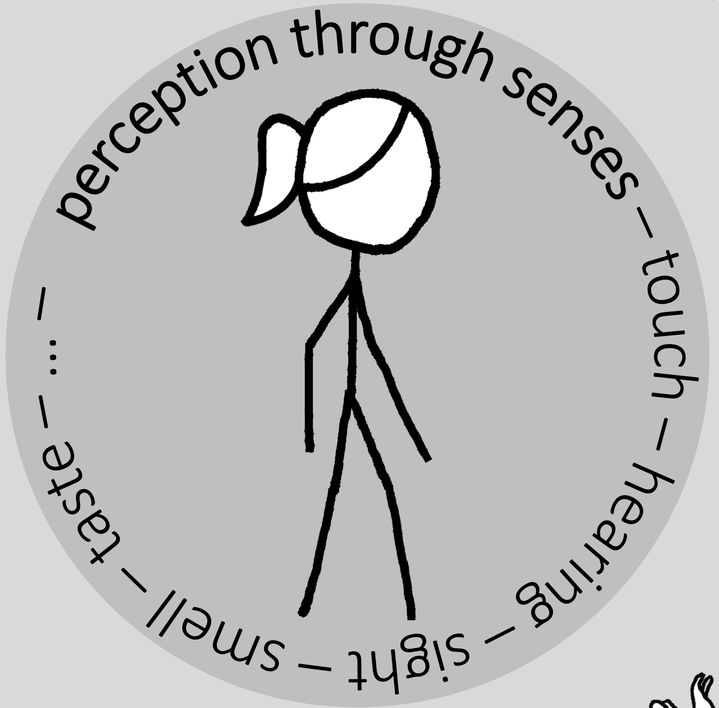
THE PATH TO WHERE WE ARE

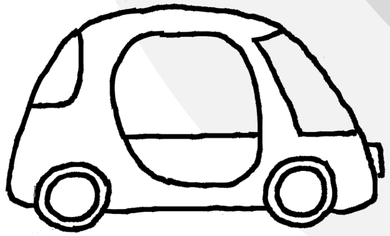
How could it come to these filter bubbles?

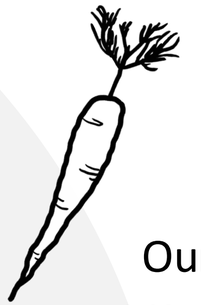
ACCESS TO THE WORLD: THE GOOD OLD TIMES

If you look out into the world, you see what is in your “reach.”

How does that affect (the correctness of) your representation of the world?

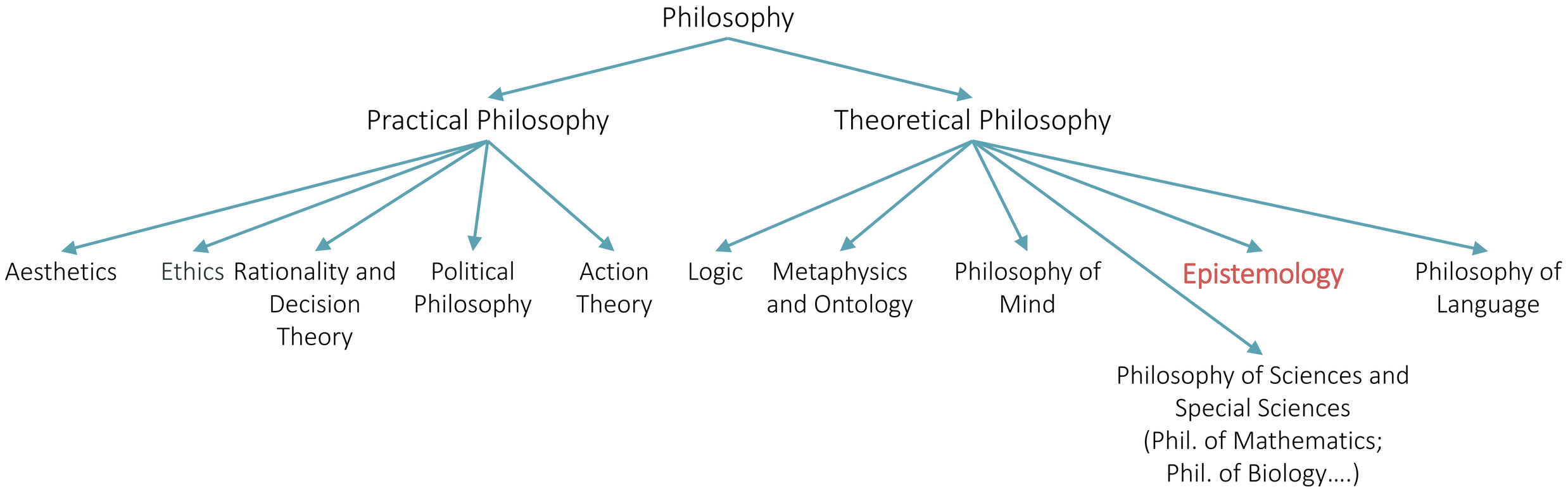






Our perceptions seem to ground the justification$_{epis}$ of our beliefs about the external world in such cases (and under normal circumstances).

Fields of Philosophy



EPISTEMOLOGY

https://plato.stanford.edu/entries/epistemology/

Epistemology: the study of knowledge and justified belief

- Main question:
  - What is *knowledge* and how can we *achieve* it?
- Epistemologists ask for example:
  - What is *knowledge*?
  - What is *justification*?
  - What are *conditions for rational* belief?
  - What are *reasons* to belief?
  - What is *evidence*?
  - And: **What are *sources* of knowledge and justification?**

    Candidates:
    Perception, *introspection*, *memory*, *reason*, ***testimony***?

Example

Knowledge – Classical Definition

Knowledge is justified true belief.

Testimony

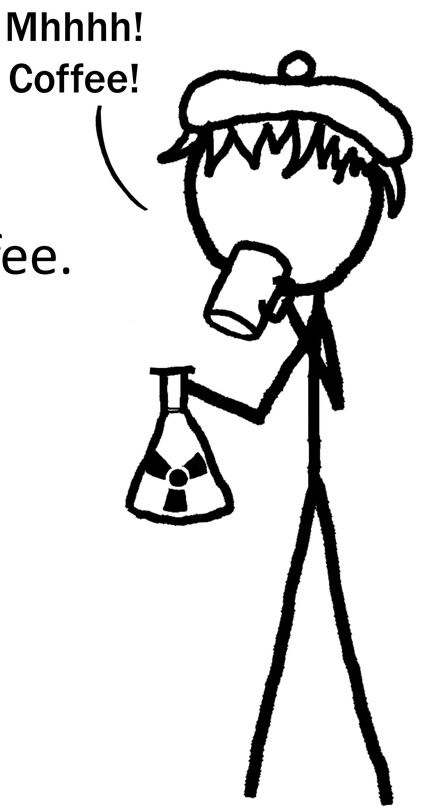
The intentional transfer of a belief from one person to another.

INTERNAL VS. EXTERNAL JUSTIFICATION

Beliefs & Propositions

Daniel, the crazy researcher,

believes that the liquid in the Erlenmeyer flask is delicious and strong coffee.



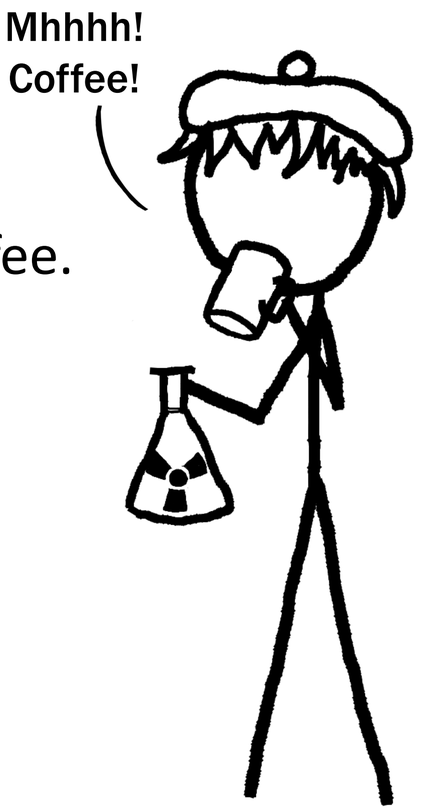
INTERNAL VS. EXTERNAL JUSTIFICATION

Beliefs & Propositions

Daniel, the crazy researcher,

believes that the liquid in the Erlenmeyer flask is delicious and strong coffee.

(Epistemic) Subject



INTERNAL VS. EXTERNAL JUSTIFICATION

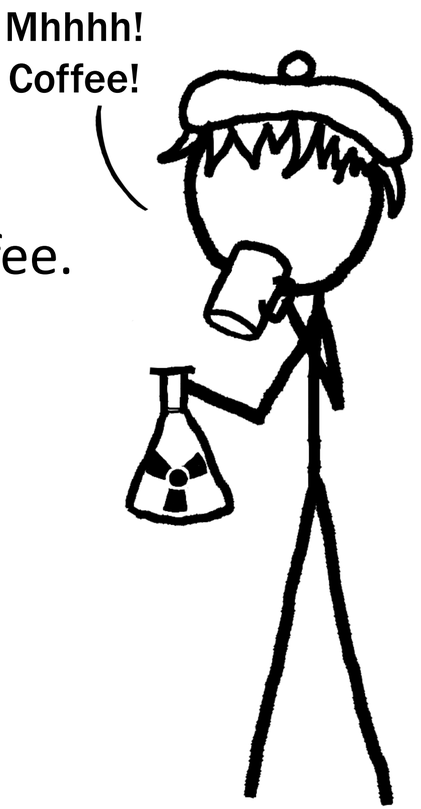
Beliefs & Propositions

Daniel, the crazy researcher,

believes that the liquid in the Erlenmeyer flask is delicious and strong coffee.

(Epistemic) Subject

Propositional Attitude (in this case: believing)



INTERNAL VS. EXTERNAL JUSTIFICATION

Beliefs & Propositions

Daniel, the crazy researcher,

believes that the liquid in the Erlenmeyer flask is delicious and strong coffee.

(Epistemic) Subject

Propositional Attitude (in this case: believing)

Proposition (object of belief)

EXCURSUS: PROPOSITIONS

Proposition (very rough working definition)

The factual content expressed by a declarative sentence on a particular occasion and the content of representational mental state (aka beliefs).

Different sentences can express the same proposition:

- “After the lecture, I will have dinner.” and “I will have dinner after the lecture.”
- “It’s quarter to six.” and “It’s 5:45.”
- “Saarbrücken is a town in Germany.” and “Saarbrücken ist eine Stadt in Deutschland.”

The same sentence can express different propositions, depending on the context:

- „The latest season of Game of Thrones was disappointing.“ (depending on time)
- „I am …“ (depending who says that [this is an example of so-called indexical])

Not every sentence expresses a proposition.

- „Are you there?“
- „Oh, no!“
- „Please go.“

INTERNAL VS. EXTERNAL JUSTIFICATION

Epistemic Justification (Justification$_{epis}$)

One central epistemological question: What kind of conditions can justify Daniel's belief? External or internal ones?

Daniel, the crazy researcher,

believes that the liquid in the Erlenmeyer flask is delicious and strong coffee.

Internalism (rough working definition)

All justification of a believer's beliefs are internal.

Externalism (rough working definition)

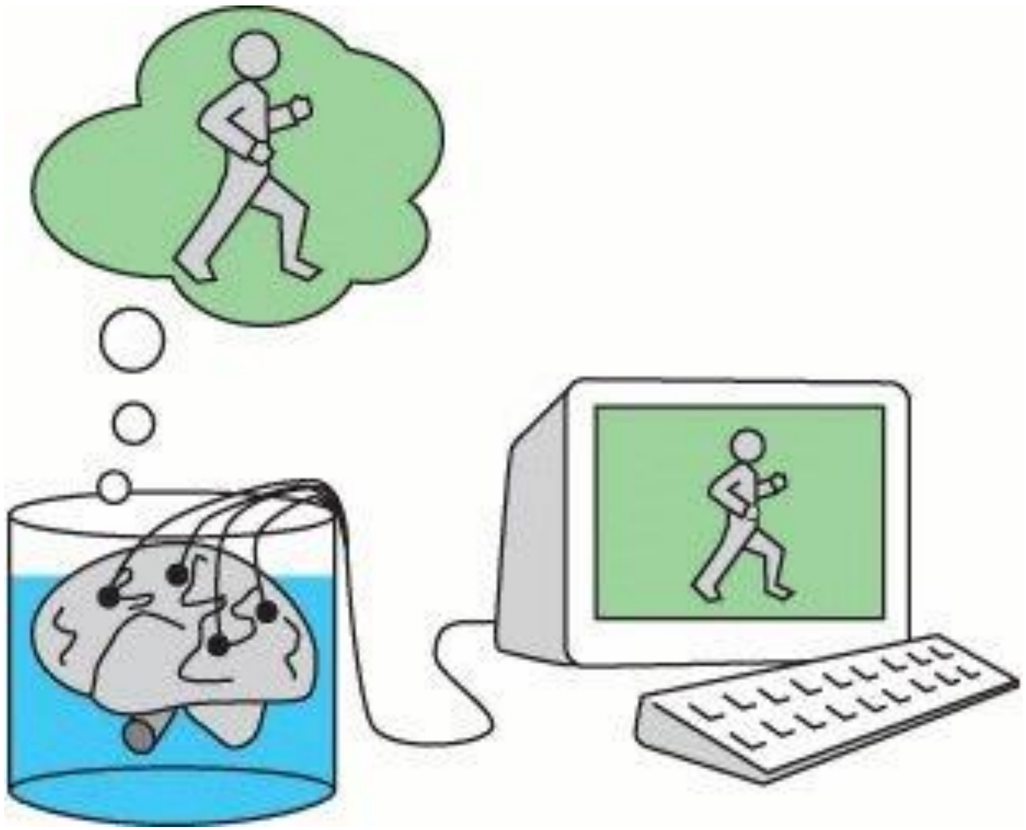
Facts external to the believer (can also) matter for the justification of beliefs.

For more details check: https://plato.stanford.edu/entries/justep-intext/

INTERNAL VS. EXTERNAL JUSTIFICATION: BRAIN IN A VAT

Famous example for clarification: **Brain in a Vat**

- Suppose we had the technology to keep your brain alive in a vat filled by some nutrient fluid.
- We can entertain you by sending electronic impulses directly to your brain in the vat.
- Imagine, we send impulses such that you believe that you are running on your favourite forest path at a sunny afternoon in late June.
- Imagine that your experiences do not differ in any way to the experience that you would have when you were actually running on your favourite forest path at a sunny afternoon in late June.



https://commons.wikimedia.org/wiki/File:Braininvat.jpg

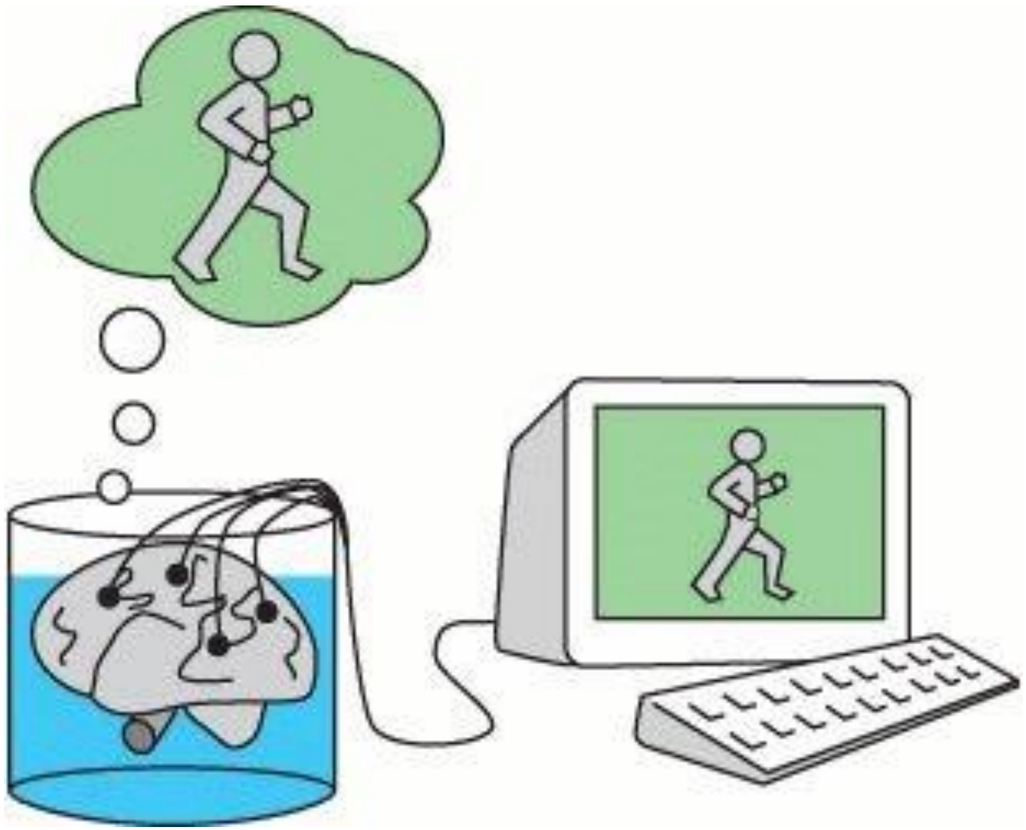
INTERNAL VS. EXTERNAL JUSTIFICATION: BRAIN IN A VAT

- Are you (as the brain in the vat) justified to belief that you are running to the same degree as you are if you really were running there?

→ Internalists affirm

→ Externalists deny

[This is Keanu Reeves when he woke up from his vat in Matrix ;)]



https://commons.wikimedia.org/wiki/File:Braininvat.jpg

INTERNAL VS. EXTERNAL JUSTIFICATION: EXAMPLE THEORIES

Justification$_{epis}$

(It's quite hard to spell out reliability in a satisfactory way. We will spare us this detail for this lecture)

Daniel, the crazy researcher,

believes that the liquid in the Erlenmeyer flask is delicious and strong coffee.

Internalism (Example Theory):

Evidentialism (Rough & Ready Definition) – cf. Conee and Feldman 2004. *Evidentialism: Essays in Epistemology*.

The justification of *S*'s belief that *p* at time *t* depends only on the evidence *S* possesses in *S*'s mind for *p* at *t*.

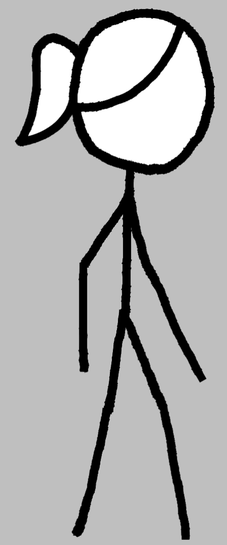
For more details check: https://www.iep.utm.edu/evidenti/

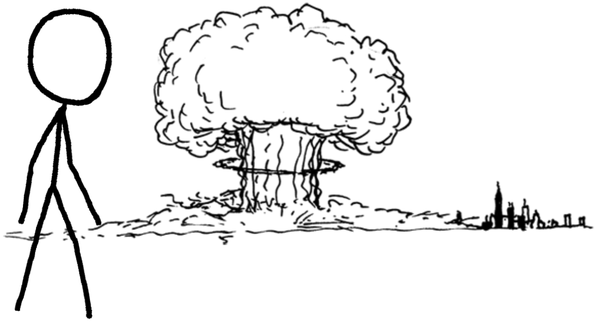
Externalism (Example Theory):

Reliabilism (Rough & Ready Definition) – cf. Goldman 2011, https://plato.stanford.edu/entries/reliabilism/

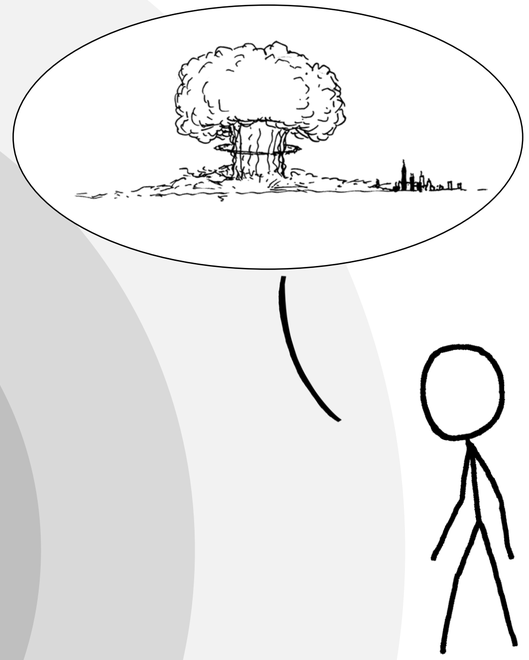
If *S*'s believing *p* at *t* results from a reliable cognitive belief-forming process (or set of processes), then *S*'s belief in *p* at *t* is justified.

THE IDEA & FUNCTION OF TESTEMONY (SOCIAL EPISTEMOLOGY)





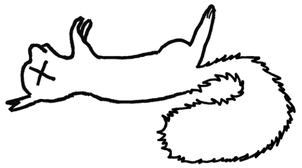
THE IDEA & FUNCTION OF TESTEMONY (SOCIAL EPISTEMOLOGY)

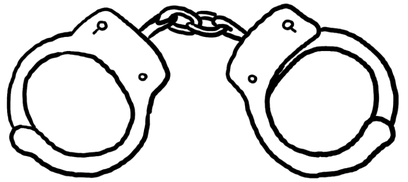


- In some circumstances, *testimony* seems to be an important second-hand kind of source of justification.
- But it is hard to specify under which circumstances it is a valid source!
- Important here: Trust!
- Central question: Under what circumstances is a recipient justified in trusting an assertion made by the sender?
- Turns out, human society is creative in institutionalizing such relations…

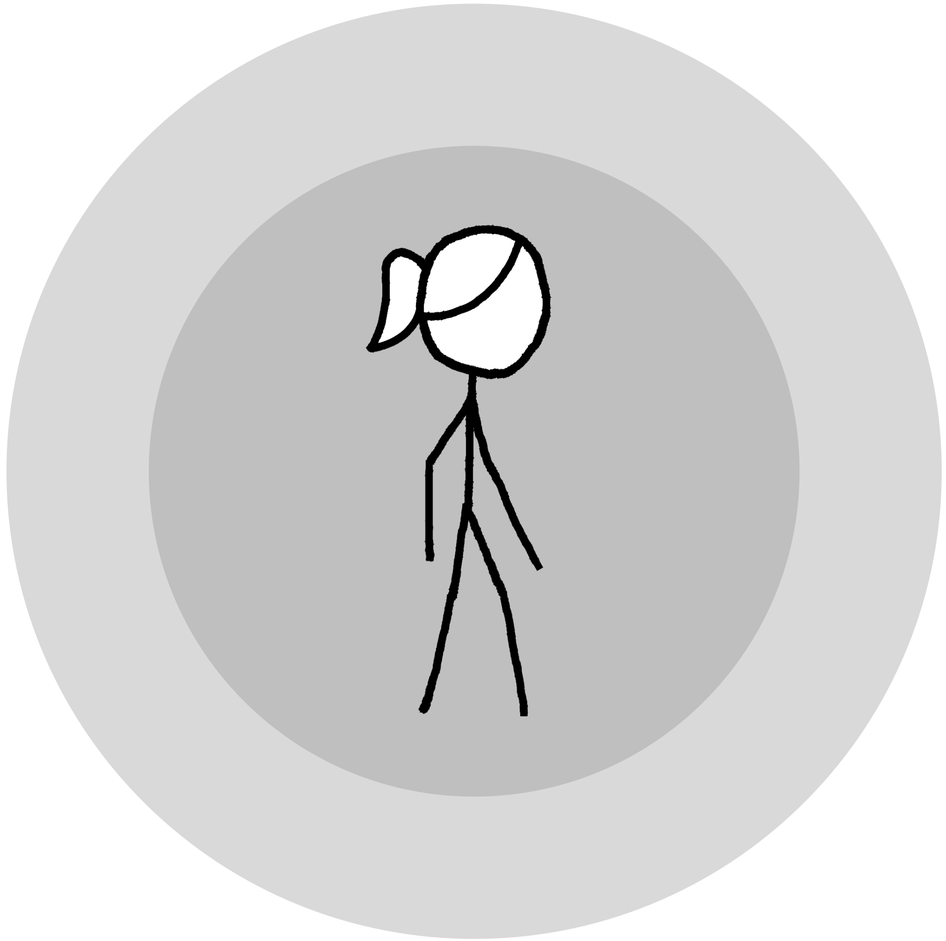
GATEKEEPER AS TRUSTED & TRUSTWORTHY SOURCES OF TESTIMONY

in the communication sense (think of an editor)



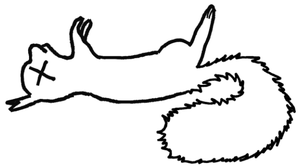


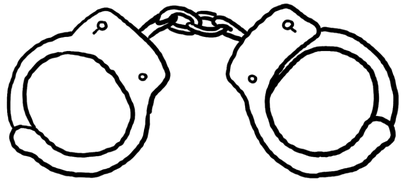
GATEKEEPER AS TRUSTED & TRUSTWORTHY SOURCES OF TESTIMONY



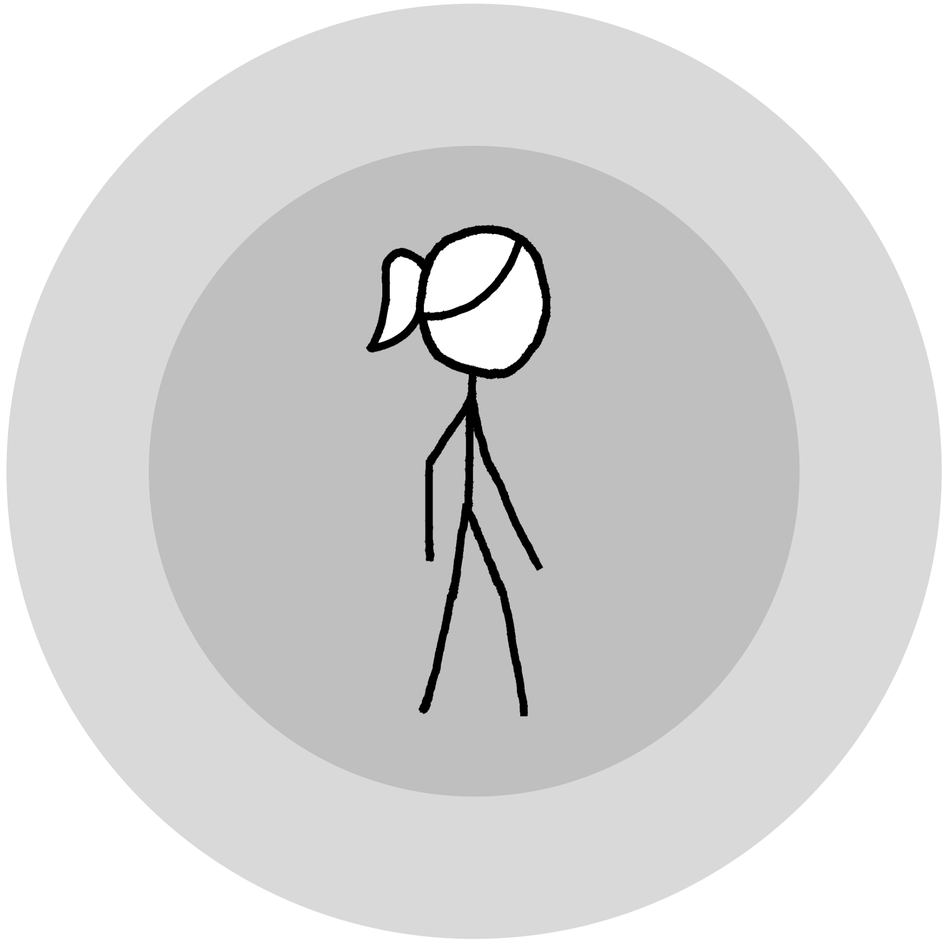
in the communication sense (think of an editor)

aggregators, broadcasters/ dispatchers

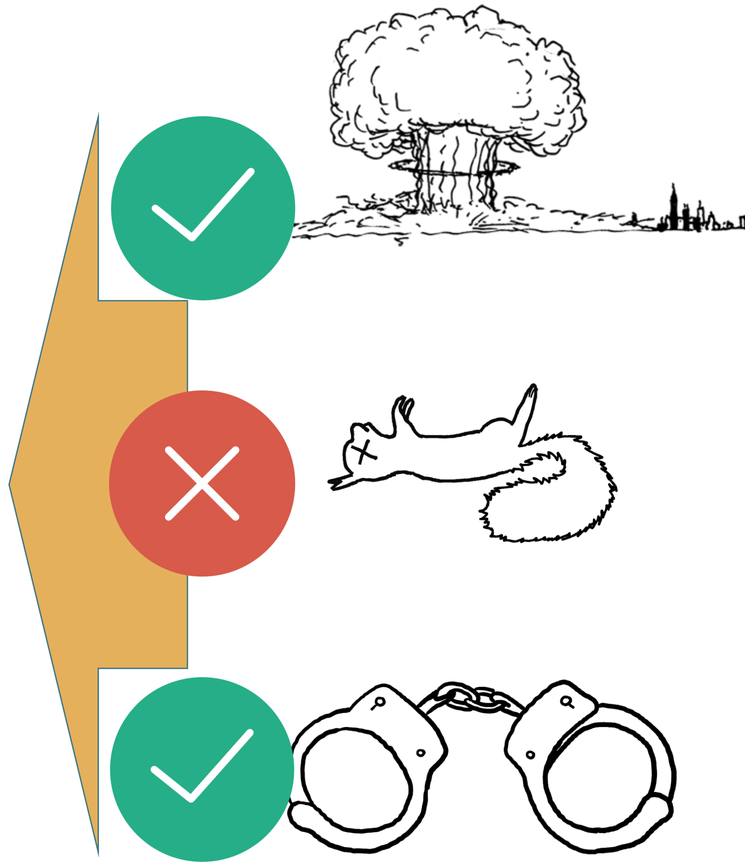




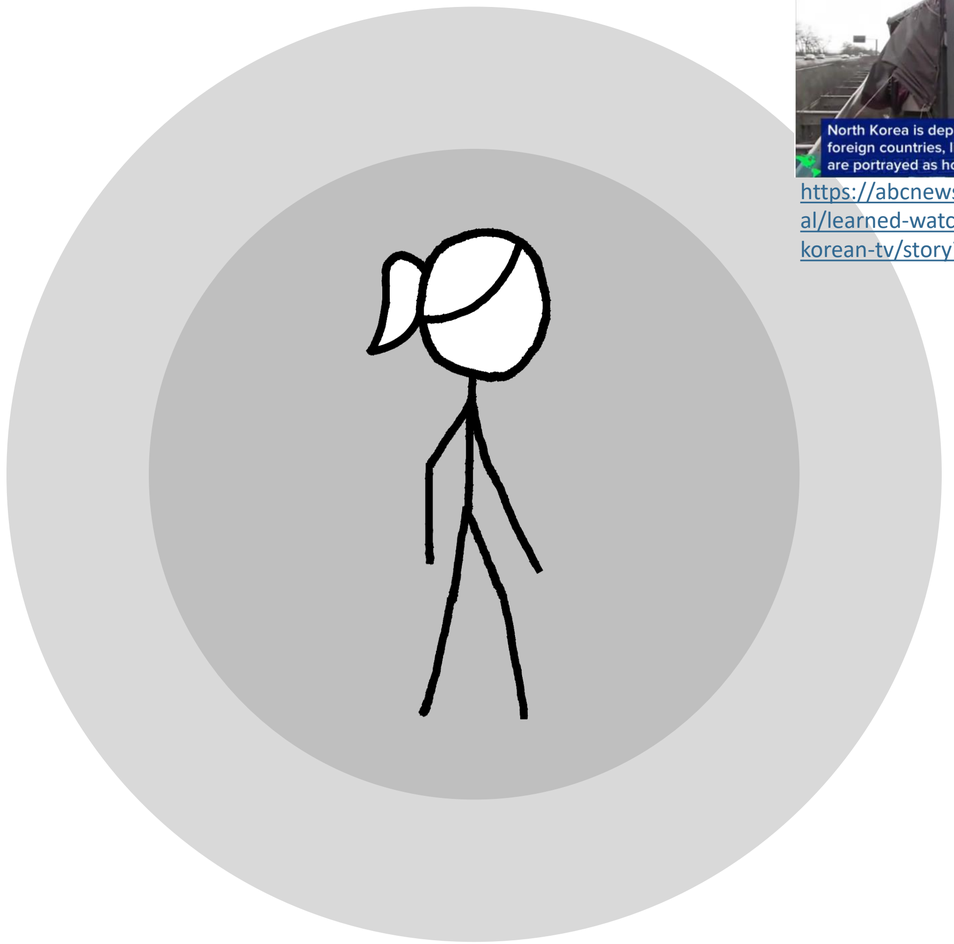
GATEKEEPER AS TRUSTED & TRUSTWORTHY SOURCES OF TESTIMONY



in the communication sense (think of an editor)



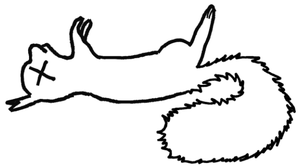
GATEKEEPER AS “TRUSTED & TRUSTWORTHY” SOURCES OF TESTIMONY

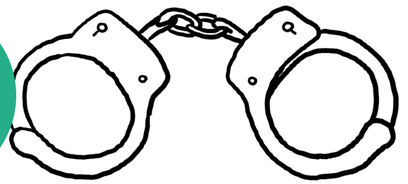


Gatekeeper in the communication sense (think of an editor)

https://abcnews.go.com/International/learned-watching-week-north-korean-tv/story?id=52714005

aggregators, broadcasters/ dispatchers, filters





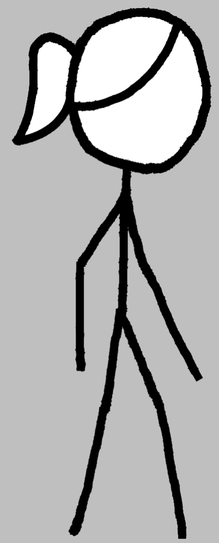
GATEKEEPERS AND RESPONSIBILITY

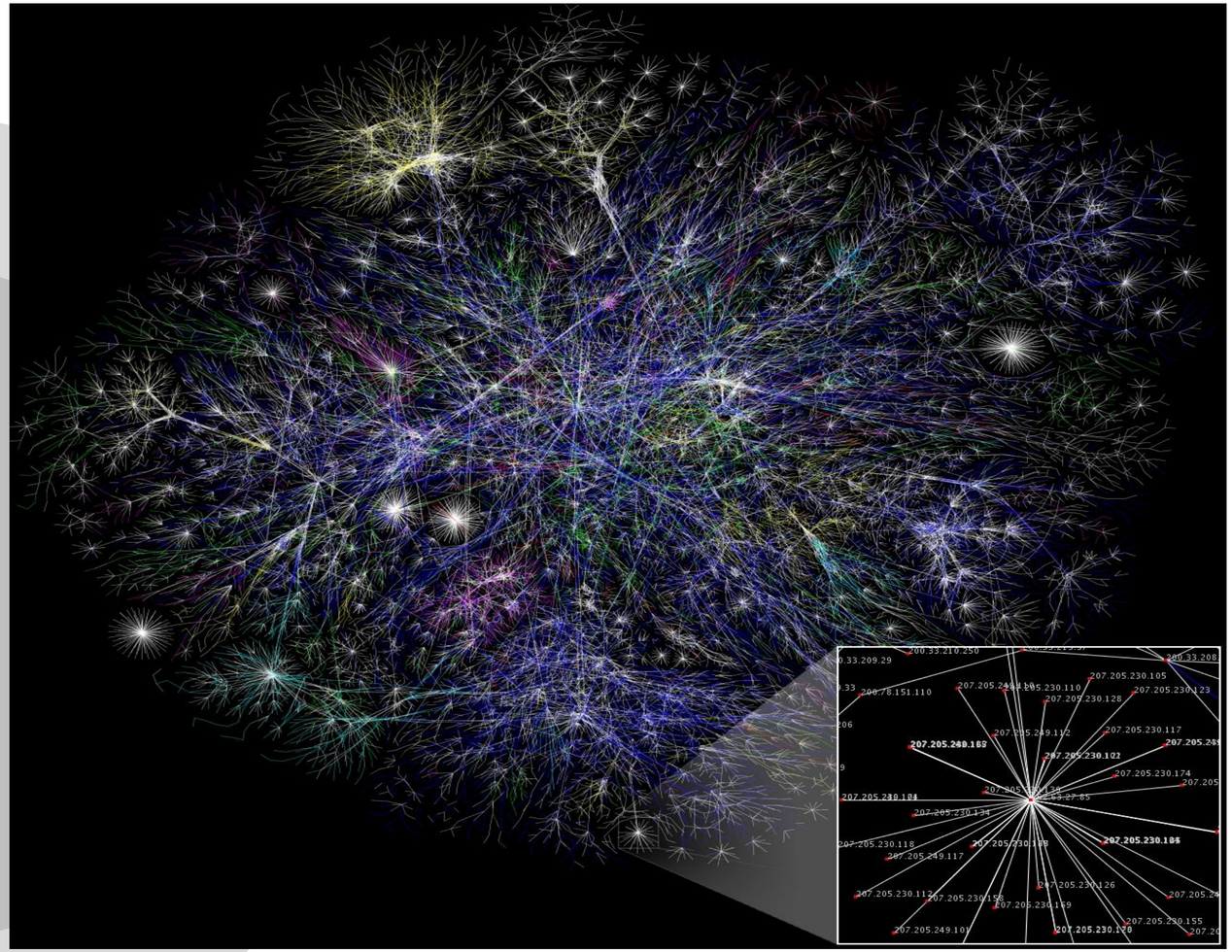
- In order to make gate keepers trustworthy and trusted, these roles are (should be) highly professionalized.
- They have codes of conduct and codes of ethics. Often these are codified in law.
- Generally, someone bears responsibility.

In Germany we have
"V. i. S. d. P."
„Verantwortlich im Sinne des Presserechts"
– "Responsible in the terms of press law"

SPIEGEL 52/2018

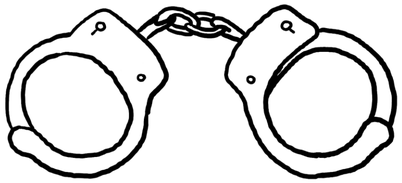
AGE OF MASS-INFORMATION





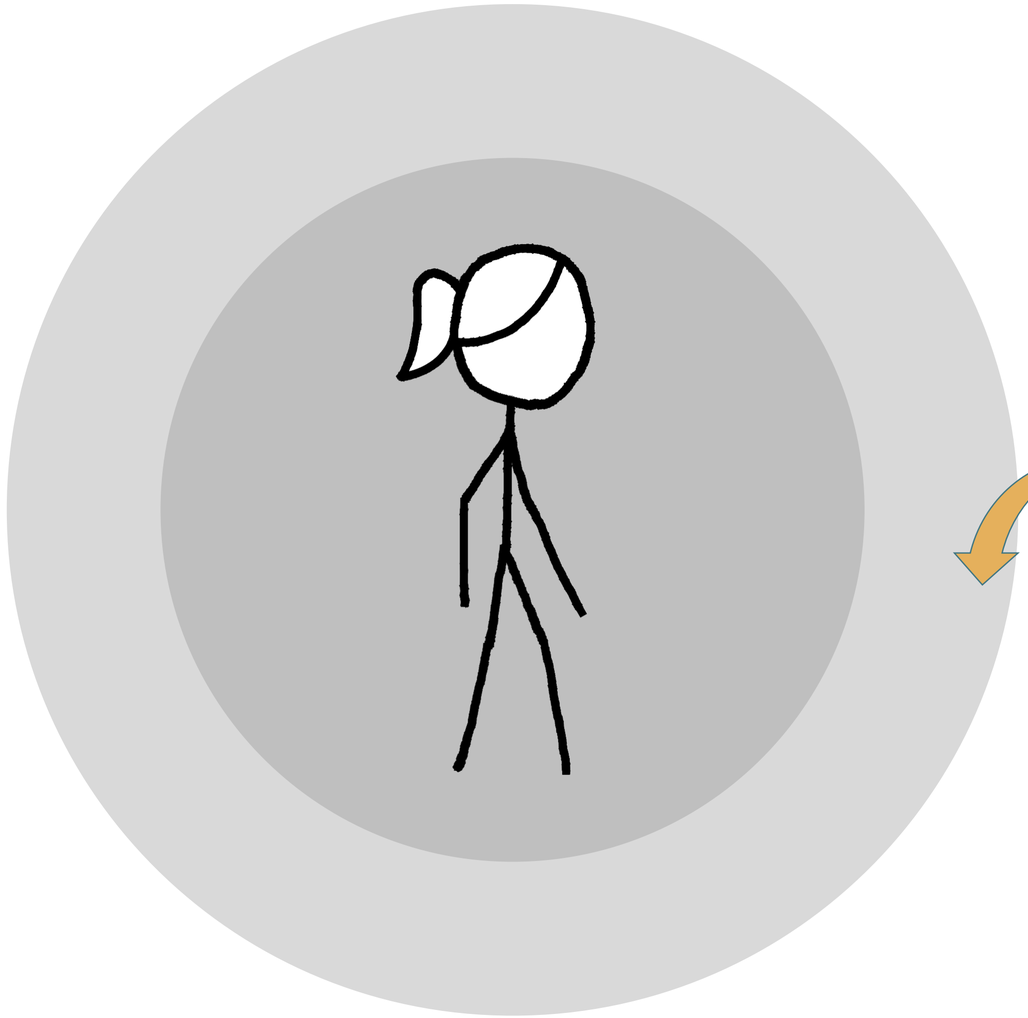
YouTube

...



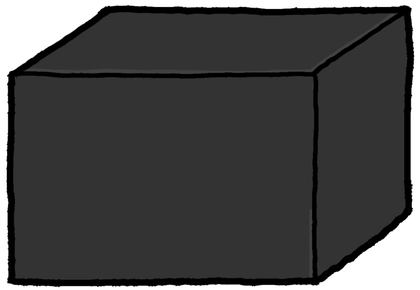
...

THE AGE OF ALGORITHMIC GATEKEEPERS



No GateKeeper?!

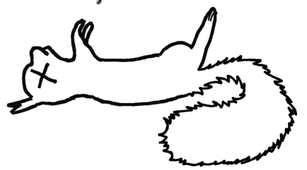
(Not in the **traditional** sense!)

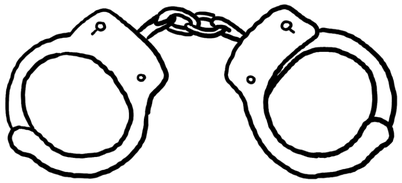


Google

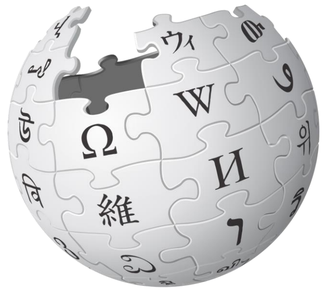
YouTube

...



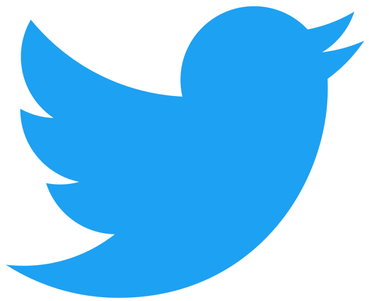


...



ALGORITHMIC GATEKEEPERS

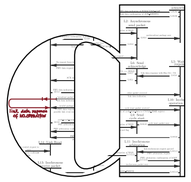
Who are the Gatekeepers nowadays?



Ethics for Nerds

An Advanced Course in Computer Science
Summer Semester 2020

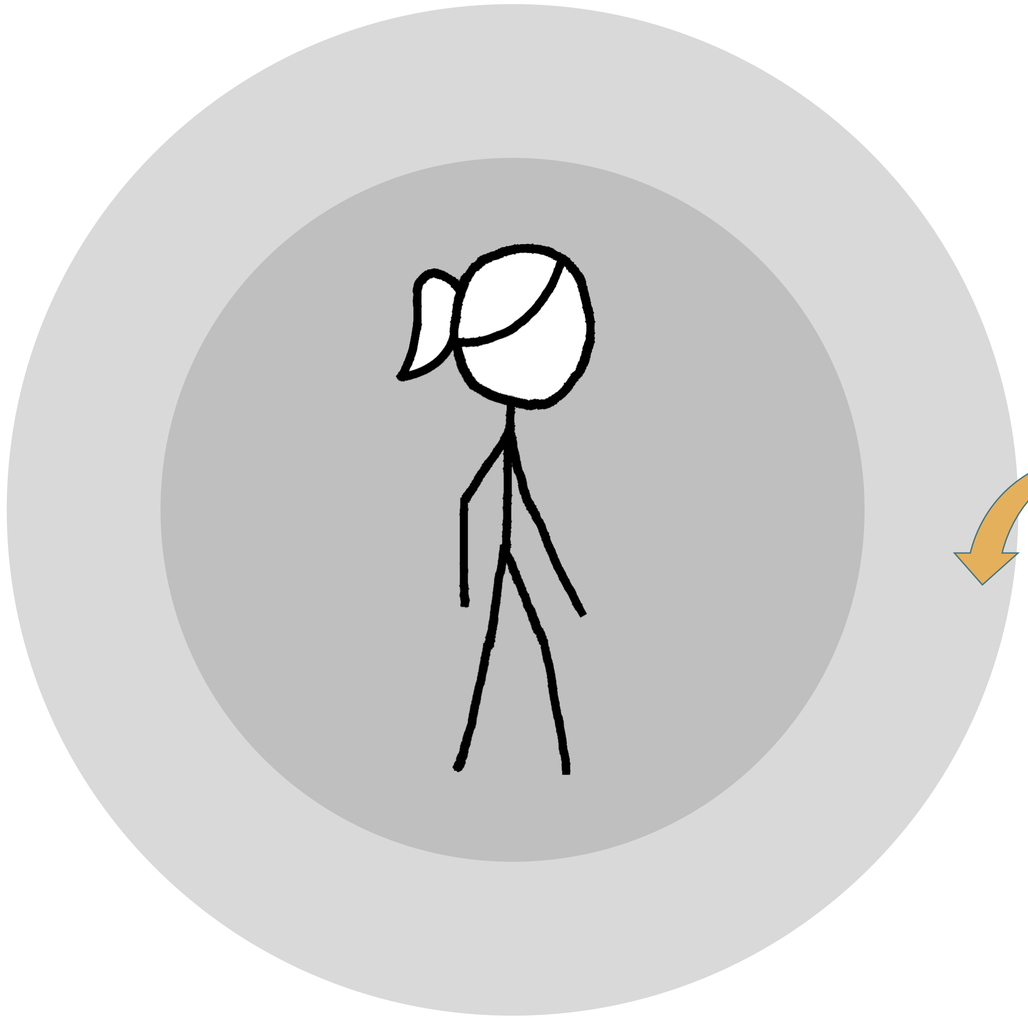
Current Topics 2.3
Filter Bubbles:
The result of optimization?



Prof. Holger Hermanns,
Kevin Baum, Sarah Sterz

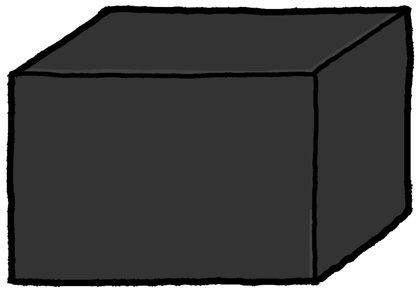
UNIVERSITÄT
DES
SAARLANDES

THE AGE OF ALGORITHMIC GATEKEEPERS



No GateKeeper?!

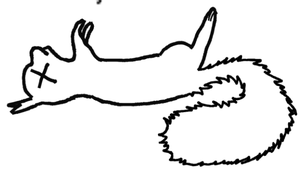
(Not in the **traditional** sense!)

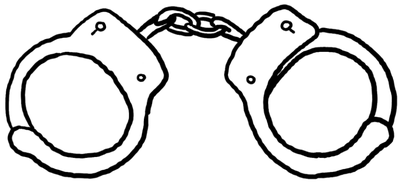


Google

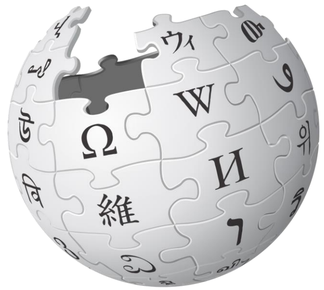
YouTube

...



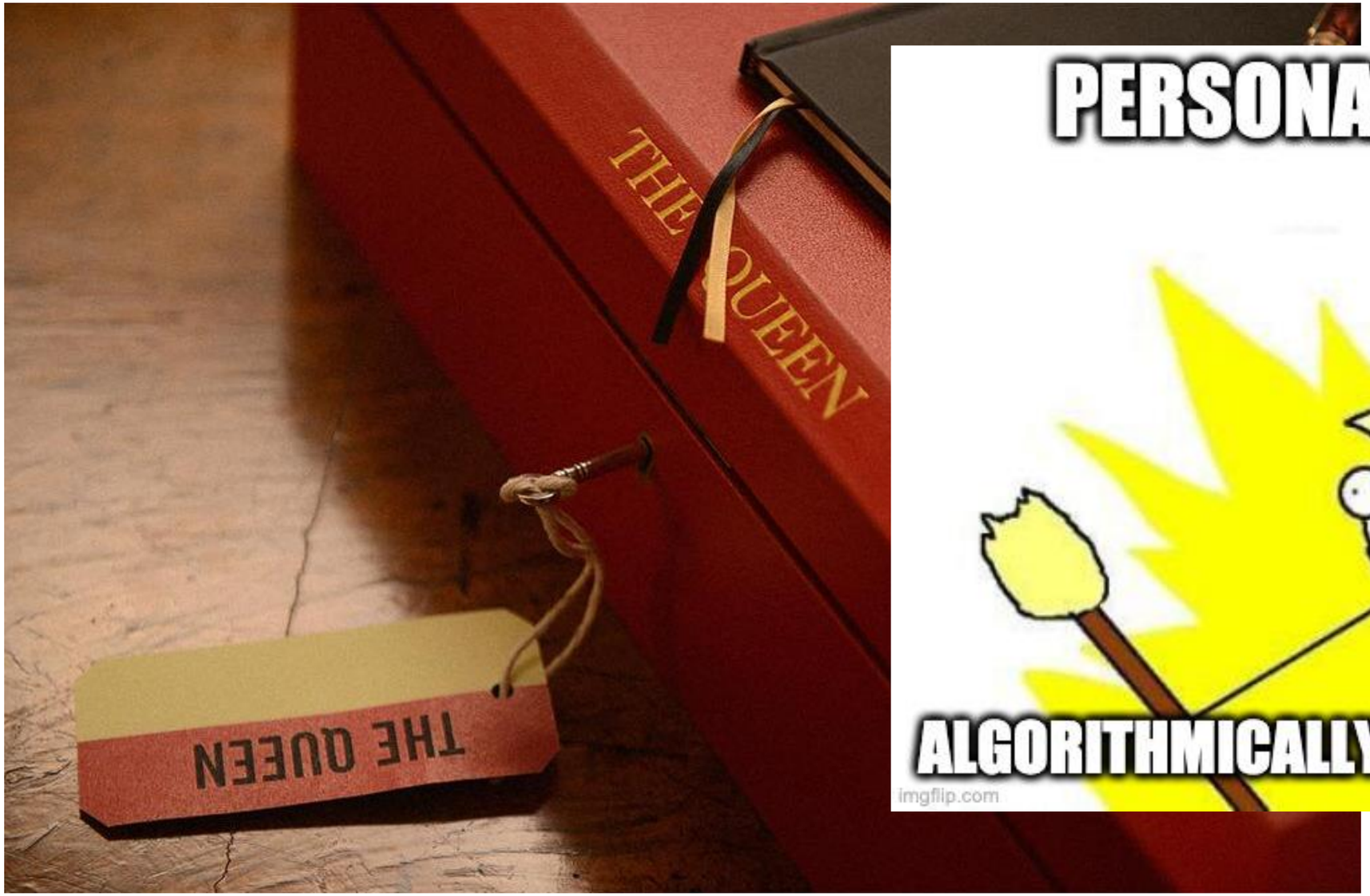


...

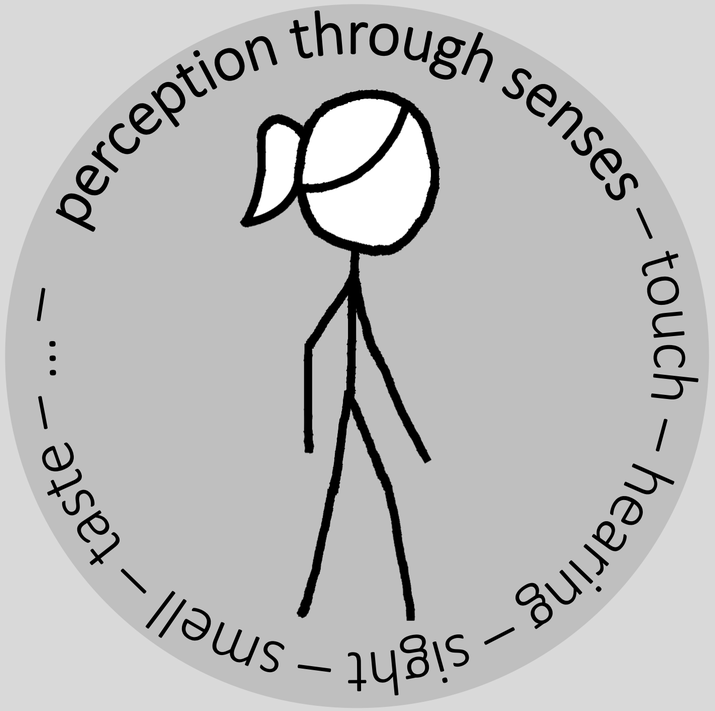


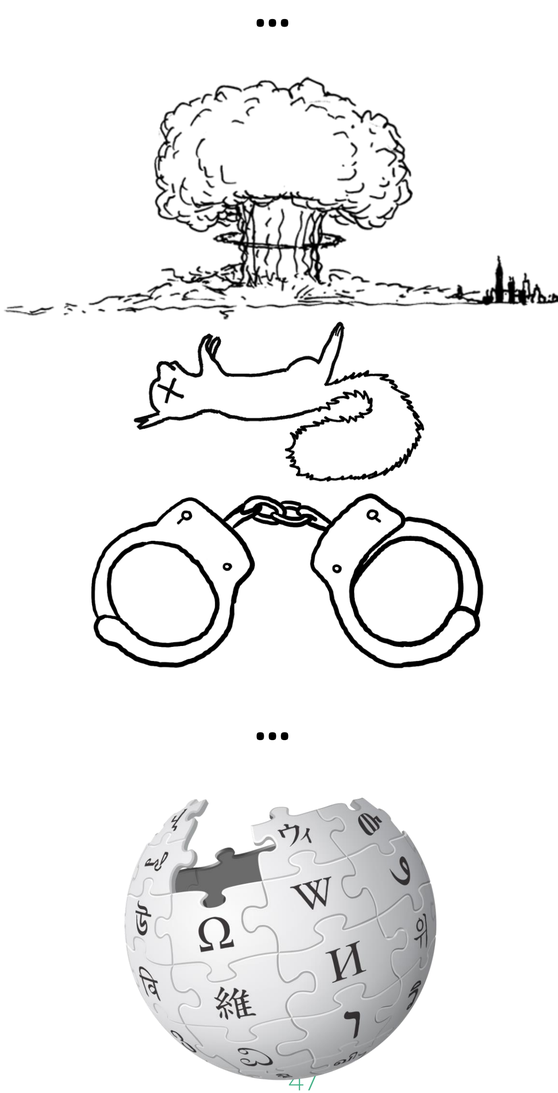
What’s the main difference now?

PERSONALIZE EVERYTHING!



THE AGE OF ALGORITHMIC GATEKEEPERS





THE CENTRAL CHANGE

- If you read a printed newspaper, that's an interaction (with that newspaper).
- If you recommend it, or hand it over to your friend, that's again an interaction (with that newspaper and your friend).
- But none of these interactions are interactions with the editorial board, the publisher, or another third party.

https://www.nytimes.com/2020/04/01/reader-center/times-satire-history.html

THE CENTRAL CHANGE

ENGLISH ESPAÑOL 中文

PLAY THE CROSSWORD Account

The New York Times

Thursday, May 28, 2020 Today's Paper

World U.S. Politics N.Y. Business Opinion Tech Science Health Sports Arts Books Style Food Travel Magazine T Magazine Real Estate Video

The Daily **Listen to 'The Daily'** The privatization of space travel.

DealBook **In the 'DealBook' Newsletter** One giant leap for SpaceX.

The Book Review Podcast **The Book Review Podcast** Jia Lynn Yang talks about "One Mighty and Irresistible Tide."

S&P 500 +1.48% ↑
Dow +2.21% ↑
Nasdaq +0.77% ↑

20°C 21° 8° Sulzbach, Germany

China Officially Expands Power to Rein In Hong Kong, Despite Outcry

- Beijing ordered that a new law be written to suppress subversion, secession, terrorism and other acts that might threaten national security in Hong Kong.
- The specifics of the legislation, to be hashed out in the coming weeks, will help determine the fate of Hong Kong.
- Early signals from Chinese authorities point to a crackdown once the law takes effect, which is expected by September.

just now 119 comments

Watching a TV news show about the new national security legislation in Hong Kong on Thursday.
Lam Yik Fei for The New York Times

U.S. to Expel Chinese Graduate Students With Ties to China's Military Schools

Many university officials say the government is paranoid, and that the United States will lose out.

27m ago

Live Updates 13 minutes ago

Economy

- Total U.S. unemployment claims are likely to surpass 40 million.
- U.S. stock futures waver ahead of new jobless data.
- For many professionals, their careers are in limbo.

Global

- As mass testing and contact tracing are rolled out in some nations, others are way ahead.
- Mexico's broken hospitals put patients and health workers at high risk.
- As caseloads soar in Indonesia, experts brace for runaway transmission.

U.S.

- The divide over the responsibility for testing.
- Forecasters expect 2 million more jobless claims.
- As jobs disappear, some professionals settle into a life in limbo.

The World Is Still Far From Herd Immunity for Coronavirus

The latest studies show that even in the hardest-hit cities, most people remain vulnerable.

55m ago

New York City Herd immunity estimate

Has antibodies

'It's Not The Virus': Mexico's Broken Hospitals Become Killers, Too

Years of neglect have hobbled many Mexican hospitals. Now, as the pandemic strikes, some patients are dying preventable deaths, doctors and nurses say.

3h ago

Nearly 50,000 people in the New York area tested positive for the virus in the last two weeks. Who are they?

3h ago

Your tracker settings

We use cookies and similar methods to recognize visitors and remember their preferences. We also use them to measure ad campaign effectiveness, target ads and analyze site traffic. To learn more about these methods, including how to disable them, view our Cookie Policy.

By clicking 'accept,' you consent to the processing of your data by us and third parties using the above methods. You can always change your tracker preferences by visiting our Cookie Policy.

ACCEPT

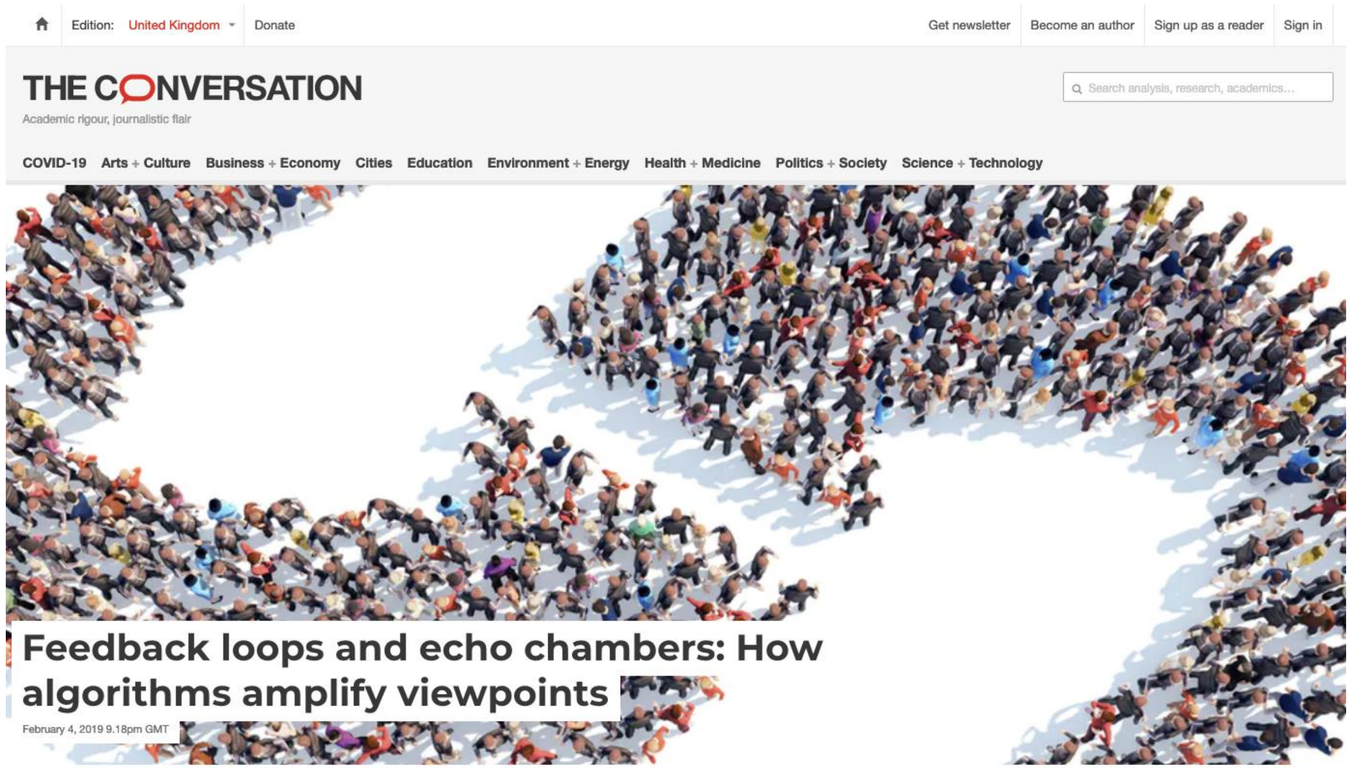
MANAGE TRACKERS

https://www.nytimes.com/

- If you read articles online, that's interaction that can be measured by third parties.
- If you read it, recommend it, or hand it over to your friend, via social media or email, that's again an interaction that usually can be measured by third parties.
- Interactions now regularly affect what you (and others!) see in the future and it does so in quite indirect and mostly opaque ways.

THE NEW DYNAMIC

https://theconversation.com/feedback-loops-and-echo-chambers-how-algorithms-amplify-viewpoints-107935



Feedback loops and echo chambers: How algorithms amplify viewpoints

February 4, 2019 9.18pm GMT

Feedback loops in algorithms amplify chosen content, to the exclusion of others. Shutterstock

Email
Twitter 24
Facebook 73
LinkedIn
Print

Whether it's allegations of ethnic cleansing in Myanmar, anti-Muslim violence in Sri Lanka or the "*gilets jaunes*" protests in France, it is clear that social media platforms are helping spread divisive messages online at an alarming rate and potentially fueling offline violence.

But the debate is about whether these platforms are an essential cause, without which these events could not have happened, or merely reflect real-world tensions.

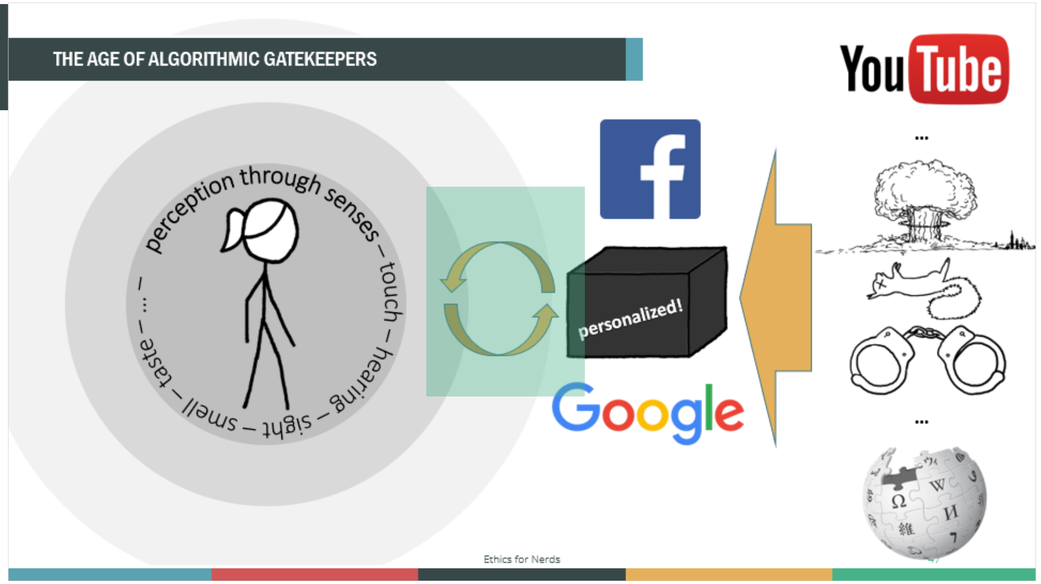
Author

Swathi Meenakshi Sadagopan
Munk Fellow, University of Toronto

Disclosure statement

Swathi Meenakshi Sadagopan is affiliated with Deloitte Canada.

Partners



Training the algorithms

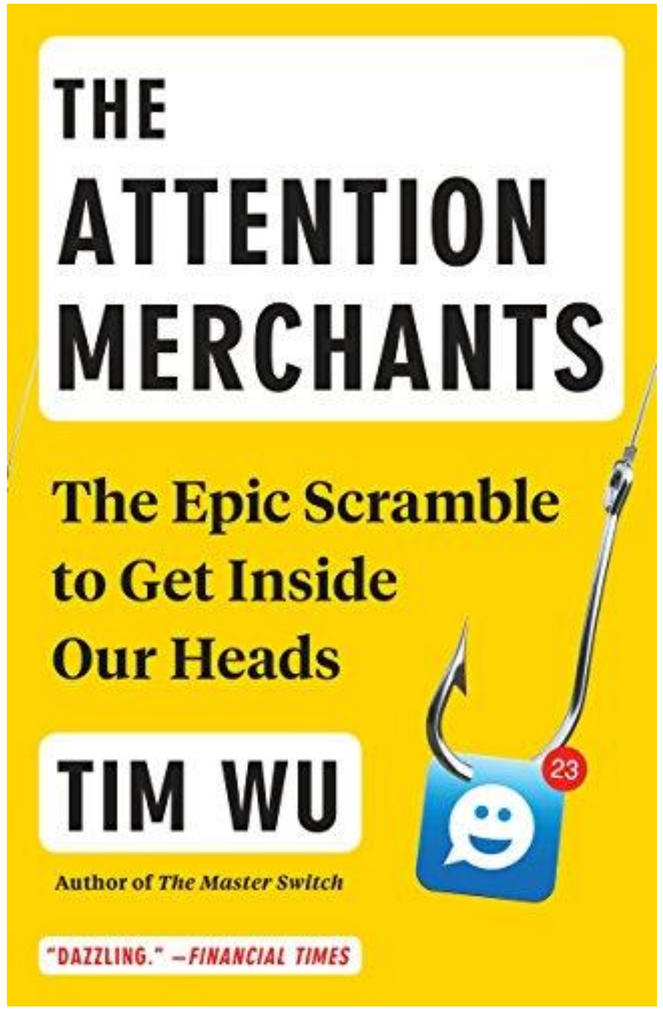
Recommendation algorithms were created by companies such as Facebook, YouTube, Netflix or Amazon for the purpose of helping people make decisions. An array of options are recommended and a choice is made by the user that is then fed as new knowledge to train the algorithm — without factoring in that the choice was in fact an output shown by the algorithm.

This creates a feedback loop, where the output of the algorithm becomes part of its input. As expected, recommendations similar to the choice that was made are shown.

This leaves us with a chicken-or-egg dilemma: Did you click on something because you were inherently interested in it, or did you click on it because you were recommended it? The answer, according to Chaney's research, lies somewhere in between.

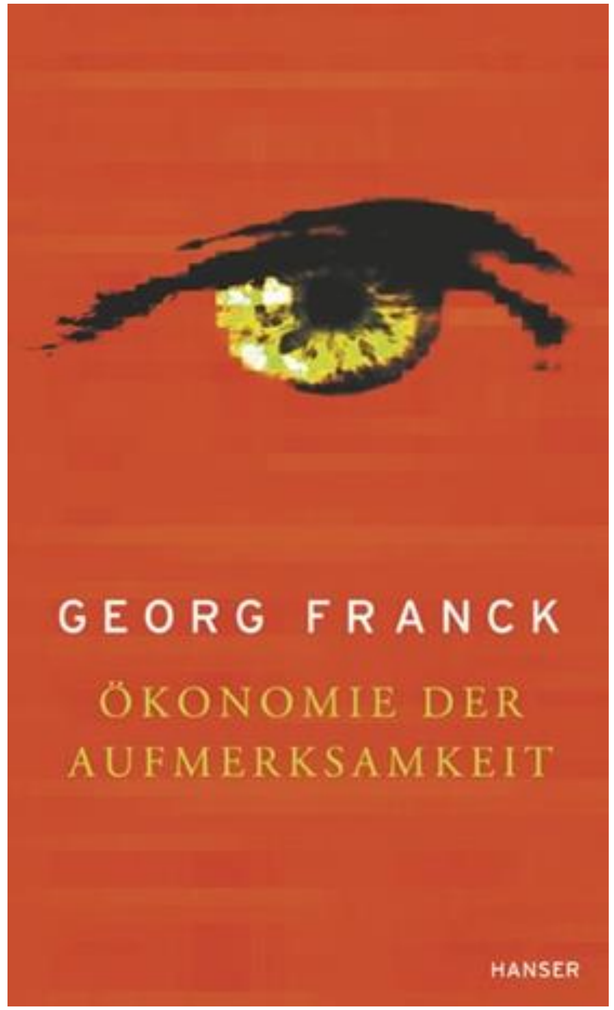
REASONS FOR AND GOALS OF FILTERING

https://www.penguinrandomhouse.com/books/234876/the-attention-merchants-by-tim-wu/



2017

https://www.hanser-literaturverlage.de/buch/oekonomie-der-aufmerksamkeit/978-3-446-19348-2/



1998

Attention is a resource—a person has only so much of it.

Matthew B. Crawford

When an online service is free, you're not the customer. You're the product.

Tim Wu

....the internet is showing us what it thinks we want to see, but not necessarily what we need to see.

Eli Pariser

INFORMATICS AND THE FILTERBUBBLE

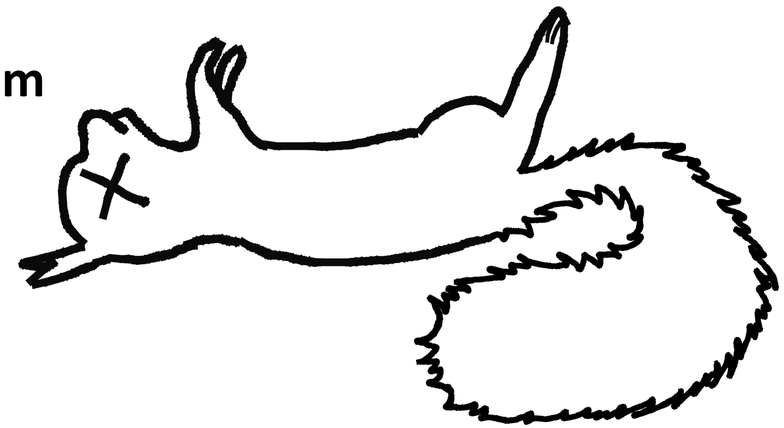
- **We want** algorithmic services that only present relevant stuff and options to us.
- **Services can** monetize your attention (first and foremost via advertisement).
- The **goal** is attention, interaction, engagement.
- These are **scarce, limited resources**!
- It is possible to measure interaction, engagement, and proxy variables like dwell time.
- This sounds like an **optimization problem** made for computer scientists!

Mandel Ngan/AFP via Getty Images

https://www.vox.com/recode/2019/10/24/20930673/mark-zuckerberg-facebook-testimony-congress-libra-currency-recode-podcast-transcript

A squirrel dying in front of your house may be more relevant to your interests right now than people dying in Africa.

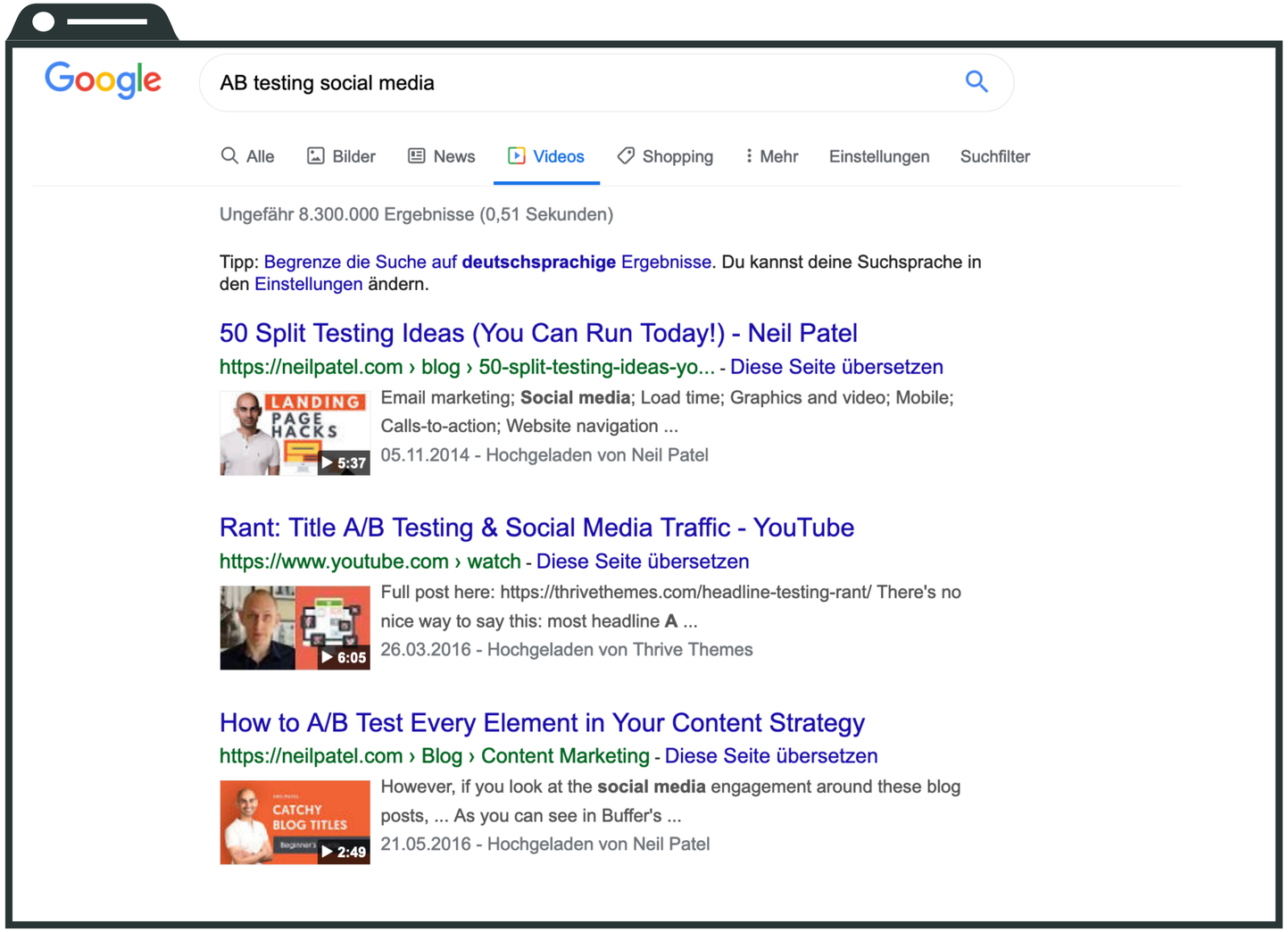
Mark Zuckerberg



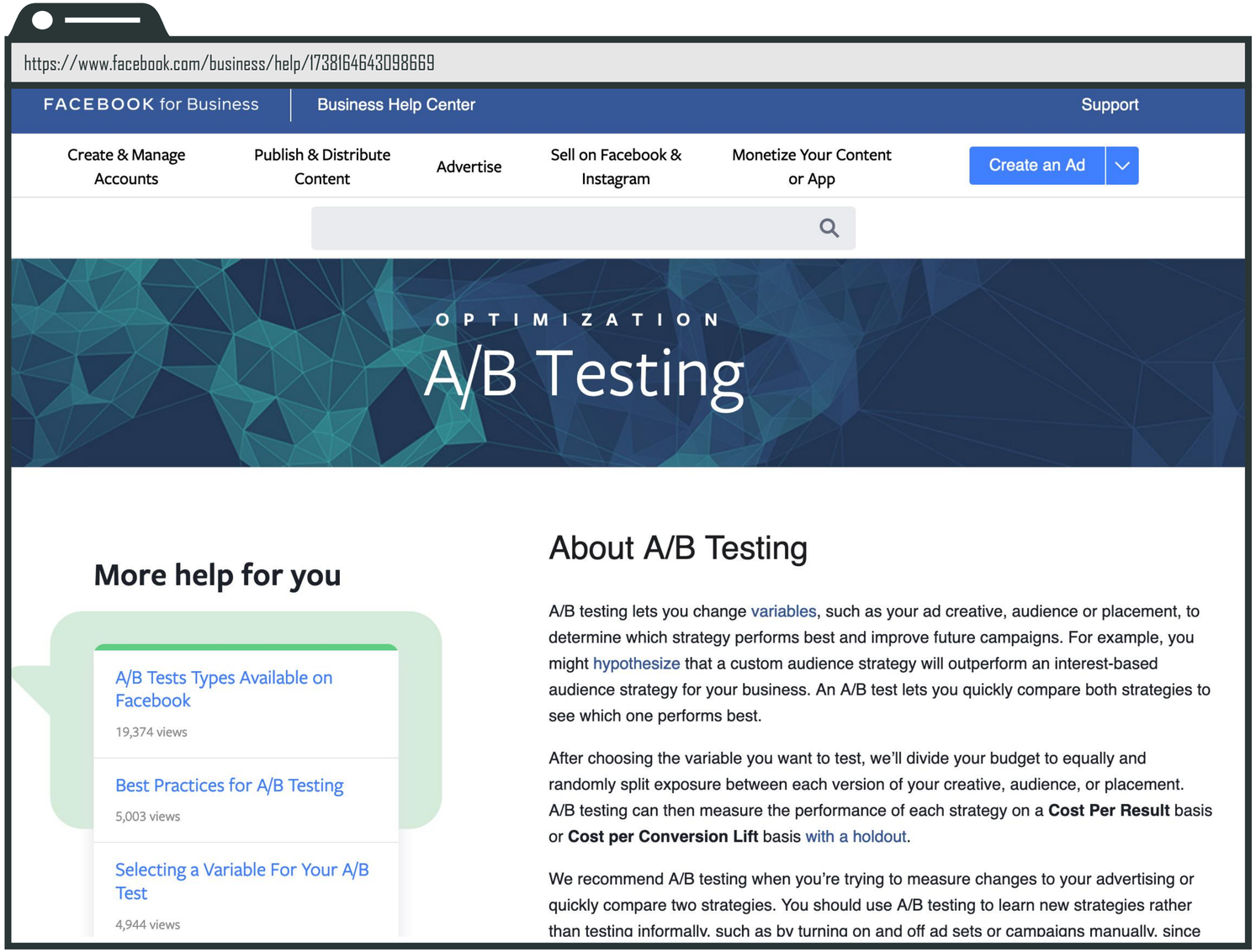
HOW TO OPTIMIZE: A STRAIGHT-FORWARD APPROACH

A/B Testing in a Nutshell

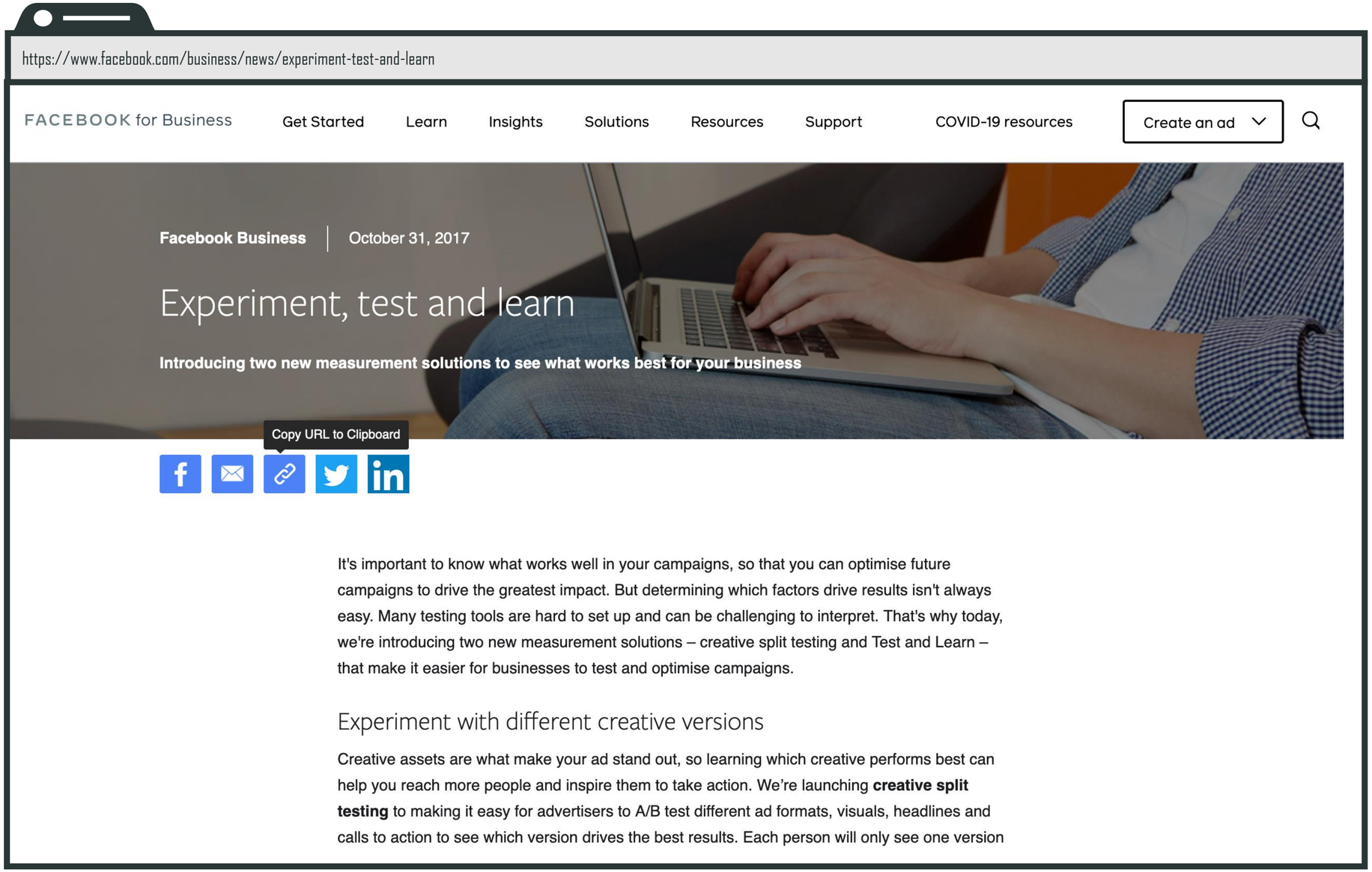
- Start with some state A (of a website, a campaign, a blog post,…).
- Vary that state into a new variant B.
- Show A and B respectively to random groups.
- Measure success of A and variant B (in our case: dwell time, engagement…).
- Throw away looser, keep winner, repeat.



HOW TO OPTIMIZE



HOW TO OPTIMIZE



RELEVANT CONCEPTS FROM (SOCIAL-)PSYCHOLOGY

Confirmation Bias (rough & ready)

The tendency to search for, interpret, favour, and recall information that confirms or support one's prior personal beliefs or values.

*Nickerson, Raymond S. (June 1998), "Confirmation bias: A ubiquitous phenomenon in many guises", Review of General Psychology, **2** (2): 175–220, doi:10.1037/1089-2680.2.2.175*

Cognitive Dissonance (rough & ready)

Cognitive dissonance occurs when a subject holds two or more contrary beliefs, ideas, or values; leads to psychological stress; subjects try to change them until they become consistent and reject, dismiss, and ignore inconsistent evidence in order to avoid cognitive dissonance (selective perception).

CONTRADICTORY VS CONTRARY

A contradicts B

$A \equiv \neg B$

“Timo is in Greece”

contradicts

“Timo is not in Greece”

$\subset$

A is contrary to B

A and B cannot both be true at the same time

“Timo is in Greece”

is contrary to, but does not contradict

“Timo is in Japan”

RELEVANT CONCEPTS FROM (SOCIAL-)PSYCHOLOGY

Confirmation Bias (rough & ready)

The tendency to search for, interpret, favour, and recall information that confirms or support one's prior personal beliefs or values.

*Nickerson, Raymond S. (June 1998), "Confirmation bias: A ubiquitous phenomenon in many guises", Review of General Psychology, **2** (2): 175–220, doi:10.1037/1089-2680.2.2.175*

Cognitive Dissonance (rough & ready)

Cognitive dissonance occurs when a subject holds two or more contrary beliefs, ideas, or values; leads to psychological stress; subjects try to change them until they become consistent and reject, dismiss, and ignore inconsistent evidence in order to avoid cognitive dissonance (selective perception).

Negativity Bias (rough & ready)

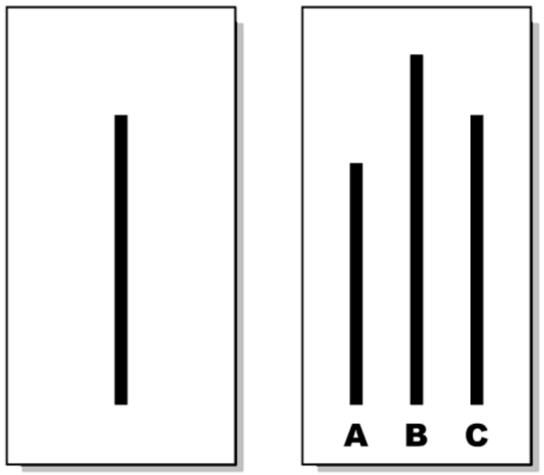
Even when two stimuli have equal intensity, things of a more negative nature (e.g. unpleasant thoughts, emotions, or social interactions; harmful/traumatic events) have a greater effect on one's psychological state and processes than neutral or positive things. (Bad news is good news!)

RELEVANT CONCEPTS FROM (SOCIAL-)PSYCHOLOGY

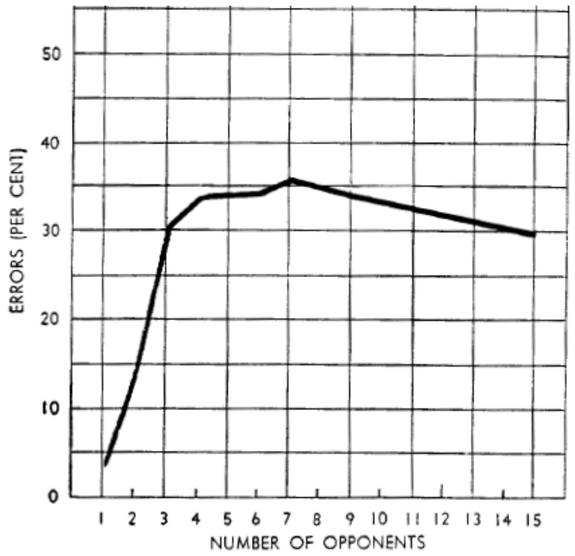
https://blogs.cornell.edu/info2040/2011/11/10/the-conformity-cascade/

Conformity Cascade (Rough & Ready)

"People look at the decisions that other people around them make, and based on that they make their own decision. If everyone else is making the same decision, then why even consider your own opinion, might as well just go with the flow."



What line from the right card has the same length as the one from the left card?



SIZE OF MAJORITY which opposed them had an effect on the subjects. With a single opponent the subject erred only 3.6 per cent of the time; with two opponents he erred 13.6 per cent; three, 31.8 per cent; four, 35.1 per cent; six, 35.2 per cent; seven, 37.1 per cent; nine, 35.1 per cent; 15, 31.2 per cent.

Networks

Course blog for INFO 2040/CS 2850/Econ 2040/SOC 2090

The Conformity Cascade

Crowd Mentality has been a big topic in the area of psychology. People in masses tend to act differently than individuals. What is the reason behind this? Do people actually lose their personalities when surrounded by other people? Do they just blindly follow others? Are they too afraid to express what they really think? Or are they actually making the best informed decision that they can?

It just so happens that crowd behavior can easily be modeled as a specific case of the accept & reject cascade. People look at the decisions that other people around them make, and based on that they make their own decision. If everyone else is making the same decision, then why even consider your own opinion, might as well just go with the flow.

Solomon Asch did a number of tests about conformity in 1958. His experiment consisted of the following:

Every participant would be placed in a group with a number of 'confederates' (people who knew the true aim of the experiment but were introduced as 'other participants' to the participant). They were told that they were participating in a number of visual tests, very simple tests (like saying whether two lines had the same length) with obvious answers and they were asked to say their answer out loud (so other participants could hear). For the first 2 trials, the confederates would give the correct answer, making the participant at ease as everyone agreed. On the following tests, the confederates would all give the incorrect answer before giving the participant a chance to say his answer. These trials were repeated a number of times to see how much an individual would conform (the confederates would all occasionally give the correct answer again to make the experiment more believable).

DIGITAL TRIBALISM & TRIABLIST CONFORMITY

https://www.ft.com/content/89f16688-fb15-11e7-a492-2c9be7f3120a

Opinion **FT Magazine**

Loneliness is contributing to our increasingly tribal politics

‘Everyone laments polarisation but what’s often overlooked is that it’s creating a novel sense of belonging and identity’

SIMON KUPER + Add to myFT

Save

Simon Kuper JANUARY 18 2018

150

Here’s an everyday event in Donald Trump’s America. Two people run into each other in their neighbourhood, or virtually on Facebook, and instantly start discussing the president. If they are liberals, one might say, “Did you see that tweet?!” and the other will tap his forehead meaningfully. If the two support Trump, they might share a grumble about lying media.

These people are participating in the political polarisation that has riven the US and, to a much lesser degree, Brexit Britain. But they are also signalling something else to each other, namely: “You and I belong to the same tribe. We have a shared identity, and something to talk about.” In other words, they are doing something that is usually considered positive: they are forging a new kind of community. Everyone rightly laments polarisation, but what’s often overlooked is that it’s creating a novel sense of belonging, and identity, in societies that were getting scarily atomised.

Many people in western countries have been struggling to define who they are, and what tribe they belong to. Fifty years ago, most people found identity through their family, church, neighbourhood and (if male) their job and trade union.

DIGITAL TRIBALISM & TRIBALIST EPISTEMOLOGY

http://www.ctrl-verlust.net/digital-tribalism-the-real-story-about-fake-news/#fn-2223-3

ctrl+verlust

Res gesta per amissionem

← Vier Jahre nach Snowden – Wird die EU-Datenschutzgrundverordnung uns vor der Überwachung retten?
Was ist Plattformpolitik? Grundzüge einer neuen Form der politischen Macht →

Digital Tribalism – The Real Story About Fake News

Publiziert am 19. Dezember 2017 von mspro

Text by: Michael Seemann / Data Visualization by: Michael Kreil

[Download as PDF]
[(original) German Version]

The Internet has always been my dream of freedom. By this I mean not only the freedom of communication and information, but also the hope for a new freedom of social relations. Despite all the social mobility of modern society, social relations are still somewhat constricting today. From kindergarten to school, from the club to the workplace, we are constantly fed through organizational forms that categorize, sort and thereby de-individualize us. From grassroots groups to citizenship, the whole of society is organized like a group game, and we are rarely allowed to choose our fellow players. It's always like, „*Find your place, settle down, be a part of us.*"

In my search for an explanation for this phenomenon, I repeatedly came across the connection between identity and truth. People who believe that Hillary and Bill Clinton had a number of people murdered and that the Democratic Party was running a child sex trafficking ring in the basement of a pizza shop in Washington DC are not simply stupid or uneducated. They spread this message because it signals membership to their specific group. David Roberts coined the term „tribal epistemology" for this phenomenon, and defines it as follows:

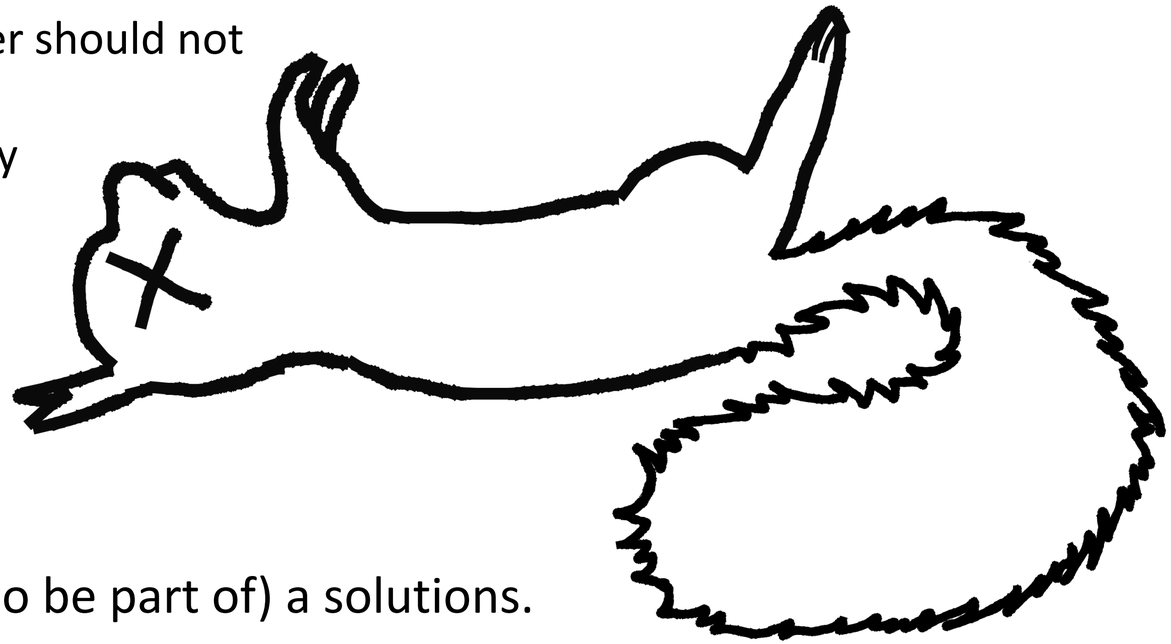
> *Information is evaluated based not on conformity to common standards of evidence or correspondence to a common understanding of the world, but on whether it supports the tribe's values and goals and is vouchsafed by tribal leaders. "Good for our side" and "true" begin to blur into one.*[2]

INFORMATICS AND THE FILTERBUBBLE

- If the resulting algorithms have bad effects and work as intended, then…

  1. because they optimize for an aspect that it should not (engagement, dwell time…) (speaking in terms of the interest of users), or

  2. because people are just responding to stuff they better should not for they are not in their (or our collective) interest (which might be true generally or in context of a highly adaptive system).



- As long as 2. is (also) true (which seems plausible), algorithms are not the singular point of application to solve filter algorithm based problems.

- Then, the change of human responses and behaviour, be it through training or by education, *plausible is* (also be part of) a solutions.

- In other words: It seems rather plausible that it is not “bad” or “evil” algorithms that cause the problems per se. They rather enhance human tendencies in a bad way.

Examples

YOUTUBE AND FILTER BUBBLES

#youtubewakeup

https://www.nytimes.com/2019/06/03/world/americas/youtube-pedophiles.html

The New York Times

THE INTERPRETER

On YouTube's Digital Playground, an Open Gate for Pedophiles

YOUTUBE: PATH INTO THE CONSPIRACY RABBIT HOLE?

BBC Sign in | News | Sport | Reel | Worklife | Travel | Future | M

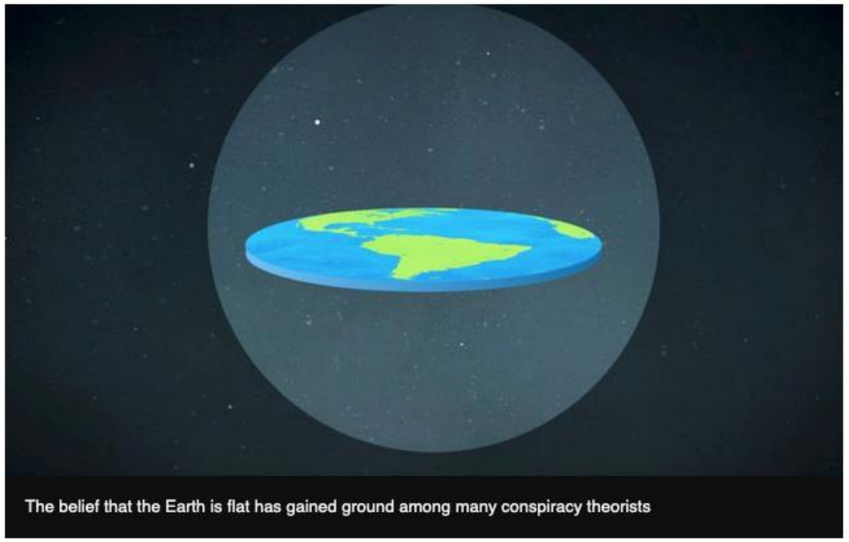
NEWS

Home | Video | World | UK | Business | Tech | Science | Stories | Entertainment & Arts

Technology

YouTube aids flat earth conspiracy theorists, research suggests

18 February 2019 Share



The belief that the Earth is flat has gained ground among many conspiracy theorists

YouTube is playing a significant role in convincing some people that the Earth is flat, research suggests.

The algorithms the site used to guide people to topics they might be interested in made it easy to "end up down the rabbit hole" of misinformation, said Prof Landrum.

"Believing the Earth is flat is of itself is not necessarily harmful, but it comes packaged with a distrust in institutions and authority more generally," she added.

The study involved interviews with 30 attendees at two conferences.

Questioning revealed YouTube had suggested the flat earth videos after attendees had watched other clips at home about conspiracy theories.

Some said they only watched the videos to criticise them but were won over by the arguments being advanced. The results from Prof Landrum's study were presented at the annual meeting of the Association for the Advancement of Science this weekend.

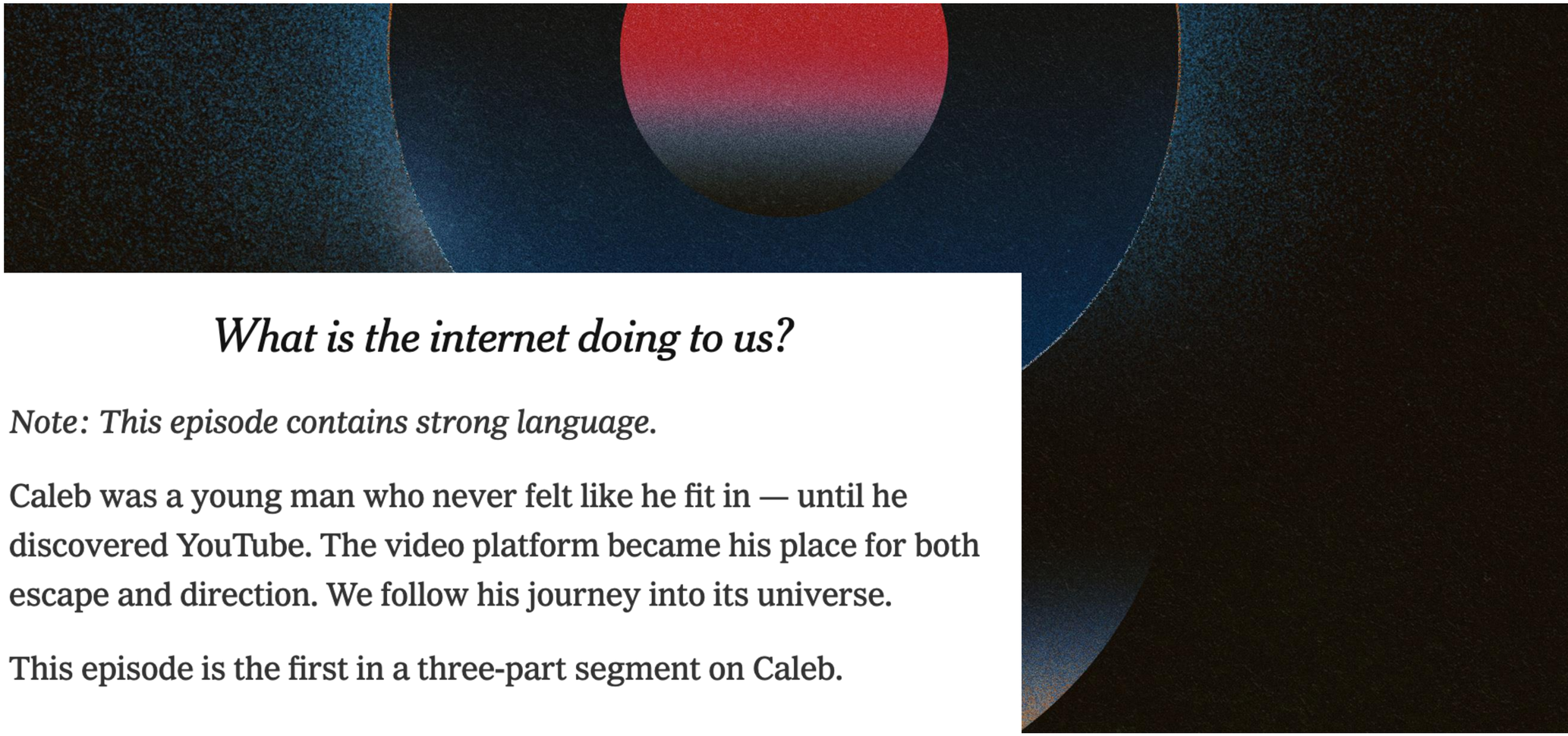
Prof Landrum said there was a need for scientists and science advocates to produce their own YouTube videos that answered and debunked the claims of flat earthers and conspiracy theorists.

"The only tool we have to battle misinformation is to try and overwhelm it with better information," said Prof Landrum.

https://www.bbc.com/news/technology-47279253

PODCAST SERIES RECOMMEDNADTION

Rabbit Hole



→ https://www.nytimes.com/column/rabbit-hole

2016 United States presidential election

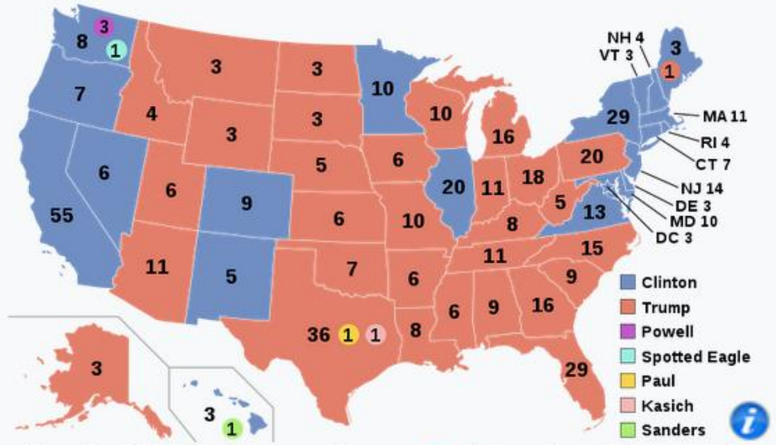
← 2012 | November 8, 2016 | 2020 →

538 members of the Electoral College
270 electoral votes needed to win

Opinion polls

Turnout 55.7%[1] ▲ 0.8 pp

Nominee	Donald Trump	Hillary Clinton
Party	Republican	Democratic
Home state	New York	New York
Running mate	Mike Pence	Tim Kaine
Electoral vote	304[a]	227[a]
States carried	30 + ME-02	20 + DC
Popular vote	62,984,828	65,853,514
Percentage	46.1%	48.2%

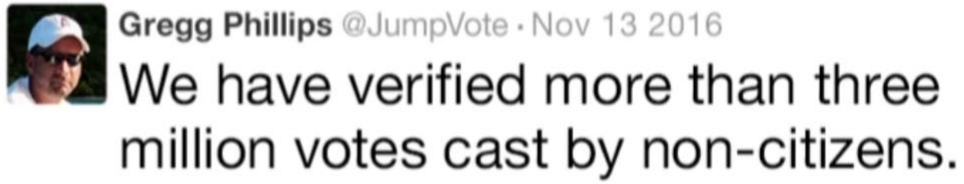


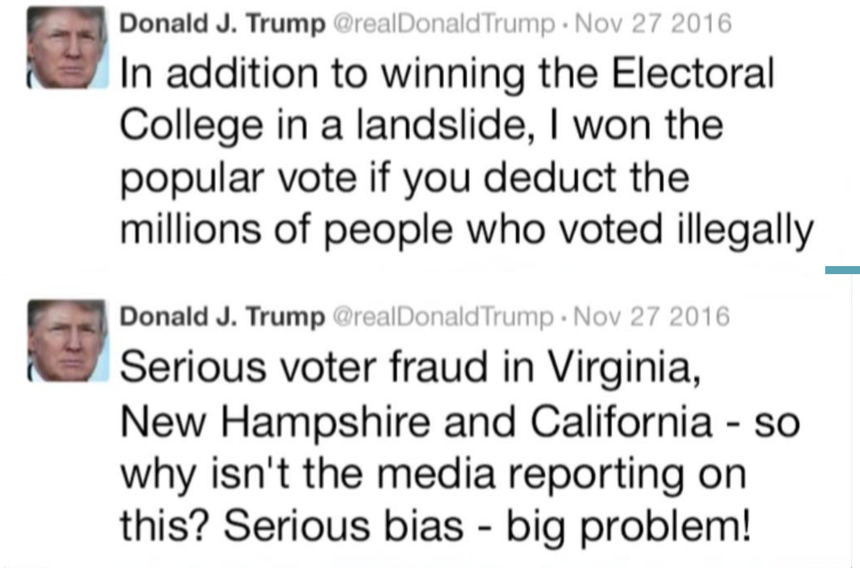
Presidential election results map. Red denotes states won by Trump/Pence and blue denotes those won by Clinton/Kaine. Numbers indicate electoral votes cast by each state and the District of Columbia. Trump received 304 and Clinton 227, as 7 faithless electors, 2 pledged to Trump and 5 to Clinton, voted for other candidates.

https://en.wikipedia.org/wiki/2016_United_States_presidential_election

FILTER BUBBLES AT WORK

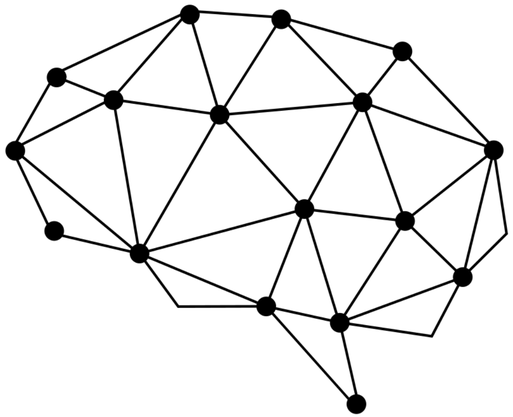
A Real Life Example





Source: https://youtu.be/xecEV4dSAXE?t=14m12s

CAMBRIDGE ANALYTICA AND THE 2016 US PRESIDENTIAL ELECTION



Cambridge Analytica

https://en.wikipedia.org/wiki/Cambridge_Analytica#/media/File:Cambridge_Analytica_logo.svg

https://www.cnet.com/news/facebook-cambridge-analytica-data-mining-and-trump-what-you-need-to-know/

c|net COVID-19 BEST PRODUCTS REVIEWS NEWS HOW TO FINANCE HEALTH SMART HOME CARS

Facebook, Cambridge Analytica and data mining: What you need to know

The world's biggest social network is at the center of an international scandal involving voter data, the 2016 US presidential election and Brexit.

Ian Sherr April 18, 2018 5:10 p.m. PT

ES 65

Consultants working for Donald Trump's presidential campaign exploited the personal Facebook data of millions.

Last month, The New York Times and the UK's Guardian and Observer newspapers broke news the social networking giant was duped by researchers, who reportedly gained access to the data of millions

Facebook CEO Mark Zuckerberg

James Martin/CNET

Manage Scripts

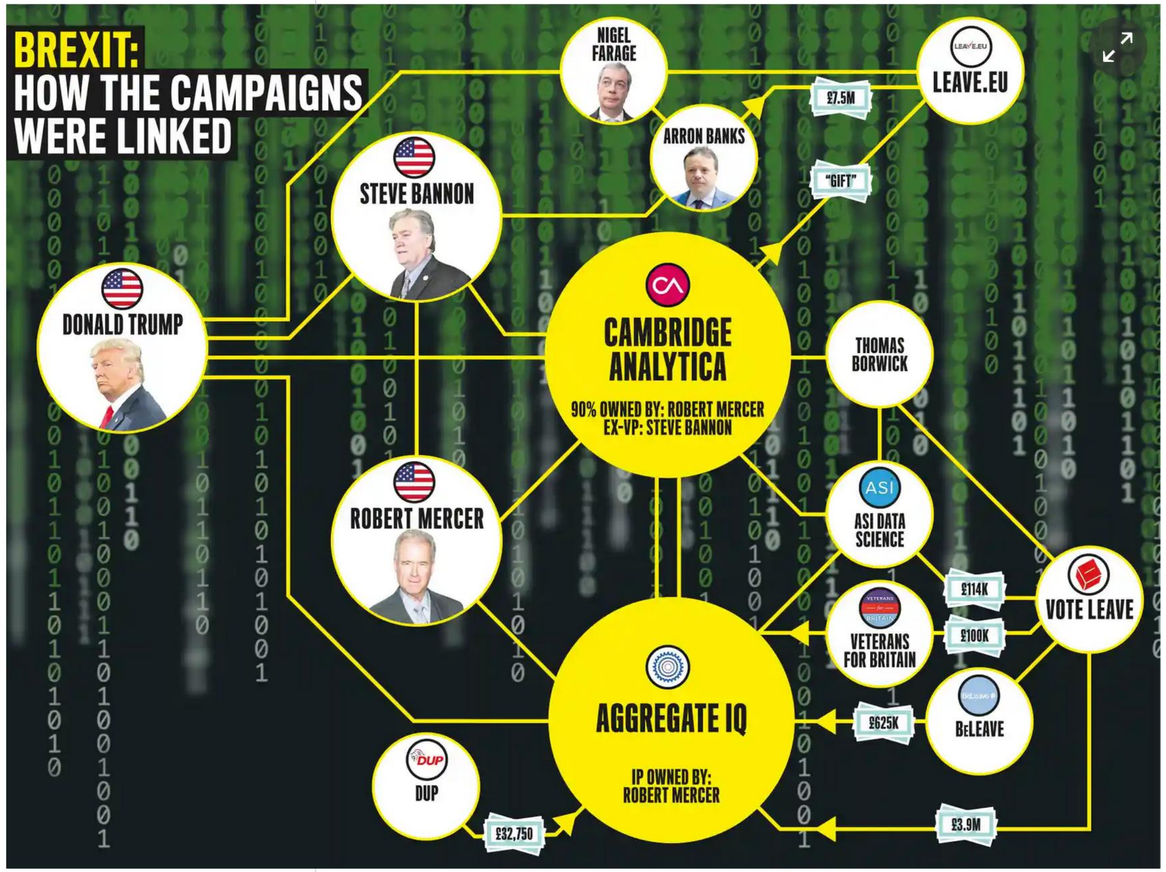
AGGREGATEiQ, CAMBRIDGE ANALYTICA, AND THE BREIXT VOTE

AGGREGATEiQ

PURE - DATA - INTELLIGENCE

https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy

Cambridge Analytica. AggregateIQ built and ran Cambridge Analytica's database platforms."



▲ Illustration: James Melaugh

Cummings won't say who did his modelling. But invoices lodged with the Electoral Commission show payments to a company called Advanced Skills

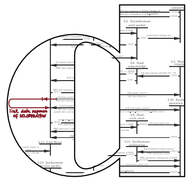
CAMBRIDGE ANALYTICA AND OUR OWN FILTER BUBBLE

https://www.dasmagazin.ch/2016/12/03/ich-habe-nur-gezeigt-dass-es-die-bombe-gibt/

Ethics for Nerds

An Advanced Course in Computer Science
Summer Semester 2020

Current Topics 2.4
An argument for our original claim
& who should do what against filter bubbles?



Prof. Holger Hermanns,
Kevin Baum, Sarah Sterz

UNIVERSITÄT
DES
SAARLANDES

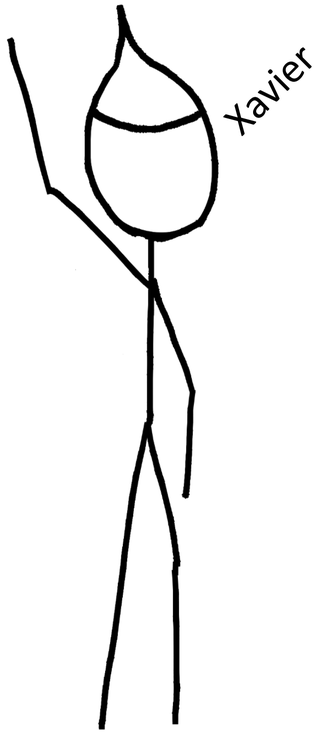
The Argument:

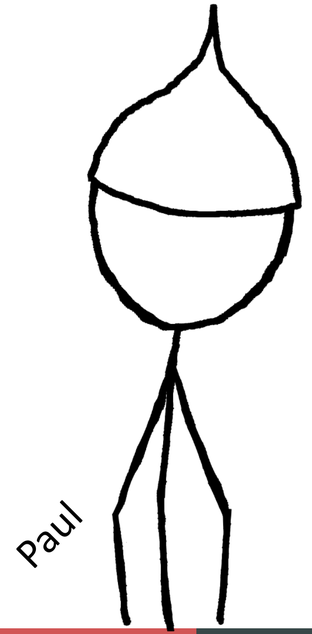
Can filter bubbles make it rational$_{epis}$ to believe in otherwise unjustified$_{epis}$ stuff?

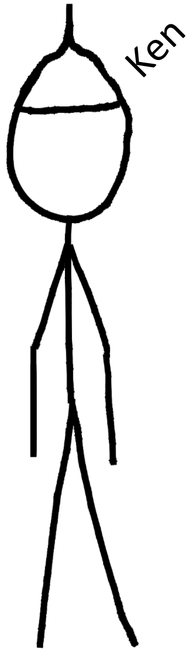
GENERAL IDEA

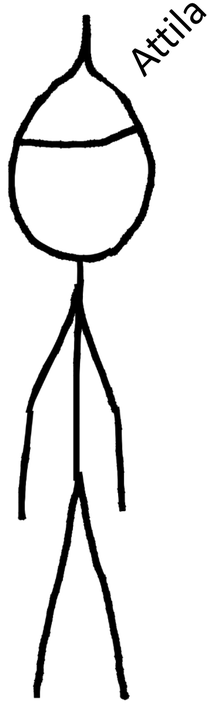
Before we present an argument for that claim, we need some preparations…

Putting a tin foil hat on his head **seems** to be **pretty reasonable** to Paul.









FILTER BUBBLE AND THE JUSTIFIED BELIEF IN FALSE POSITIONS

P1: **There are propositions such that**

ceteris paribus no subject$_{epis}$ with sufficient intellectual abilities and enough rationality$_{epis}$ is justified$_{epis}$ to believe these propositions,

and

for some such propositions and such subjects$_{epis}$, filter algorithm-related effects can make it justified$_{epis}$ to believe in these propositions.

FILTER BUBBLE AND THE JUSTIFIED BELIEF IN FALSE POSITIONS

P1: **There are propositions such that**

ceteris paribus no subject$_{epis}$ with sufficient intellectual abilities and enough rationality$_{epis}$ is justified$_{epis}$ to believe these propositions,

and

for some such propositions and such subjects$_{epis}$, filter algorithm-related effects can make it justified$_{epis}$ to believe in these propositions.

(We might call this 'filter bubble situations')

P2: **If** P1, **then** filter algorithms can lead to circumstances where it would be in-apt to criticize such subjects$_{epis}$ even though they believe in such propositions.

FILTER BUBBLE AND THE JUSTIFIED BELIEF IN FALSE POSITIONS

P1: **There are propositions such that**
ceteris paribus no subject$_{epis}$ with sufficient intellectual abilities and enough rationality$_{epis}$ is justified$_{epis}$ to believe these propositions,
and
for some such propositions and such subjects$_{epis}$, filter algorithm-related effects can make it justified$_{epis}$ to believe in these propositions.

(We might call this 'filter bubble situations')

P2: **If** P1, **then** filter algorithms can lead to circumstances where it would be in-apt to criticize such subjects$_{epis}$ even though they believe in such propositions.

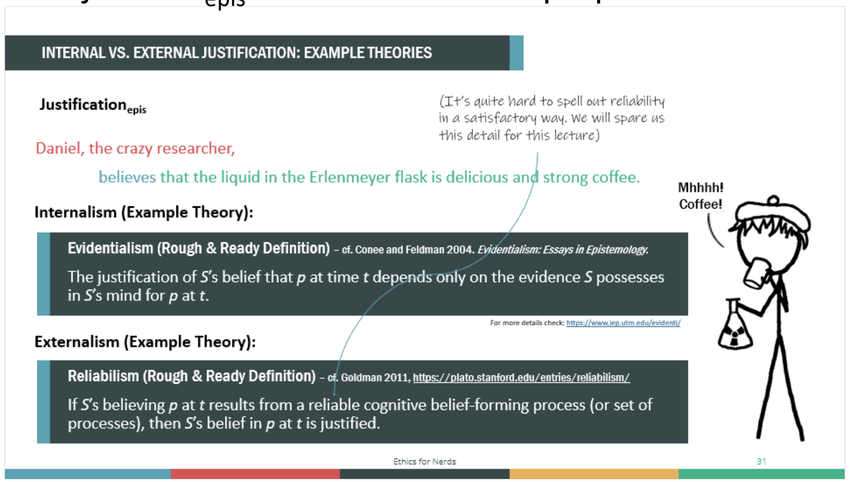
C1: Filter algorithms can lead to circumstances where it would be in-apt to criticize subjects with sufficient intellectual abilities and enough rationality$_{epis}$ even though they believe in propositions that ceteris paribus no subject$_{epis}$ with sufficient intellectual abilities and enough rationality$_{epis}$ is justified$_{epis}$ to believe in these propositions.

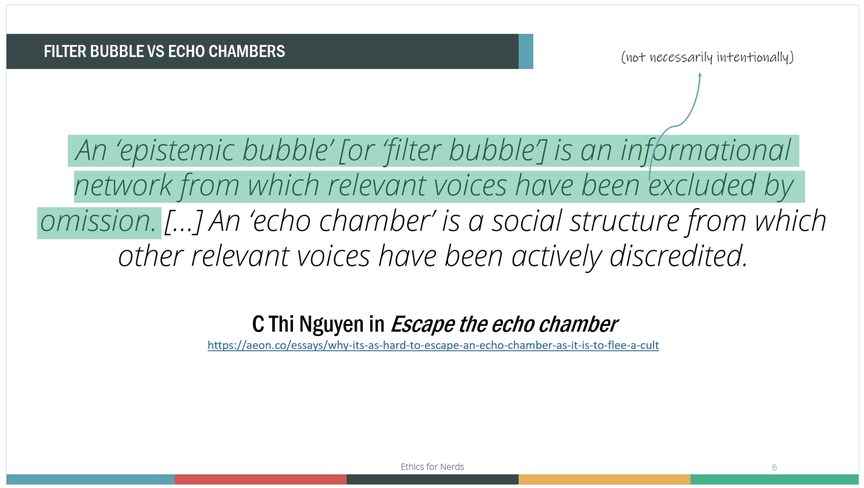
FILTER BUBBLES AND EVIDENCE

P1: **There are propositions such that**
ceteris paribus no subject$_{epis}$ with sufficient intellectual abilities and enough rationality$_{epis}$ is justified$_{epis}$ to believe these propositions,
and
for some such propositions and such subjects$_{epis}$, filter algorithm-related effects can make it justified$_{epis}$ to believe in these propositions.

Why to believe P1?

- The first conjunct is plausible enough and does not need further elaboration.
- The second conjunct does.
- What do we need?
  1. an account of justification$_{epis}$ (of which we introduced two, we restrict ourselves to the easier to apply variant: Evidentialism)
  2. some understanding of what filter bubbles are.





FILTER BUBBLES AND EVIDENCE

Evidentialism (Rough & Ready Definition) – cf. Conee and Feldman 2004. *Evidentialism: Essays in Epistemology*.

The justification of *S*'s belief that *p* at time *t* depends only on the evidence *S* possesses in *S*'s mind for *p* at *t*.

What counts as evidence? An example:

(1) If you read an information in a certain newspaper, you can take that as evidence that this information is true.

(2) If you read the information in more than one newspaper (token from another type), this is even better evidence.

(3) But if you read the information in multiple copies of the *same* newspaper (different tokens of the same kind), you have no more evidence than in (1).

https://pixnio.com/objects/newspaper-paper-pile-information-journalism-news

FILTER BUBBLES AND EVIDENCE

Evidentialism (Rough & Ready Definition) – cf. Conee and Feldman 2004. *Evidentialism: Essays in Epistemology.*

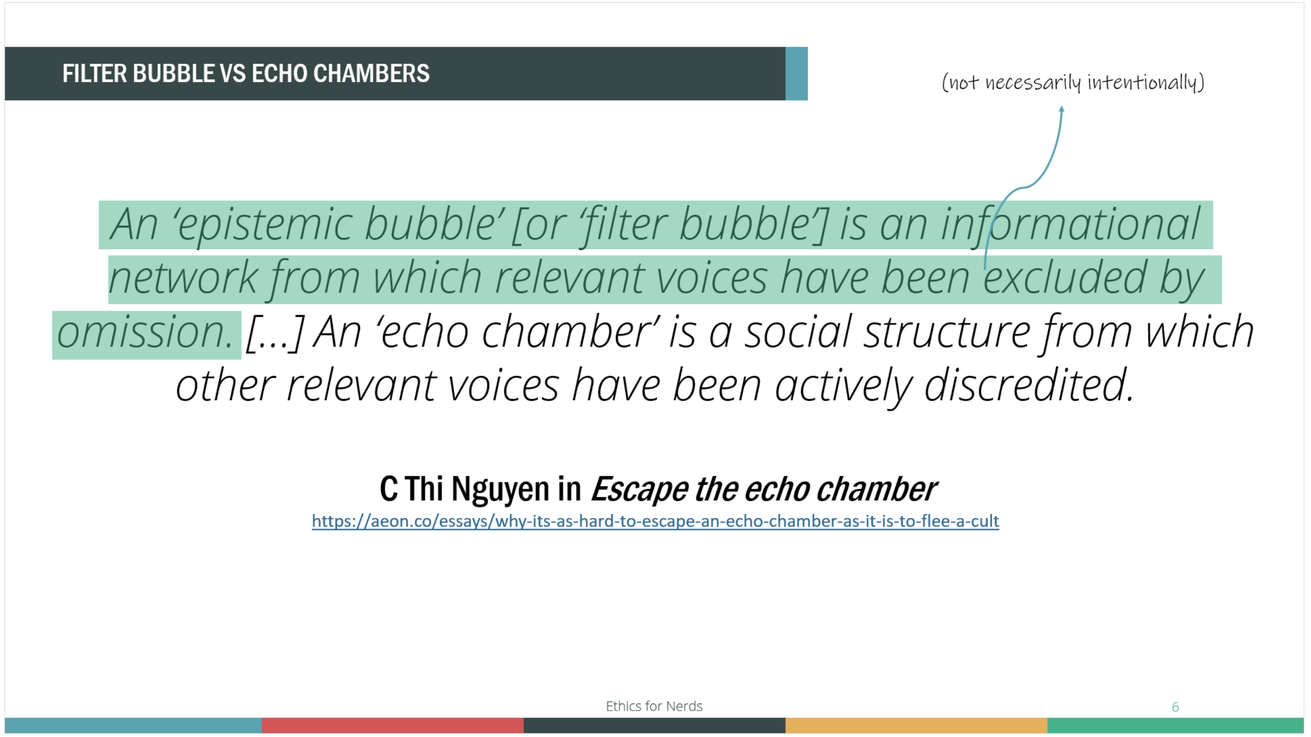
The justification of *S*'s belief that *p* at time *t* depends only on the evidence *S* possesses in *S*'s mind for *p* at *t*.

Example:

(1) If you find an information thanks to/with a certain algorithm, you can take that as evidence that this information is true.

(2) If you find distinct sources supporting the same information thanks to/with a certain algorithm, is this even better evidence?

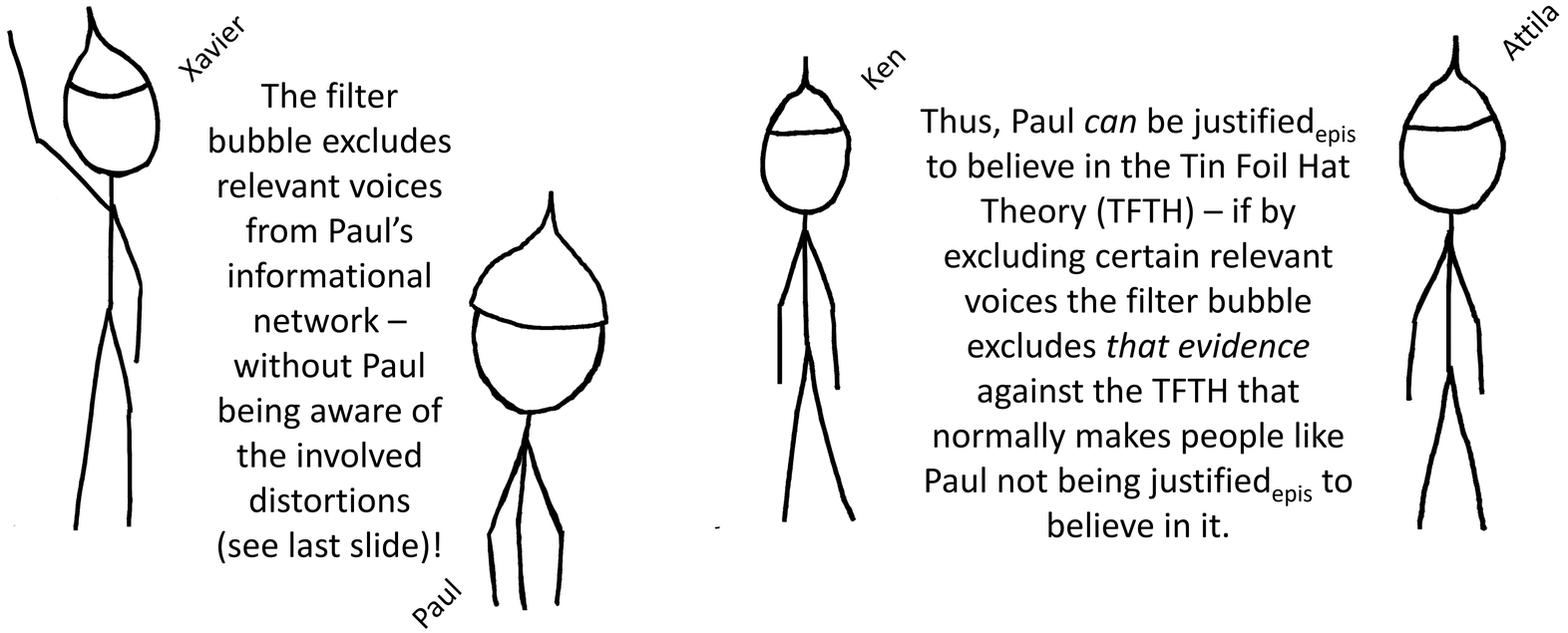
- It depends! If the algorithms exclude counter-evidence or if the algorithm makes it hard for you to see that the sources are heavily dependent on each other, then certainly not.
- It seems that there are incentives to do so in certain context (regarding certain users and information), because this will increase engagement!



FILTER BUBBLES AND EVIDENCE

P1: **There are propositions such that**
ceteris paribus no subject$_{epis}$ with sufficient intellectual abilities and enough rationality$_{epis}$ is justified$_{epis}$ to believe these propositions,
and
for some such propositions and such subjects$_{epis}$, filter algorithm-related effects can make it justified$_{epis}$ to believe in these propositions.

Why to believe P1? With Evidentialism



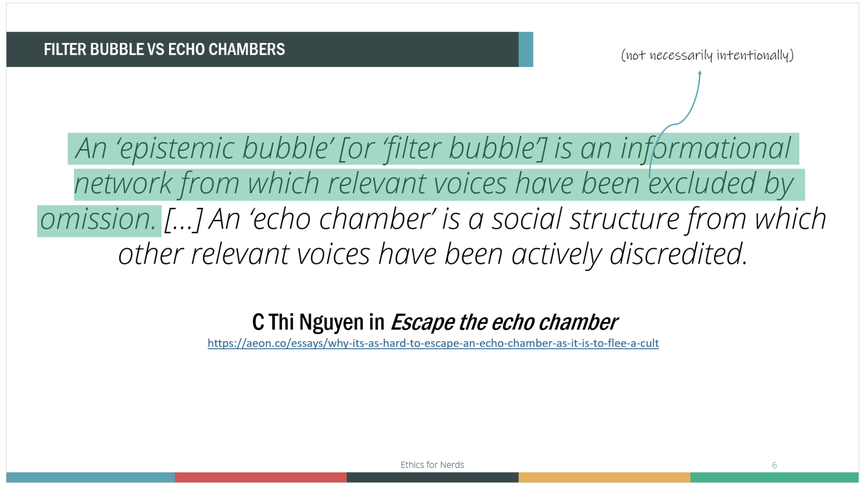
The filter bubble excludes relevant voices from Paul's informational network – without Paul being aware of the involved distortions (see last slide)!

Thus, Paul *can* be justified$_{epis}$ to believe in the Tin Foil Hat Theory (TFTH) – if by excluding certain relevant voices the filter bubble excludes *that evidence* against the TFTH that normally makes people like Paul not being justified$_{epis}$ to believe in it.

Evidentialism (Rough & Ready Definition) – cf. Conee and Feldman 2004. *Evidentialism: Essays in Epistemology.*

The justification of *S*'s belief that *p* at time *t* depends only on the evidence *S* possesses in *S*'s mind for *p* at *t.*

Therefore: For some propositions which are such that ceteris paribus no subject$_{epis}$ with sufficient intellectual abilities and enough rationality$_{epis}$ is justified$_{epis}$ to believe them, filter algorithm-related effects can make it justified$_{epis}$ to believe them



REMARKS ON THE ARGUMENT

- There are accounts of justification$_{epis}$ do not support P1 (i.e., the so-called *responsibility view*).
- The argument's conclusion is *morally* significant, for instance,…
  - because it undermines autonomy (it interferes with our access to the facts and, thus, we become easier to manipulate and less capable to make decisions in our own interest);
  - because it undermines public discourse and threatens democracy (loss of common agreement about facts; increasing polarization; enables hidden campaigns and manipulation of voters);
- → we will return in forthcoming lectures to these topics.

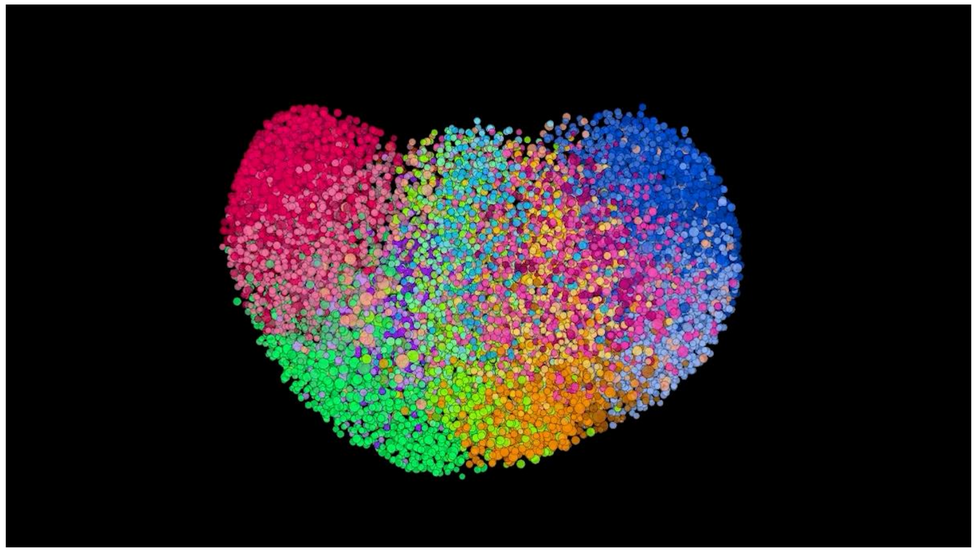
Episteme, 10, 2 (2013) 117–134 © Cambridge University Press
doi:10.1017/epi.2013.11

JUSTIFIED BELIEF IN A DIGITAL AGE: ON THE EPISTEMIC IMPLICATIONS OF SECRET INTERNET TECHNOLOGIES

BOAZ MILLER AND ISAAC RECORD
boaz.miller@gmail.com and isaac.record@gmail.com

Whether a subject's belief that *p* is justified at time *t* may depend on whether he performed certain relevant actions prior to *t*. That is, sometimes responsible subjects are not only required to form beliefs responsibly, but also to have conducted some inquiry before they come to believe that *p*. This isn't just to say that a subject's ignorance of the shortcomings of his evidence is not always a valid defence against the charge of being epistemically irresponsible. Rather, it is to say that some forms of ignorance are culpable, and the subject may be blamed on epistemic grounds.

APPROACHES



https://www.technologyreview.com/2018/08/22/140661/this-is-what-filter-bubbles-actually-look-like/

Who can do what against these problems?

Who can solve these problems?

Users

- If users behaved in certain ways, we probably would not have filter bubbles and echo chamber effects, or they would not be so bad.
- Users can partially solve this problem (especially if they are aware of the bubble and what to do against it)
- Some things they cannot change…

Online services

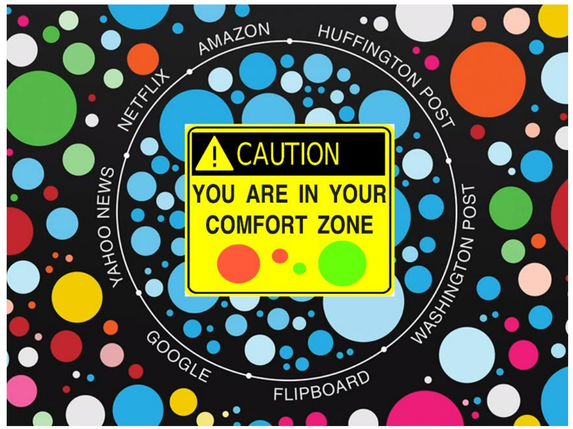
- If the implementations were changed in certain ways or countermeasures were taken, we probably would not have filter bubbles and echo chamber effects.
- The question remains: change in what way and take what countermeasures?
- Online services can help solve this problem.

Society

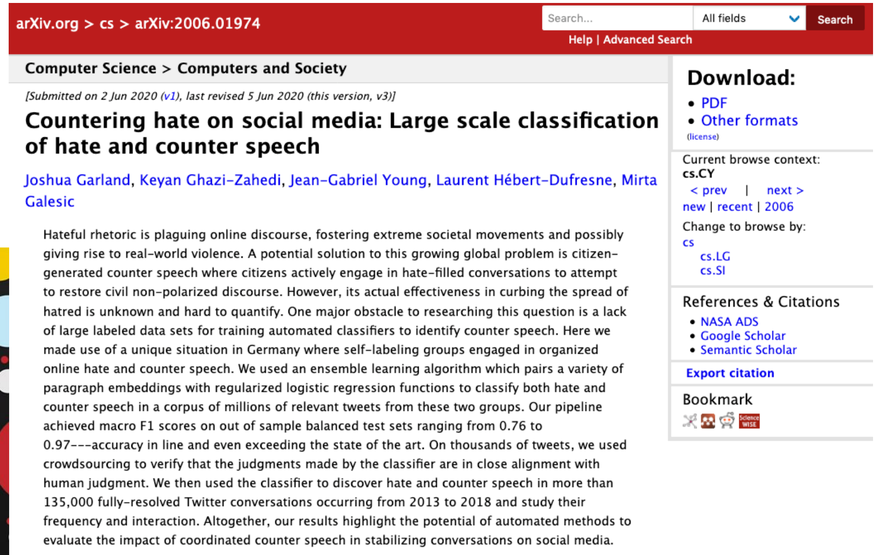
- If people were more educated, less lonely, and digitally literate, they might be less vulnerable to the effects of filter bubbles and echo chambers.
- This will take time, but the problem is already here.
- And how to achieve that?
- Should be done but cannot be everything that should be done.

FILTER BUBBLES: USERS

- Don’t be part of the problem I: Do not upload, repost, share bad stuff intentionally. Maybe counter fake news, hate, and tribalism.
- Don’t be part of the problem II (anti-personalization):
  - Don’t let them track you (use anti-tracking software!)
  - Obfuscation and active training: Maybe, read, share, ‘like’ stuff of sources with contrary, but still acceptable viewpoints to train a more diverse filter algorithm.
- Set yourself free:
  - Go analog, go outside!
  - Talk to people with other opinions, read newspapers one normally would not read…
  - Deliberately try to exit the bubble!



https://medium.com/filter-bubbles-vs-democracy-in-the-age-of-social/filter-bubbles-vs-democracy-5b0e4fae6837



https://arxiv.org/abs/2006.01974

FILTER BUBBLES: USERS

Is that a suitable solution?

- *in theory*: probably yes, if sufficiently enough users would fulfill one's duties to do so
- *in reality*: lots of people will probably refuse to do so, are already too deep in, or do not have the required skills and knowledge → not viable in real life; lacks scalability

https://twitter.com/elonmusk/status/1262076474565242880

the problem is so important, that we want to have a real-life solution!

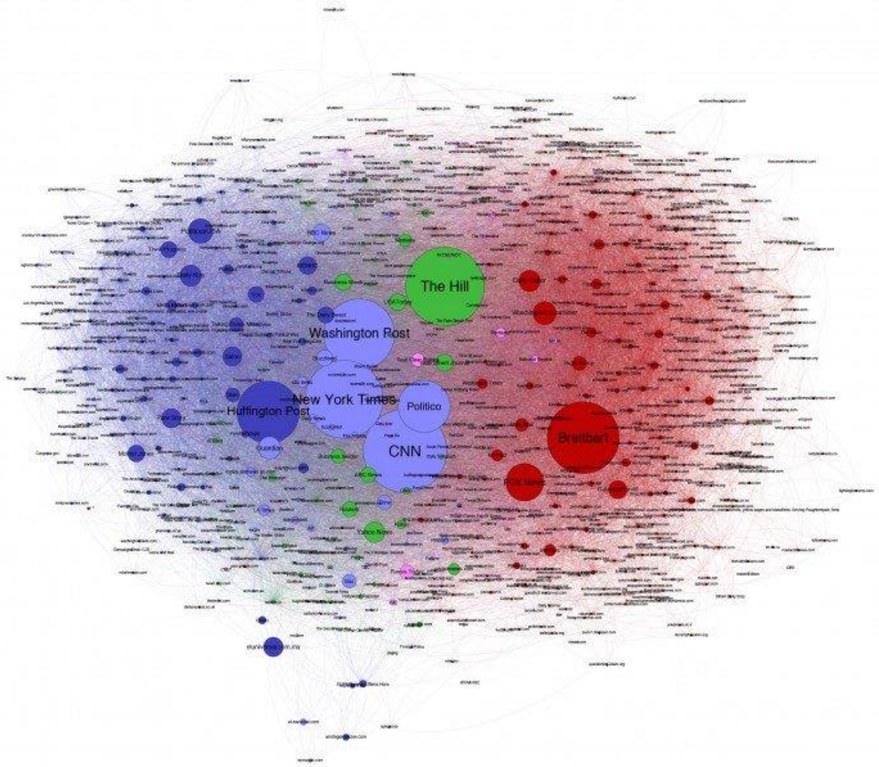
https://www.technologyreview.com/2018/08/22/140661/this-is-what-filter-bubbles-actually-look-like/

FILTER BUBBLES: DEVELOPERS AND SERVICE PROVIDERS

- Can we find an implementation
  - that has most of the benefits of filter algorithms,
  - but eliminates most of the drawbacks?
- (If we do not find one: Switch social media off/forbid it?)

→ some approaches

- fact checker and elimination of “fake news”
- beneath every post that includes an opinion of a certain topic, show a contrary opinion for the same topic…



MOST RECENT DEVELOPMENT/COUNTER MEASURE

Donald Trump

Twitter labels Trump's false claims with warning for first time

President rails against decision after his tweets on mail-in voting are marked with message: 'Get the facts'

Julia Carrie Wong *in San Francisco and* Sam Levine *in New York*

Wed 27 May 2020 01.02 BST

4,133

▲ Some of Donald Trump's tweets now feature a link highlighting false claims. Photograph: Jonathan Ernst/Reuters

Twitter has for the first time taken action against a series of tweets by Donald Trump, labeling them with a warning sign and providing a link to further information.

Since ascending to the US presidency, Trump has used his Twitter account to threaten a world leader with war, amplify racist misinformation by British hate figures and, as recently as Tuesday morning, spread a lie about the 2001 death of a congressional aide in order to smear a cable news pundit. Throughout it all, Twitter has remained steadfast in its refusal to censor the head of state, even going so far as to write a new policy to allow itself to leave up tweets by "world leaders" that violate its rules.

https://www.theguardian.com/us-news/2020/may/26/trump-twitter-fact-check-warning-label

https://twitter.com/realDonaldTrump/status/1265255835124539392

Donald J. Trump ✓
@realDonaldTrump

There is NO WAY (ZERO!) that Mail-In Ballots will be anything less than substantially fraudulent. Mail boxes will be robbed, ballots will be forged & even illegally printed out & fraudulently signed. The Governor of California is sending Ballots to millions of people, anyone.....

Tweet übersetzen

Get the facts about mail-in ballots

2:17 nachm. · 26. Mai 2020 · Twitter for iPhone

50.326 Retweets **130.307** „Gefällt mir"-Angaben

Donald J. Trump ✓ @realDonaldTrump · 26. Mai
Antwort an @realDonaldTrump

....living in the state, no matter who they are or how they got there, will get one. That will be followed up with professionals telling all of these people, many of whom have never even thought of voting before, how, and for whom, to vote. This will be a Rigged Election. No way!

Get the facts about mail-in ballots

https://twitter.com/i/events/1265330601034256384

Trump makes unsubstantiated claim that mail-...

Event · 26. Mai 2020

Trump makes unsubstantiated claim that mail-in ballots will lead to voter fraud

On Tuesday, President Trump made a series of claims about potential voter fraud after California Governor Gavin Newsom announced an effort to expand mail-in voting in California during the COVID-19 pandemic. These claims are unsubstantiated, according to CNN, Washington Post and others. Experts say mail-in ballots are very rarely linked to voter fraud.
Foto via @CNNPolitics

What you need to know

- Trump claimed that mail-in ballots would lead to "a Rigged Election." However, fact-checkers say there is no evidence that mail-in ballots are linked to voter fraud.
- Trump falsely claimed that California will send mail-in ballots to "anyone living in the state, no matter who they are or how they got there." In fact, only registered voters will receive ballots.
- Five states already vote entirely by mail and all states offer some form of mail-in absentee voting, according to NBC News.

The Hill @thehill · 26. Mai
Trump blasts California over mail-in voting following Republican lawsuit
hill.cm/4FO4sqH

2.239 5.382 19.955

"Trump, who has increasingly leveled unsubstantiated claims about widespread fraud in mail-in voting, zeroed in on California in a string of tweets Tuesday morning, two days after the Republican National Committee (RNC) sued [Governor Gavin] Newsom over his effort to expand mail-in voting in California during the coronavirus pandemic." — The Hill

Jennifer Jacobs @JenniferJJacobs · 26. Mai
Trump says mail-in voting system in California will be "substantially

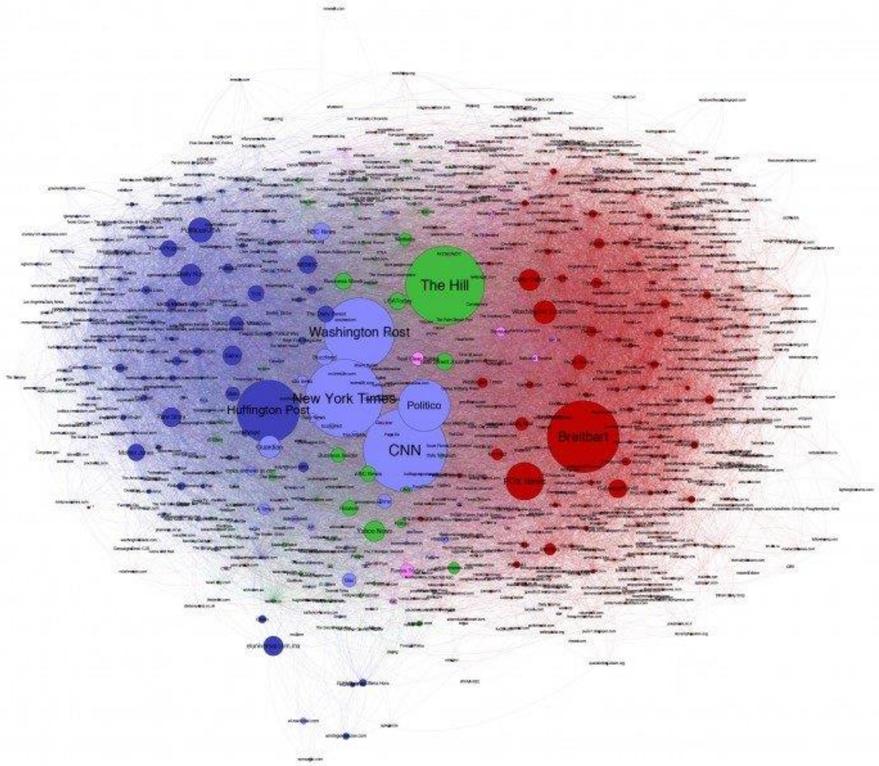
https://www.technologyreview.com/2018/08/22/140661/this-is-what-filter-bubbles-actually-look-like/

FILTER BUBBLES: DEVELOPERS AND SERVICE PROVIDERS

- Can we find an implementation
  - that has most of the benefits of filter algorithms,
  - but eliminates most of the drawbacks?
- (If we do not find one: Switch social media off/forbid it?)

→ some approaches

- fact checker and elimination of “fake news”
- beneath every post that includes an opinion of a certain topic, show a contrary opinion for the same topic
- Accept/embrace liability and responsibility for contents.
- Make the algorithms less personalized.
- ….



Are these suitable approaches?

At least some seem to be, since the approach would be implementable and could somehow (e.g. through law) be enforced on online services, if they are not willing to implement it…