

gesis

Leibniz-Institut
für Sozialwissenschaften



Auditing Algorithms: How Platform Technologies Shape our Digital Environment

Meet the Experts! – GESIS online talks

Dr. Roberto Ulloa · October 7, 2021

Speaker



Dr. Roberto Ulloa

- Researcher in the team Social Analytics and Services, Department Computational Social Science
- Interested in Digital Institutions, Computer Simulations, Algorithm Auditing
- Contact: roberto.ulloa@gesis.org

gesis

Leibniz-Institut
für Sozialwissenschaften



Auditing Algorithms: How Platform Technologies Shape our Digital Environment

Meet the Experts! – GESIS online talks

Dr. Roberto Ulloa · October 7, 2021

Agenda

- How do platform technologies shape our digital environment?
- Why do these problems occur?
- Algorithm auditing
- General considerations
- Takeaways and final remarks

How do platform technologies shape our digital environment?

Google Results for “Black girls” in Google (Noble, 2013; Noble, 2018)

- all Ads and 5 results link to pornography
- and 2 to a UK rock band
- 3 are relevant (ranked lower)

About 140,000,000 results (0.07 seconds) [Advanced search](#)

Sugary Black Pussy .com-Black girls in a hardcore action galleries
[sugaryblackpussy.com/](#) - Cached
 (black pussy and hairy black pussy,black sex,black booty,black ass,black teen pussy,big black ass,black porn star,hot black girl) ...

Black Girls -- ((100% Free Black Girls Chat))
[www.woome.com/people/girls/crowds/black/](#) - Cached
Black Girls Online // (100% Free Black Girls Chat) -- Black Girl Chat Rooms, Meet a Black Girl Online Now!!

Black Girls | Big Booty Black Girls | Black Porn | Black Pussy
[www.blackgirls.com/](#) - Cached
BlackGirls.com is the top spots for black porn online. Hottest big Booty black girls sucking black cocks, in black ebony porn movies.

HOME | THE OFFICIAL HOME OF BLACK GIRLS ROCK!
[www.blackgirlsrockinc.com/](#) - Cached
 Jun 24, 2011 – **BLACK GIRLS ROCK!** Inc. is 501(c)3 non-profit youth empowerment and mentoring organization established to promote the arts for young ...

Two black girls love cock | Redtube Free Big Tits Porn Videos, Anal ...
[www.redtube.com/7310](#) - Cached
 Watch Two **black girls** love cock on Redtube Home of free big tits porn videos, anal movies & group clips.

Black Girls | Free Music, Tour Dates, Photos, Videos
[www.myspace.com/blackgirlsband](#) - Cached
Black Girls's official profile including the latest music, albums, songs, music videos and more updates.

BOOTY ON THE BEACH, BLACK GIRLS GONE WILD, GOONCITY ...
[www.youtube.com/watch?v=h7lqV7z8Wrs](#) - Cached
 Mar 11, 2010 – DJ NOLAN AND FANS HIT THE BEACH ,GOONCITYDANCE.COM , I JUST SHOW LOVE TO MY FRIENDS, GET THE DVD IT HAS MORE ...

Black Girl Problems.
[black-girl-problems.tumblr.com/](#) - Cached
 The problems **black girls** have. Some of its funny, some of its serious. Click the follow button, you know you want to. twitter: @blackgirlprobss people can relate.

Black Girls | Facebook
[www.facebook.com/blackgirlsband](#) - Cached
 Sat, Sep 24, 2011 - NYC
Black Girls - follow us!!! get ready for the seafood special spring break tour 2k11 - General Manager: Erica - Booking Agent: blackgirlsbooking@gmail.com ...

Black Girl with Long Hair
[bglhonline.com/](#) - Cached
 18 September 2011 ~ Posted By **Black Girl With Long Hair** ~ 83 Comments by ERIKA NICOLE KENDALL of A **Black Girl's** Guide to Weight Loss. Earlier ...

Ads

Hot Black Dating
[www.blackcrush.com](#)
 Hook Up Tonight & Get Busy with a Hot **Black Girl** Near You. Join Free

Local Ebony Sex
[www.amateurmatch.com](#)
 The Sexiest Ebony Dating Online. Chat Browse and Get Laid. Free Join

Black Women Seeking Men
[www.blacksexmatch.com](#)
 Find **Black Women** Near You Who Want a Lover in Only 5 mins!

Big Booty Black Porn
[www.bigbootyblackvideos.com](#)
 A must see **black** booty porn site. Watch uncensored videos - 100% Free

Black XXX - uncensored
[www.dabigblackdonkeybooty.com](#)
 Hardcore **Black** Porn tube videos. Extremely good - 100% Free.

Black Girls
[www.aebn.net](#)
 Watch **Black** Adult PayPerView Choose From Over 100,000 Porn Films

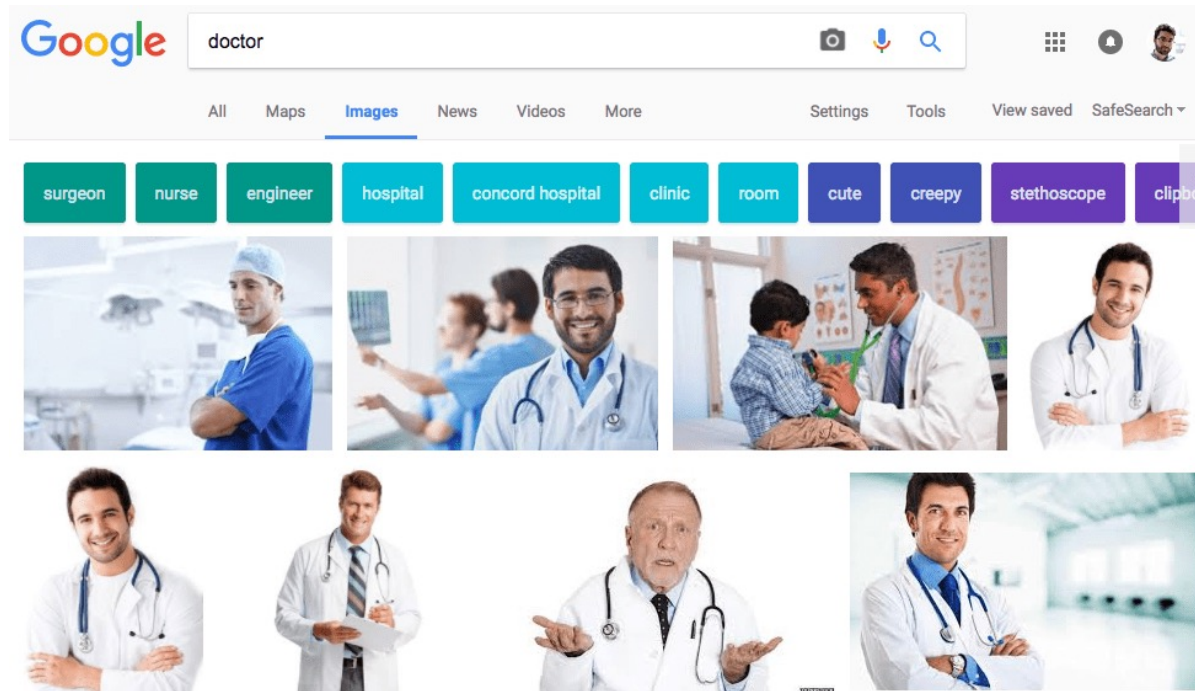
Naughty Black Wives
[www.affairsclub.com/Black](#)
 Husband Out For Work: You In For Naughty Pleasure! Join For Free.

[See your ad here >](#)

Noble, S. (2013). Google Search: Hyper-visibility as a Means of Rendering Black Women and Girls Invisible. *InVisible Culture*.
 Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.

6

Google Under-representation of women in image search results (Kay, 2015)



- gender stereotypes exaggeration (compared to the U.S. Bureau of Labor and Statistics)
- effect on perceptions (shifting estimations ~7%)

Source: <https://qz.com/958666/the-reason-why-most-of-the-images-are-men-when-you-search-for-doctor/>

 Microsoft Bing **Under-representation of women in image search results**


- **1983** (Archer et. al): 1/3 women in 4 magazines




- **2017** (Otterbacher et al): 1/3 women



- **2021** (Ulloa et al):

 Microsoft Bing 1/3 women for “person”
1/3 women for “intelligent person”

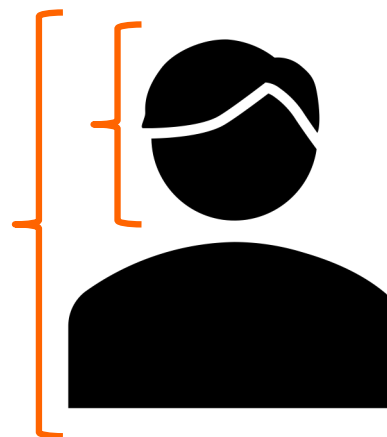
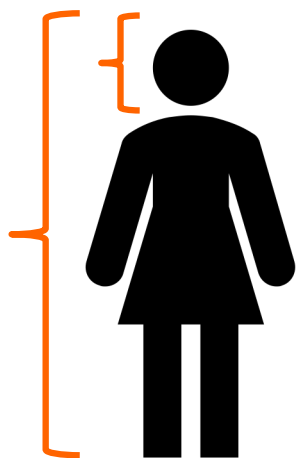
 **Google** half pictures were of women for “person”
1/3 women for “intelligent person”

Archer, D., Iritani, B., Kimes, D. D., & Barrios, M. (1983). Face-ism: Five studies of sex differences in facial prominence. *Journal of Personality and Social Psychology*, 45(4), 725–735.

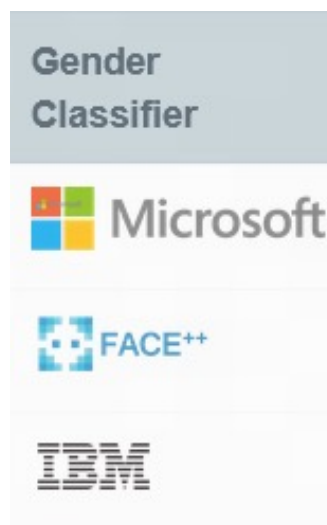
Otterbacher, J., Checco, A., Demartini, G., & Clough, P. (2018). Investigating User Perception of Gender Bias in Image Search: The Role of Sexism. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (pp. 933–936). Association for Computing Machinery.

Ulloa, R., Richter, A. C., Makhortykh, M., Urman, A., & Kacperski, C. (n.d.). Representativeness and Face-ism: Gender Bias in Image Search. *Unpublished Manuscript*.

Face-ism: women are represented with a lower face-to-body ratio (Archer et. al, 1983)

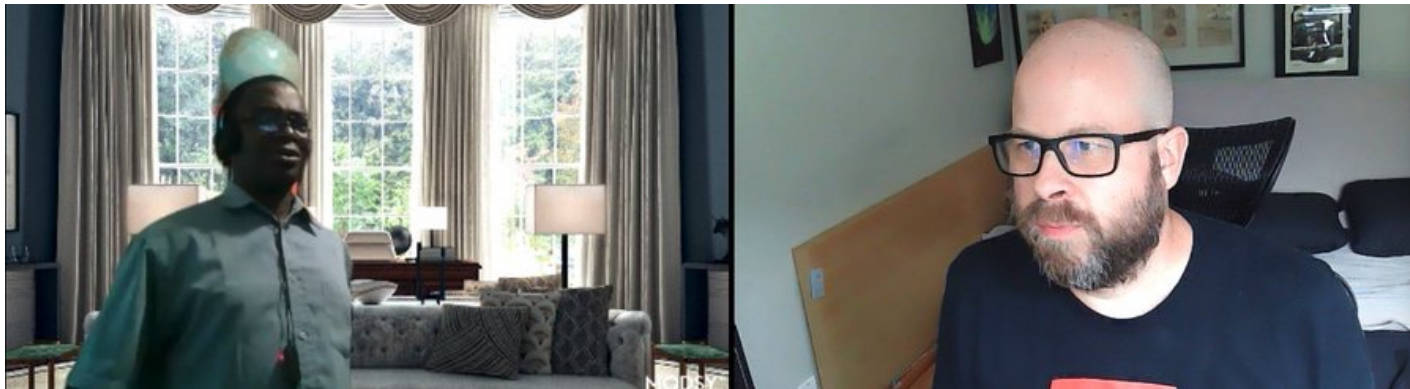


Gender Shades: gender recognition performed poorly for women with darker skin (Buolamwini & Gebru, 2018)



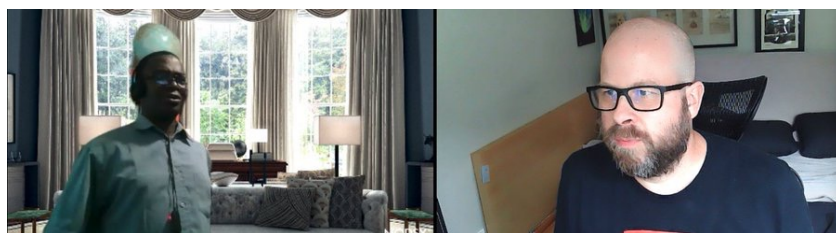


Zoom Virtual Background (September 2020)





Twitter Automatic Cropping (September 2020)



 **Colin Madland, PhD(c)** @colinmadland

Replying to @colinmadland



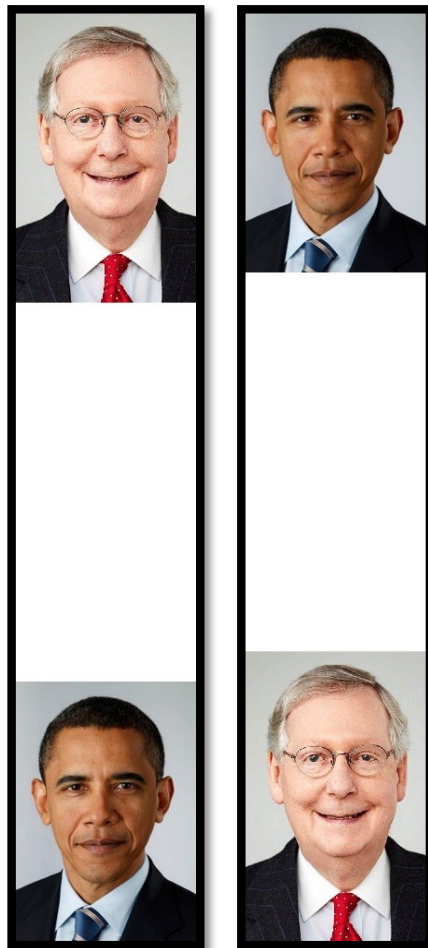
2:18 AM · Sep 19, 2020

5.8K 27 Copy link to Tweet

[Tweet your reply](#)



Twitter Automatic Cropping (September 2020)



Gender and race issues on online labor market search websites

- TaskRabbit and Fiverr (Hannak et al, 2017)
- Indeed, Monster, CareerBuider (Chen et al., 2018)

Hannák, A., Wagner, C., Garcia, D., Mislove, A., Strohmaier, M., & Wilson, C. (2017). Bias in Online Freelance Marketplaces: Evidence from TaskRabbit and Fiverr. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1914–1933.

Chen, L., Ma, R., Hannák, A., & Wilson, C. (2018). Investigating the Impact of Gender on Rank in Resume Search Engines. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14.

 Automatic labelling (Cabañas et al., 2018)



misterbnb Sponsored · 

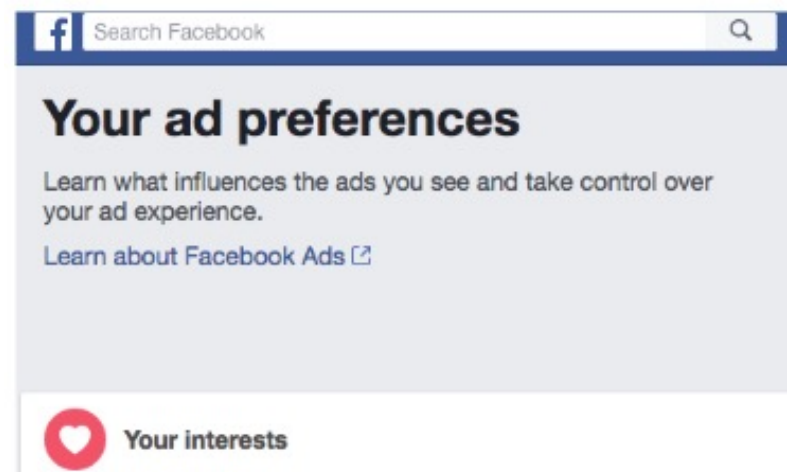
Connect with the gay community & rent affordable places from people like you. Book now!




misterbnb
Largest gay hotelier in the world

Stay Like a Gay Local
Live like a gay local - feel welcome anywhere you go in over 130 countries. Check out our top cities New York, Paris, Barcelona, and Rome.






Search Facebook 

Your ad preferences

Learn what influences the ads you see and take control over your ad experience.
[Learn about Facebook Ads](#)

 Your interests

Facebook automatic labelling (Cabañas et al., 2018)

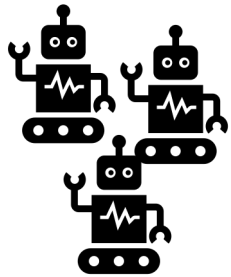
- 2092 sensitive attributes
 - For example: communism, islam, quran, suicide prevention, socialism, judaism, homosexuality, alternative medicine, christianity, illegal immigration, oncology, lgbt community, gender identity, bible, pregnancy, nationalism, veganism, buddhism, feminism
- only 0.03% were added by users (users 4577 users, >100K sensitive label instances)
- these interests are used to sell ads
- but also put people at risk:
 - hate campaigns, e.g., countries in which homosexuality is illegal
 - identification attacks (Hong, 2012)
- only 27% of labels were accurate (Bashir et al., 2019)

Cabañas, J. G., Cuevas, Á., & Cuevas, R. (2018). *Unveiling and Quantifying Facebook Exploitation of Sensitive Personal Data for Advertising Purposes*. 479–495.

Hong, J. (2012). The state of phishing attacks. *Communications of the ACM*, 55(1), 74–81.

Bashir, M., Farooq, U., Shahid, M., Zaffar, M. F., & Wilson, C. (2019). Quantity vs. Quality: Evaluating User Interest Profiles Using Ad Preference Managers. *NDSS*

Google Intransparent labelling (Datta et al., 2018)

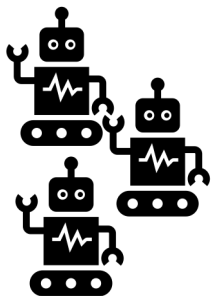


Group A
visited **substance abuse**
websites

**RQ1: substance abuse
Ads?**
yes

**RQ2: differences in the
Google Ads preferences?**

no

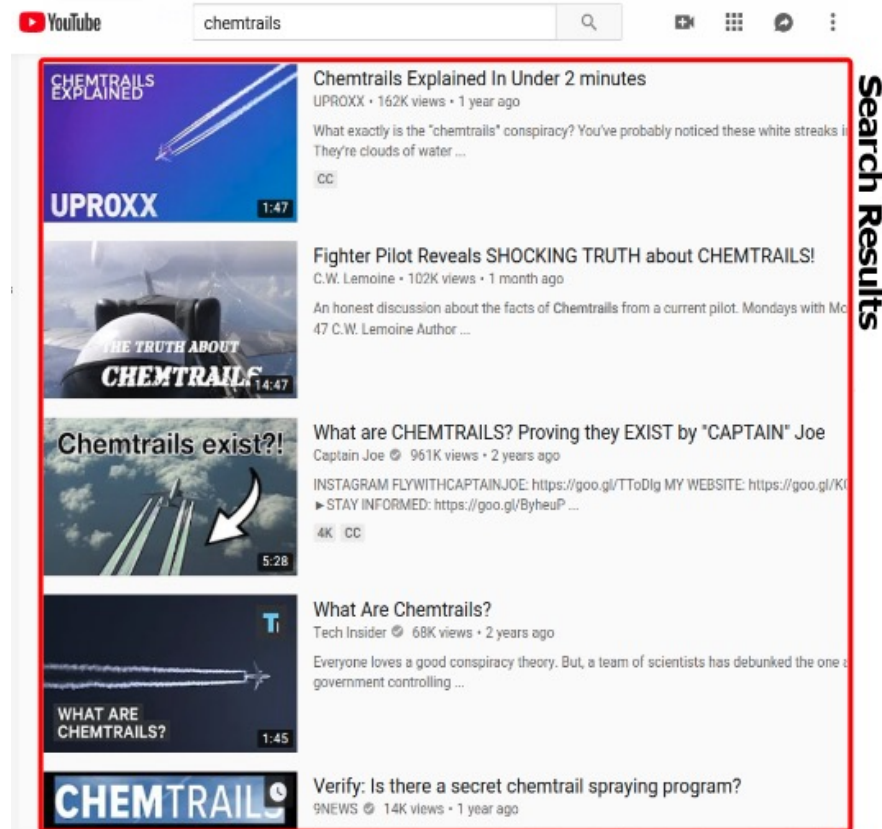
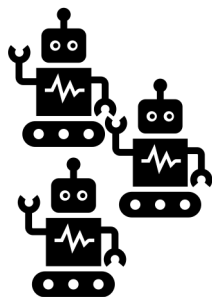
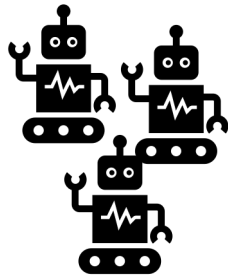


Group B
did not visit any website

no



“Filter Bubble” effect (Hussein et al., 2020)



YouTube chemtrails

CHEMTRAILS EXPLAINED
UPROXX • 162K views • 1 year ago
What exactly is the “chemtrails” conspiracy? You’ve probably noticed these white streaks in the sky. They’re clouds of water ...
CC

FIGHTER PILOT REVEALS SHOCKING TRUTH ABOUT CHEMTRAILS!
C.W. Lemoine • 102K views • 1 month ago
An honest discussion about the facts of Chemtrails from a current pilot. Mondays with Mo ...
47 C.W. Lemoine Author ...

CHEMTRAILS EXIST?!
Captain Joe • 961K views • 2 years ago
INSTAGRAM FLYWITHCAPTAINJOE: <https://goo.gl/TTToDlq> MY WEBSITE: <https://goo.gl/KI...>
► STAY INFORMED: <https://goo.gl/ByheuP> ...
4K CC

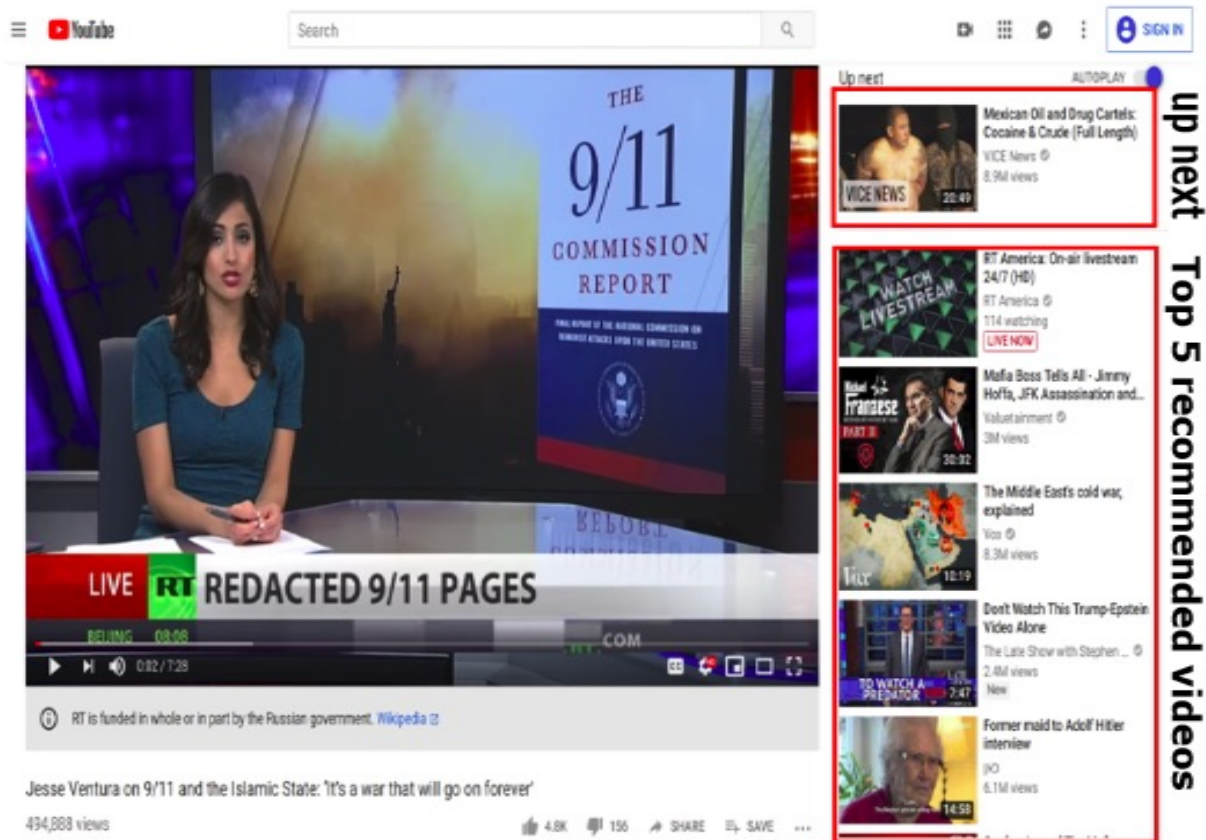
WHAT ARE CHEMTRAILS?
Tech Insider • 68K views • 2 years ago
Everyone loves a good conspiracy theory. But, a team of scientists has debunked the one ...
government controlling ...

CHEMTRAILS
9NEWS • 14K views • 1 year ago
Verify: Is there a secret chemtrail spraying program?

Search Results



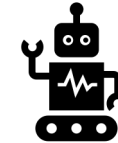
“Filter Bubble” effect (Hussein et al., 2020)



The screenshot shows a YouTube video player with a news report titled "THE 9/11 COMMISSION REPORT" and "REDACTED 9/11 PAGES". The video is from RT and has 434,883 views. To the right of the video player is a list of recommended videos, including "Mexican Oil and Drug Cartels: Cocaine & Crude (Full Length)", "RT America: On-air livestream 24/7 (HD)", "Mafia Boss Tells All - Jimmy Hoffa, JFK Assassination and...", "The Middle East's cold war, explained", "Don't Watch This Trump-Epstein Video Alone", and "Former maid to Adolf Hitler interview". A red box highlights the recommended videos, and a vertical text label "up next Top 5 recommended videos" is placed to the right of the list.

Hussein, E., Juneja, P., & Mitra, T. (2020). Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1), 048:1-048:27.

Google Personalization



- pricing (Mikians et al, 2012):
 - browsing history
 - geo-location
- accounted for ~12% of result differences (Hannak et al, 2013):
 - being logged in Google
 - geo-location (Kliman-Silver et al, 2015)

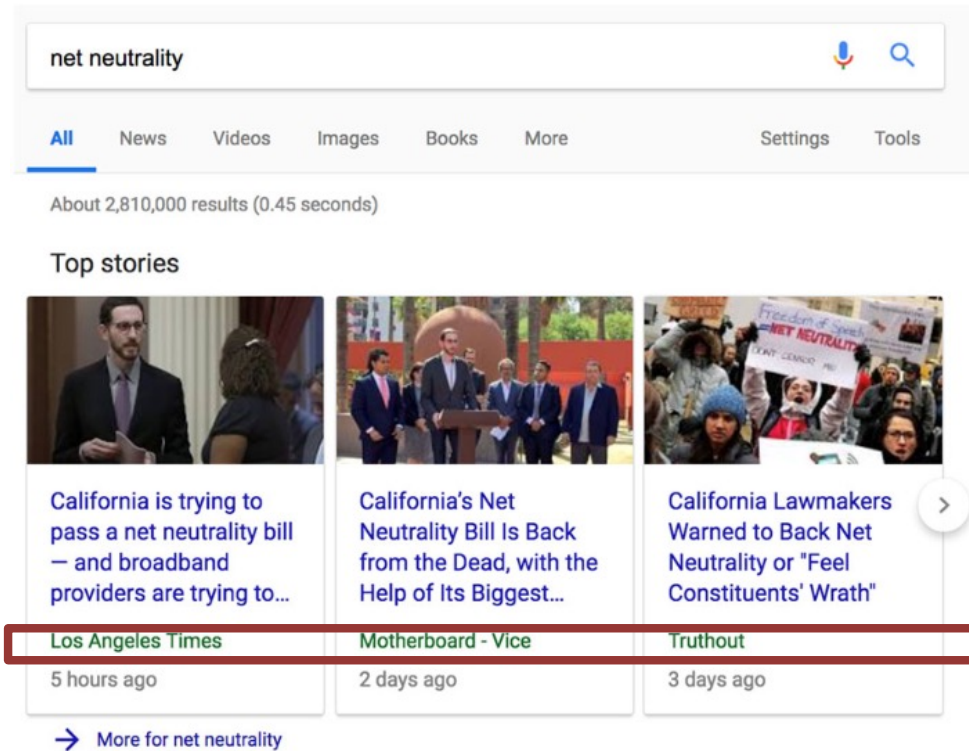
Mikians, J., Gyarmati, L., Erramilli, V., & Laoutaris, N. (2012). Detecting price and search discrimination on the internet. *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, 79–84.

Hannak, A., Sapiezynski, P., Molavi Kakhki, A., Krishnamurthy, B., Lazer, D., Mislove, A., & Wilson, C. (2013). Measuring personalization of web search. *Proceedings of the 22nd International Conference on World Wide Web*, 527–538.

Kliman-Silver, C., Hannak, A., Lazer, D., Wilson, C., & Mislove, A. (2015). Location, Location, Location: The Impact of Geolocation on Web Search Personalization. *Proceedings of the 2015 Internet Measurement Conference*, 121–127.

Google Source concentration

- top 20% of new sources account for 86% of all top stories (Trielli & Diakopoulos, 2019)






net neutrality

All News Videos Images Books More Settings Tools

About 2,810,000 results (0.45 seconds)

Top stories

 <p>California is trying to pass a net neutrality bill — and broadband providers are trying to...</p> <p>Los Angeles Times</p> <p>5 hours ago</p>	 <p>California's Net Neutrality Bill Is Back from the Dead, with the Help of Its Biggest...</p> <p>Motherboard - Vice</p> <p>2 days ago</p>	 <p>California Lawmakers Warned to Back Net Neutrality or "Feel Constituents' Wrath"</p> <p>Truthout</p> <p>3 days ago</p>
--	--	---

→ More for net neutrality

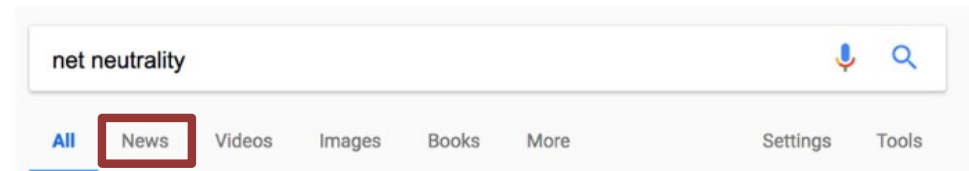
Trielli, D., & Diakopoulos, N. (2019). Search as News Curator: The Role of Google in Shaping Attention to News Information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15.

Fischer, S., Jaidka, K., & Lelkes, Y. (2020). Auditing local news presence on Google News. *Nature Human Behaviour*, 4(12), 1236–1244.

Urman, A., Makhortykh, M., & Ulloa, R. (2021). Auditing Source Diversity Bias in Video Search Results Using Virtual Agents. *Companion Proceedings of the Web Conference 2021*, 232–236.

Google Source concentration

- top 20% of new sources account for 86% of all top stories (Trielli & Diakopoulos, 2019)
- national outlets dominate news search results (Fischer et al, 2020)



The New York Times



The Washington Post

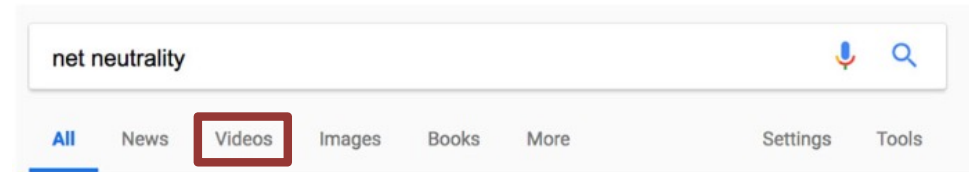
Trielli, D., & Diakopoulos, N. (2019). Search as News Curator: The Role of Google in Shaping Attention to News Information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15.

Fischer, S., Jaidka, K., & Lelkes, Y. (2020). Auditing local news presence on Google News. *Nature Human Behaviour*, 4(12), 1236–1244.

Urman, A., Makhortykh, M., & Ulloa, R. (2021). Auditing Source Diversity Bias in Video Search Results Using Virtual Agents. *Companion Proceedings of the Web Conference 2021*, 232–236.

Google Source concentration

- top 20% of new sources account for 86% of all top stories (Trielli & Diakopoulos, 2019)
- national outlets dominate news search results (Fischer et al, 2020)
- YouTube dominates video search results. In Google, it is followed by big media outlets (Urman, 2021)



The New York Times

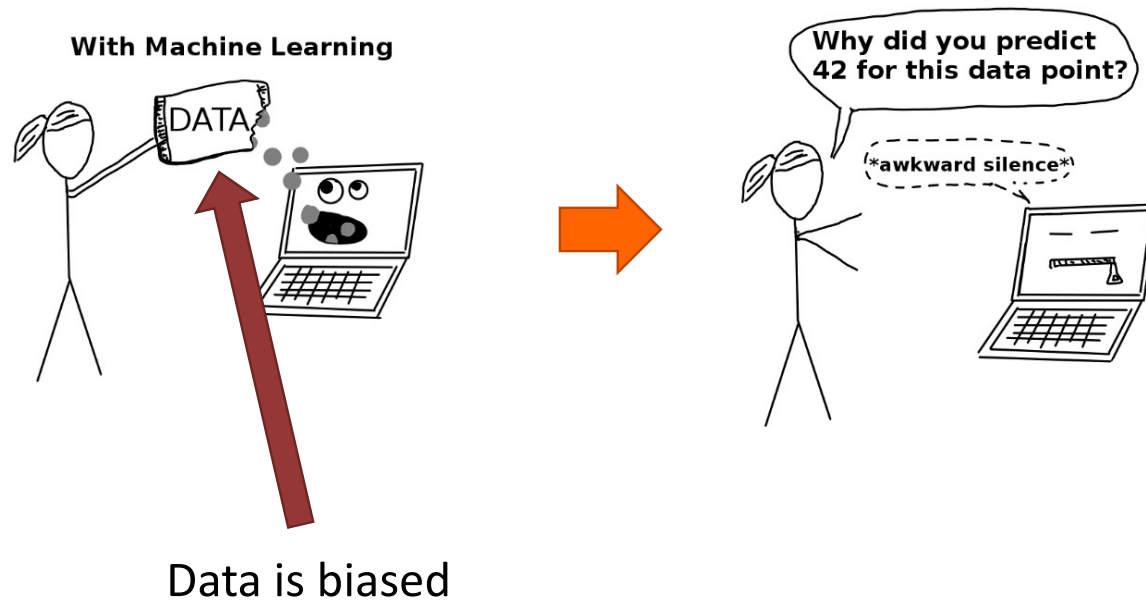
Trielli, D., & Diakopoulos, N. (2019). Search as News Curator: The Role of Google in Shaping Attention to News Information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15.

Fischer, S., Jaidka, K., & Lelkes, Y. (2020). Auditing local news presence on Google News. *Nature Human Behaviour*, 4(12), 1236–1244.

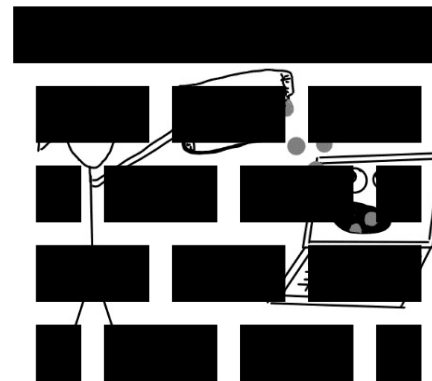
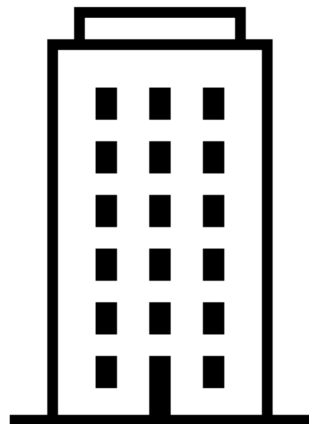
Urman, A., Makhortykh, M., & Ulloa, R. (2021). Auditing Source Diversity Bias in Video Search Results Using Virtual Agents. *Companion Proceedings of the Web Conference 2021*, 232–236.

Why does all of this occur?

Machine learning algorithms are opaque



Details are owned by a company





Twitter automatic cropping (September 2020)

- apologized (Hern, 2020)
- reported on the issue, and change the cropping mechanism (Chowdhury, 2021)
- open a competition to investigate their algorithm (Dang, 2021; Hern, 2021)

Hern, A. (2020, September 21). Twitter apologises for 'racist' image-cropping algorithm. *The Guardian*.

<https://www.theguardian.com/technology/2020/sep/21/twitter-apologises-for-racist-image-cropping-algorithm>

Chowdhury, Rumman (2021). Sharing learnings about our image cropping algorithm.

https://blog.twitter.com/engineering/en_us/topics/insights/2021/sharing-learnings-about-our-image-cropping-algorithm

Dang, S. (2021, July 30). Twitter launches competition to find biases in its image-cropping algorithm. *Reuters*.

<https://www.reuters.com/technology/twitter-launches-competition-find-biases-its-image-cropping-algorithm-2021-07-30/>

Hern, A. (2021, August 10). Student proves Twitter algorithm 'bias' toward lighter, slimmer, younger faces. *The Guardian*.

<https://www.theguardian.com/technology/2021/aug/10/twitters-image-cropping-algorithm-prefers-younger-slimmer-faces-with-lighter-skin-analysis>

 Facebook hinders research projects



AlgorithmWatch forced to shut down Instagram monitoring project after threats from Facebook. (2021.). *AlgorithmWatch*. Retrieved 3 October 2021, from <https://algorithmwatch.org/en/instagram-research-shut-down-by-facebook/>
Ad Observatory by NYU Tandon School of Engineering. (2021). Retrieved 4 October 2021, from <https://adobservatory.org/>
Facebook Rolls Out News Feed Change That Blocks Watchdogs from Gathering Data – The Markup. (2021). Retrieved 3 October 2021, from <https://themarkup.org/citizen-browser/2021/09/21/facebook-rolls-out-news-feed-change-that-blocks-watchdogs-from-gathering-data>

 Hinders research projects

July 13, 2021

ALGORITHM WATCH OPEN MENU ///

/ AlgorithmWatch forced to shut down Instagram monitoring project after threats from Facebook
 by Nicolas Kayser-Bril

Digital platforms play an ever-increasing role in structuring and influencing public debate. Civil society watchdogs, researchers and journalists need to be able to hold them to account. But Facebook is increasingly fighting those who try. It shut down New York University's Ad Observatory last week, and went after AlgorithmWatch, too. The European Parliament and EU Member States must act now to prevent further bullying.

STORY 13 AUGUST 2021 AUF DEUTSCH LESEN #CAMPAIGN #DATADONATION #FACEBOOK #INSTAGRAM



August 6, 2021

 **NYU Ad Observatory**
 By NYU Cybersecurity for Democracy

Explore Facebook political ads

Facebook has effectively stalled our Ad Observatory project by suspending Facebook accounts of Cybersecurity for Democracy team members. Lawmakers, regulators, and civil society groups are stepping up to support this project.

Please check [Cybersecurity for Democracy](https://cybersecurityfor democracy.org) for updates.

September 21, 2021

We need your input: Consider our survey

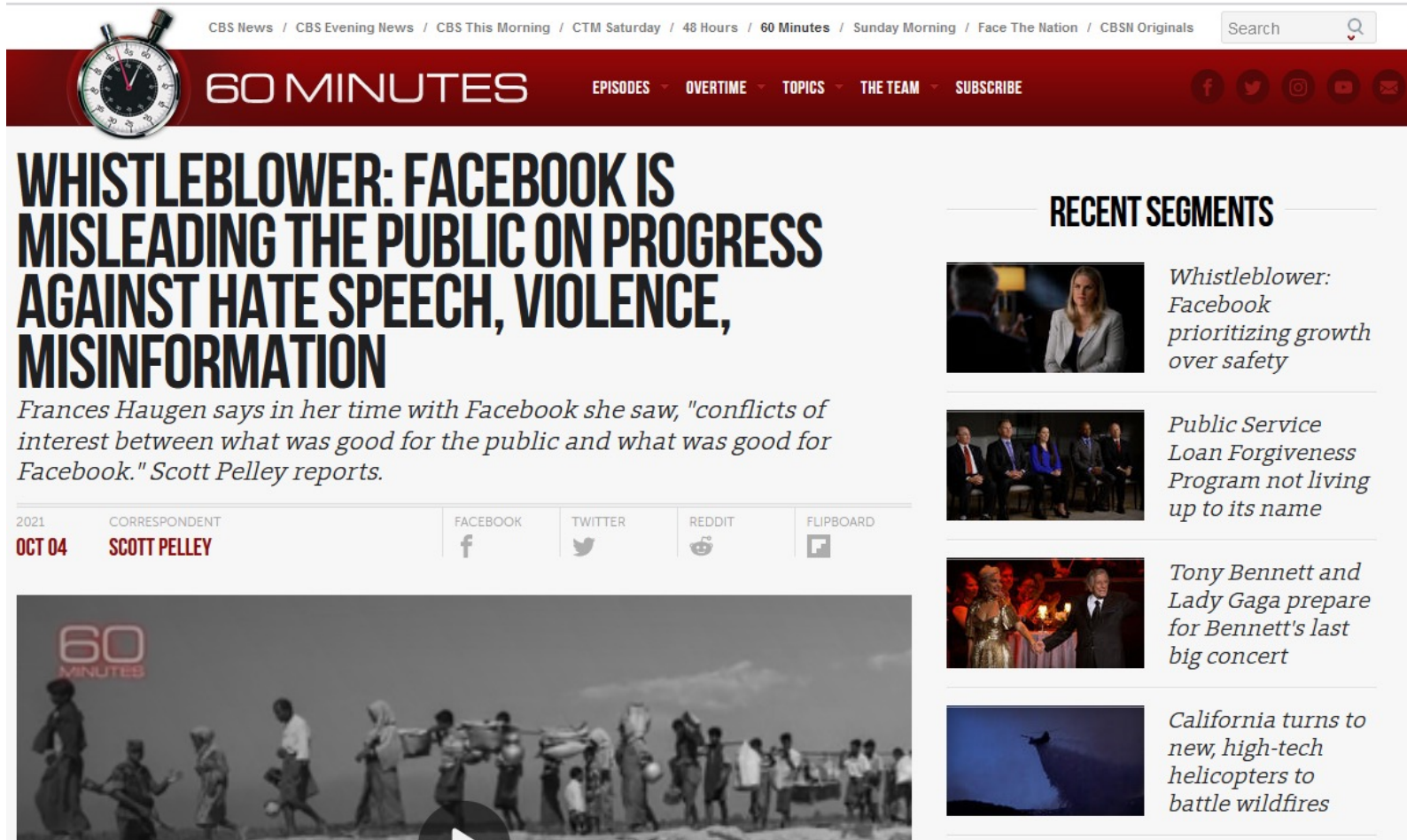
 **The Markup** Menu Donate

Citizen Browser
Facebook Rolls Out News Feed Change That Blocks Watchdogs from Gathering Data

The tweak, which targets the code in accessibility features for visually impaired users, drew ire from researchers and those who monitor the platform
 By [Corin Faife](#)
 September 21, 2021 08:00 ET
 Updated September 21, 2021 13:42 ET

AlgorithmWatch forced to shut down Instagram monitoring project after threats from Facebook. (2021.). *AlgorithmWatch*. Retrieved 3 October 2021, from <https://algorithmwatch.org/en/instagram-research-shut-down-by-facebook/>
 Ad Observatory by NYU Tandon School of Engineering. (2021). Retrieved 4 October 2021, from <https://adobservatory.org/>
 Facebook Rolls Out News Feed Change That Blocks Watchdogs from Gathering Data – *The Markup*. (2021). Retrieved 3 October 2021, from <https://themarkup.org/citizen-browser/2021/09/21/facebook-rolls-out-news-feed-change-that-blocks-watchdogs-from-gathering-data>

 Conflicts of interest (Pelley, October 4th 2021)



The screenshot shows the CBS 60 Minutes website. At the top, there is a navigation bar with links to various CBS programs and a search box. Below this is a red header with the '60 MINUTES' logo and navigation options like 'EPISODES', 'OVERTIME', 'TOPICS', 'THE TEAM', and 'SUBSCRIBE'. The main article title is 'WHISTLEBLOWER: FACEBOOK IS MISLEADING THE PUBLIC ON PROGRESS AGAINST HATE SPEECH, VIOLENCE, MISINFORMATION'. A sub-headline reads: 'Frances Haugen says in her time with Facebook she saw, "conflicts of interest between what was good for the public and what was good for Facebook." Scott Pelley reports.' Below the text are social media sharing icons for Facebook, Twitter, Reddit, and Flipboard. To the right, there is a 'RECENT SEGMENTS' section with four video thumbnails and their corresponding titles: 'Whistleblower: Facebook prioritizing growth over safety', 'Public Service Loan Forgiveness Program not living up to its name', 'Tony Bennett and Lady Gaga prepare for Bennett's last big concert', and 'California turns to new, high-tech helicopters to battle wildfires'. At the bottom of the article, there is a large image showing a group of people carrying supplies on their heads, with the '60 MINUTES' logo overlaid in the top left corner.

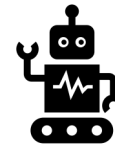
Algorithmic auditing

“process of investigating the **functionality and impact** of decision-making algorithms”(Mittelstadt, 2016)

How to perform algorithm auditing?

Bandy (2021)

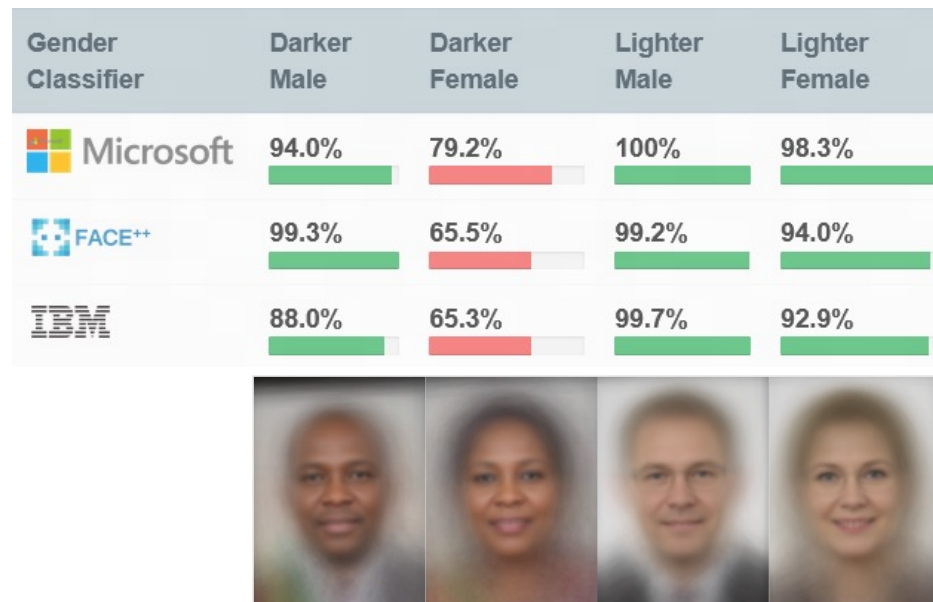
- source code
- **direct scrape using APIs**
- **sock puppets**
- **carrier puppets**
- **crowdsourcing**



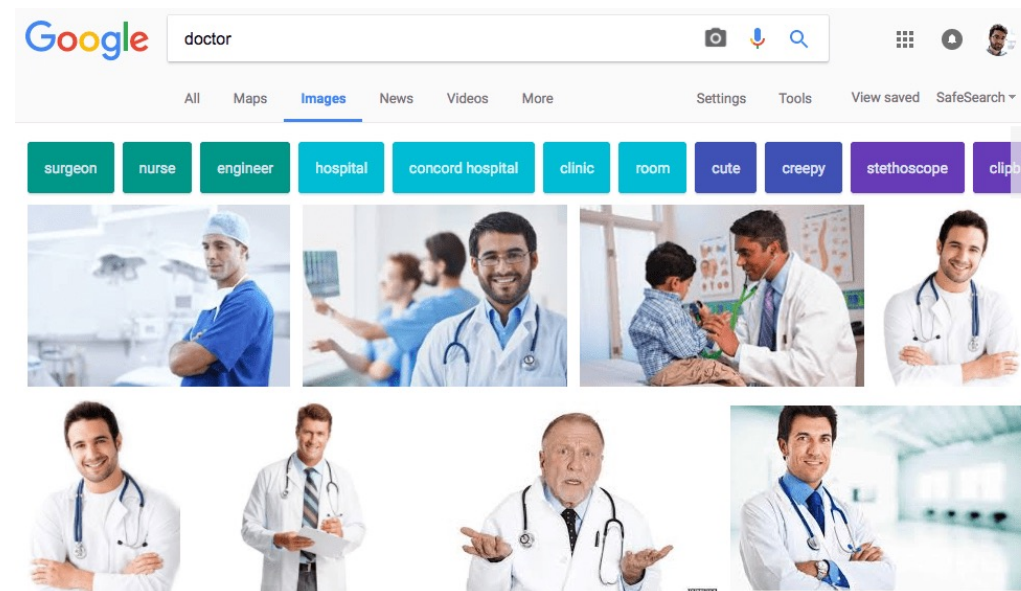
What has been investigated? (Bandy, 2021)



What has been investigated? (Bandy, 2021)

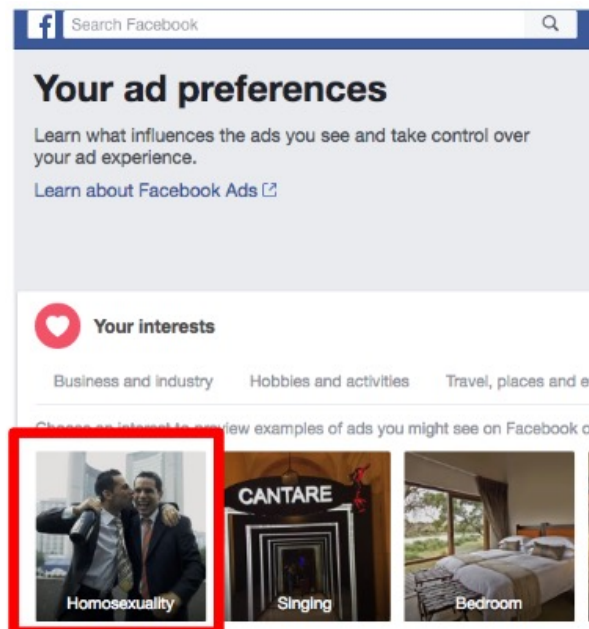


What has been investigated? (Bandy, 2021)



Kay, M., Matuszek, C., & Munson, S. A. (2015). Unequal Representation and Gender Stereotypes in Image Search Results for Occupations. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3819–3828. <https://doi.org/10.1145/2702123.2702520>

What has been investigated? (Bandy, 2021)



What has been investigated? (Bandy, 2021)



Silva et al., (2020)



NYU Ad Observatory (Wazman, 2021)
By NYU Cybersecurity for Democracy



Silva, M., Santos de Oliveira, L., Andreou, A., Vaz de Melo, P. O., Goga, O., & Benevenuto, F. (2020). Facebook Ads Monitor: An Independent Auditing System for Political Ads on Facebook. *Proceedings of The Web Conference 2020*, 224–234.

Watzman, N. (2021, May 12). The political ads Facebook won't show you. *Cybersecurity for Democracy*.

<https://medium.com/cybersecurity-for-democracy/the-political-ads-facebook-wont-show-you-e0d6181bca25>

What has been investigated? (Bandy, 2021)



What has been investigated? (Bandy, 2021)

Platform	Discrimination	Distortion	Exploitation	Misjudgment
Search				
Advertising				
Recommendation				
Pricing				
Vision				

Bandy, J. (2021). Problematic Machine Behavior: A Systematic Literature Review of Algorithm Audits. *ArXiv:2102.04256 [Cs]*.

What has been investigated? (Bandy, 2021)

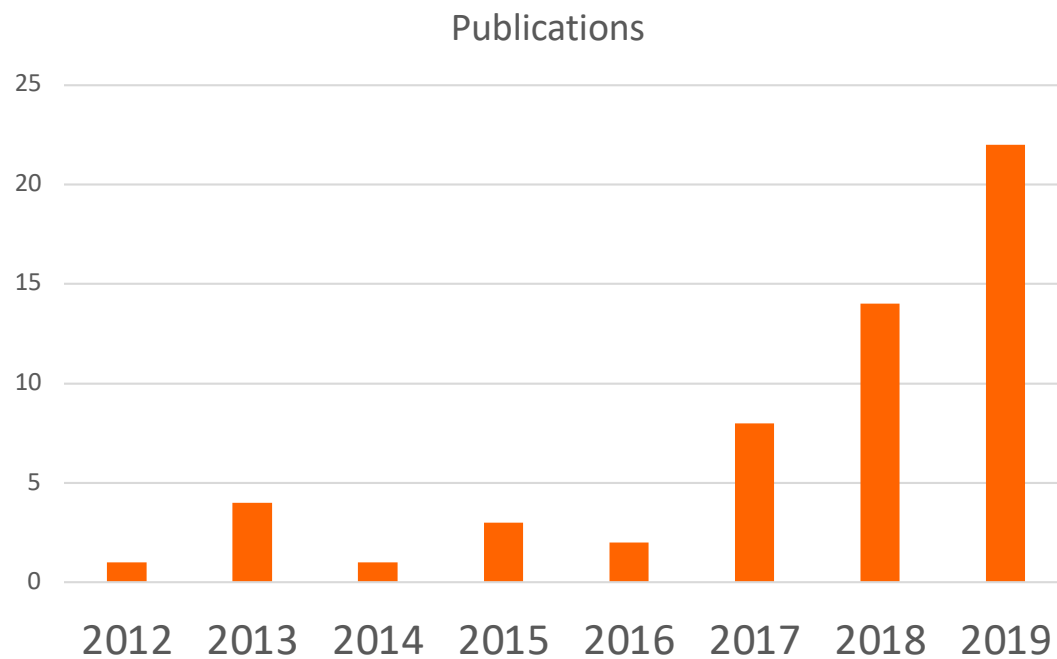
Platform	Discrimination	Distortion	Exploitation	Misjudgment
Search				
Advertising				
Recommendation				
Pricing				
Vision				
Criminal Justice				
Language Processing				
Mapping				

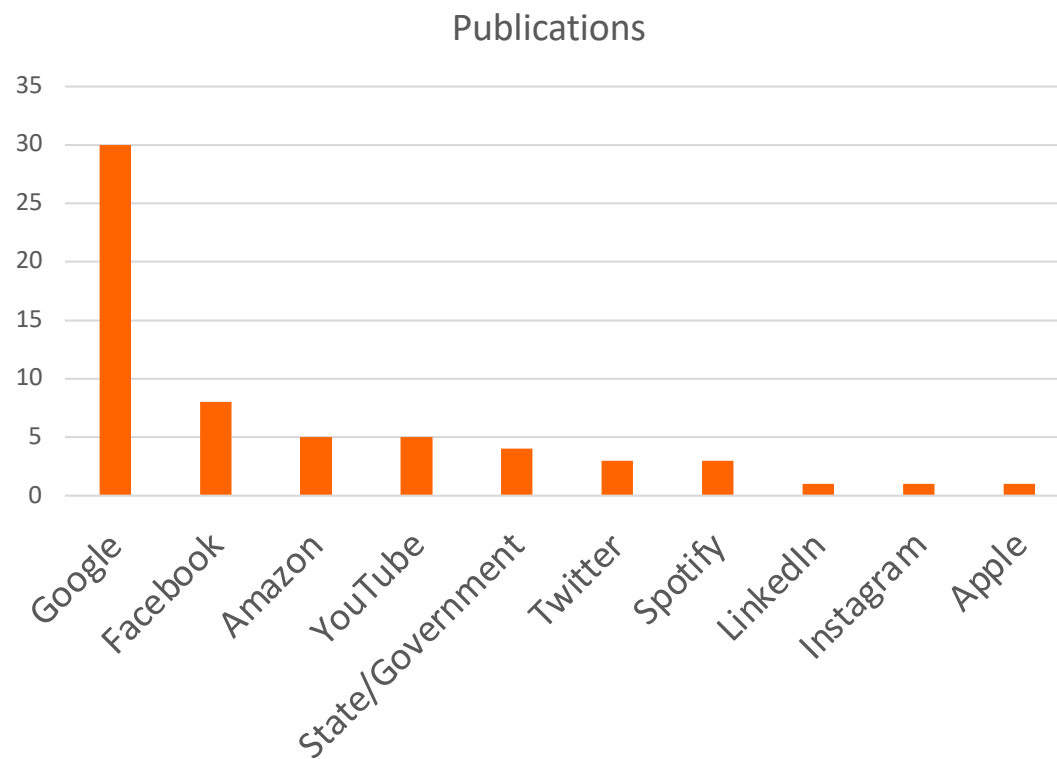
Bandy, J. (2021). Problematic Machine Behavior: A Systematic Literature Review of Algorithm Audits. *ArXiv:2102.04256 [Cs]*.

What has been investigated? (Bandy, 2021)

Platforms	Discrimination	Distortion	Exploitation	Misjudgment	Total
Search	5	18	2		25
Advertising	3	2	3	4	12
Recommendation	1	7			8
Pricing	5				5
Vision	5				5
Criminal Justice	1			3	4
Language Processing	1	1			2
Mapping		1			1
Total	21	29	5	7	62

Bandy, J. (2021). Problematic Machine Behavior: A Systematic Literature Review of Algorithm Audits. *ArXiv:2102.04256 [Cs]*.





Fill the gap with more search engines



- Makhortykh, M., Urman, A., & Ulloa, R. (2020). How search engines disseminate information about COVID-19 and why they should do better. *Harvard Kennedy School Misinformation Review*, 1(COVID-19 and Misinformation).
- Makhortykh, M., Urman, A., & Ulloa, R. (2021). Detecting Race and Gender Bias in Visual Representation of AI on Web Search Engines. In L. Boratto, S. Faralli, M. Marras, & G. Stilo (Eds.), *Advances in Bias and Fairness in Information Retrieval* (pp. 36–50). Springer International Publishing.
- Urman, A., Makhortykh, M., & Ulloa, R. (2021). Auditing Source Diversity Bias in Video Search Results Using Virtual Agents. *Companion Proceedings of the Web Conference 2021*, 232–236.
- Ulloa, R., Makhortykh, M., & Urman, A. (2021). Algorithm Auditing at a Large-Scale: Insights from Search Engine Audits. *ArXiv:2106.05831 [Cs]*.
- Urman, A., Makhortykh, M., & Ulloa, R. (2021). The Matter of Chance: Auditing Web Search Results Related to the 2020 U.S. Presidential Primary Elections Across Six Search Engines. *Social Science Computer Review*, 08944393211006863.
- Makhortykh, M., Urman, A., & Ulloa, R. (2021). Hey, Google, is this what the Holocaust looked like? Auditing algorithmic curation of visual historical content on Web search engines. *First Monday*.

Experimental design (case of the search engines)

- time of query (Hannak et al, 2013)
- browsing history (Mikians, 2012)
- location (Hannak et al, 2013; Kliman-Silver et al., 2015)
- ranking (Kulshrestha, 2017)
- browser type (Urman et al., 2021)
- language (Makhortykh et al, 2020)
- randomization (Makhortykh et al, 2020; Urman et al., 2021)
- cookie consent and captchas (Ulloa et al, 2021)

Hannak, A., Sapiezynski, P., Molavi Kakhki, A., Krishnamurthy, B., Lazer, D., Mislove, A., & Wilson, C. (2013). Measuring personalization of web search. *Proceedings of the 22nd International Conference on World Wide Web*, 527–538.

Kliman-Silver, C., Hannak, A., Lazer, D., Wilson, C., & Mislove, A. (2015). Location, Location, Location: The Impact of Geolocation on Web Search Personalization. *Proceedings of the 2015 Internet Measurement Conference*, 121–127.

Kulshrestha, J., Eslami, M., Messias, J., Zafar, M. B., Ghosh, S., Gummadi, K. P., & Karahalios, K. (2017). Quantifying Search Bias: Investigating Sources of Bias for Political Searches in Social Media. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 417–432.

Makhortykh, M., Urman, A., & Ulloa, R. (2020). How search engines disseminate information about COVID-19 and why they should do better. *Harvard Kennedy School Misinformation Review*, 1(COVID-19 and Misinformation).

Mikians, J., Gyarmati, L., Erramilli, V., & Laoutaris, N. (2012). Detecting price and search discrimination on the internet. *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, 79–84.

Ulloa, R., Makhortykh, M., & Urman, A. (2021). Algorithm Auditing at a Large-Scale: Insights from Search Engine Audits. *ArXiv:2106.05831 [Cs]*.

Urman, A., Makhortykh, M., & Ulloa, R. (2021). The Matter of Chance: Auditing Web Search Results Related to the 2020 U.S. Presidential Primary Elections Across Six Search Engines. *Social Science Computer Review*, 08944393211006863.

General considerations

- If you have participants data:
 - ethics approval
 - informed consent of the participant
- Auditing might break the Terms & Conditions of platforms:
 - false accounts
 - automated data collection
- Varies according to platform
 - search engines generally open
 - social media platforms not so
- Keep in mind data rights specially in the case of images, not everything that is posted is legal

Takeaways & final remarks

- Algorithm auditing is a set of methods to expose problems in platform technologies
- Opportunity to steer the future by avoiding the perpetuation of existent inequalities
- Conflict of interest between companies and researchers
- Ethical considerations, but not many alternatives
- Some regulations are being put into place:
 - General Data Protection Regulation (GDPR)
 - European Artificial Intelligence Act (proposal Apr 2021)
 - Algorithmic Accountability Act (US, rejected 2019, resubmitted 2021)
 - Algorithmic Justice and Online Platform Transparency Act of 2021 (US, submitted)

Thank you !

gesis

Leibniz-Institut
für Sozialwissenschaften

Leibniz
Leibniz
Gemeinschaft

Expert Contact & GESIS Consulting




Contact: you can reach the speaker/s via e-mail:

roberto.ulloa@gesis.org

GESIS Consulting: GESIS offers individual consulting in a number of areas – including survey design & methodology, data archiving, digital behavioral data & computational social science – and across the research data cycle.

Please visit our website www.gesis.org for more [detailed information](#) on available services and terms.

More Services from GESIS

- Get materials for [capacity building in computational social science](#) and take advantage of our expanding expertise and resources in [digital behavioral data](#).
- Use GESIS data services for [finding data](#) for secondary analysis and [sharing your own data](#).
- Check out the [GESIS blog](#) "Growing Knowledge in the Social Sciences" for topics, methods and discussions from the GESIS cosmos – and beyond.
- Keep up with GESIS activities and subscribe to the monthly [newsletter](#).
-  for publications, tools & services.

*** Upcoming online workshop; Nov 2-5, 2021 ***

[Introduction to Social Media as Research Data: Potentials and Pitfalls](#)

More from CSS Experts in the Series

- June 24 Katrin Weller: **A Short Introduction to Computational Social Science and Digital Behavioral Data**
- July 01 Fabian Flöck, Indira Sen: **Digital Traces of Human Behavior from Online Platforms – Research Designs and Error Sources**
- July 08 Sebastian Stier, Johannes Breuer: **Combining Survey Data and Digital Behavioral Data**
- Sept 16 Oliver Watteler, Katrin Weller: **Research Ethics and Data Protection in Social Media Research**
- Sept 30 Roberto Ulloa: **Introduction to Online Data Acquisition**
- Oct 07 Roberto Ulloa: **Auditing Algorithms: How Platform Technologies Shape our Digital Environment**
- Oct 14 Marius Sältzer, Sebastian Stier: **The German Federal Election: Social Media Data for Scientific (Re-)Use**
- Nov 04 Arnim Bleier: **Introduction to Text Mining**
- Nov 11 Haiko Lietz: **Social Network Analysis with Digital Behavioral Data**
- Dec 2 Olga Zagovora, Katrin Weller: **Altmetrics: Analyzing Academic Communications from Social Media Data**
- Dec 16 Andreas Schmitz: **Online Dating: Data Types and Analytical Approaches**
- Jan 13 Gizem Bacaksizlar: **Political Behavior and Influence in Online Networks**
- Jan 27 David Brodesser: **SocioHub – A Collaboration Platform for the Social Sciences**