Two cases of algorithmic biases: Youtube’s polarization and Pornhub’s heteronormativity.
Contents

A tool for passive scraping

tracking.exposed

How to measure the filter bubble on Youtube

FilterTube

Heteronormativity and algorithms on PornHub

PorNorm

Discussion on future developments

What’s next?
Let’s talk about algorithms

Algorithms are the technological solution to the information overload.

Unfortunately, they can also end up reinforcing the **preexisting beliefs** of the person who programmed them.

Those intrinsically embedded assertions, called *bias*, do have an impact on the dissemination of ideas and culture.
For example, the spreading of online news (and other contents like porn) consumption is increasing year after year; **Algorithmic mediation** is a crucial issue for our democracy.
Roughly half of Americans prefer to get news on a digital platform; about a third prefer TV

52% of US adults consume news through Digital Platform,
23% of U.S. adults regularly getting news on Youtube.

70% of the watch time is from YouTube's suggested videos.
(Youtube, 2018)

Note: Digital platforms includes respondents who said they prefer news websites or apps, social media, search, or podcasts.

52% of US adults consume news through Digital Platform,
23% of U.S. adults regularly getting news on Youtube.

70% of the watch time is from YouTube's suggested videos.
(Youtube, 2018)

Pew Research Center survey
(Shearer and Mitchell, 2020)
"Deep Neural Networks for YouTube Recommendations" (Covington et al., 2016)

"The Black Box Society: The Secret Algorithms That Control Money and Information" (Pasquale, 2015)
We are a no-profit, free software project aimed to analyze evidence of algorithm personalization.

Sponsored and collaborating with (past|present):
Our browser extensions:

- Facebook Tracking Exposed
- Pornhub Tracking Exposed
- Youtube Tracking Exposed
- amazon Tracking Exposed
How contribution and collection works

Browser extension parse the HTML (JavaScript)

Recommendations/evidences

Software released under AGPL-3
Free license.

https://github.com/tracking-exposed
From the extension pop-up you access to a personal page via an authentication token:

https://youtube.tracking.exposed/personal/#59M4Qyevek5dGH4GJyqymK5T81XrdRSdDSks1bZUquUsE6pQ
Users and researchers can compare the personalization seen on a specific video with auto-generated graphs and datafiles.
Why you can’t trust API
Controlled experiments
(socks puppets audit)

Researchers design an experiment, they keep under control as many variables as possible, sometimes they use bots.

Collective observations
(crowdsourced audit)

People of the web perform the same actions on the platform, recording evidence about the algorithm with TREX

“Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms” (Sandvig et al., 2014)
Links to previous researches:

**Facebook:**
- Italian elections
- Dutch elections, a new crowdsourced approach.

**Youtube:**
- Personalization on Youtube
- YouTube Tracking Exposed: Investigating Brexit polarization
- Covid and misinformation in non English language (pre-print).
- FilterTube: Investigating echo chambers, filter bubbles and polarization on YouTube

**Pornhub:**
- Collective observation / a bit of algorithm reversing.
- Heteronormativity and Pornography: an Algorithmic Analysis of PornHub (coming soon)

**Amazon:**
- Price discrimination research(es).
Echo Chambers, Filter Bubble and Polarization on Youtube: the USA’s post-electoral debate

Theoretical framework

*Echo Chamber* → The political ideology of the subject and the consequent watching choices. Based on ideological affinity, are created by users (Dubois ’18).

*Filter bubble* → Direct effect of algorithmic personalization, based on the users’ behavior (Eli Pariser ’11).

*Polarization* → Process of increased segregation into distinct social groups, separated along racial, economic, political, religious or other lines (Gallacher & Heerdink ‘19).
Methodology

New browser with YouTube Tracking Exposed Extension

Selection of channels to watch

- 9 Progressive browsers
  - Progressive Channel 1
    - video 1
    - video 2
  - Progressive Channel 2
    - video 1
    - video 2
  - Progressive Channel 3
    - video 1
    - video 2

- 9 Conservative browsers
  - Conservative Channel 1
    - video 1
    - video 2
  - Conservative Channel 2
    - video 1
    - video 2
  - Conservative Channel 3
    - video 1
    - video 2

Query search on YouTube

- American elections
  - Search results
  - Search results' comments download
  - Natural Language Processing
  - Tag Cloud Sentiment analysis
- Corona virus
  - Search results
  - Search results' comments download
  - Downloaded with YouTube Tracking Exposed Extension
  - Coded
  - Statistics
  - Network analysis
Search results according to political orientation and media type
Heteronormativity and Pornography: an Algorithmic Analysis of PornHub
Popular pornographic websites are deemed to reinforce a male, white and heterosexual point of view, and thus contribute to foster hegemonic masculinity, the sexualization of minorities, and a heteronormative porn culture.

Algorithms contribute to manage content visibility and can reiterate the gender bias coded into them, reifying a specific view of the world due to their social embeddedness.

(Bucher 2012. Noble 2018)
Methodology

10 different profiles

20 homepages collected per day

1 week
Account sign-up choice:

How do you like it?
 Tell us about yourself and what turns you on so we can provide you with the content that suits your tastes.

I am a:

- None
- Male
- Female
- Couple
- Same Sex Couple (Female)
- Same Sex Couple (Male)
- Trans Woman
- Trans Man
- Other
- Non-Binary

I like:

- Girls
- Guys
- Guys and Girls
Homepage collection

- Pornhub's Homepage
- Sections
  - Hot Porn Videos in [Your Country]
  - Most Viewed Videos in [Your Country]
  - Recommended for You
  - Recommended Category for You - [Category]
  - Recently Featured XXX Videos
Videos suggested in shared sections:

- Hot Porn Videos in Your Country
- Most Viewed Videos in Your Country
- Recently Featured XXX Videos

All the users’ profiles received the same videos.
Videos in **Recommended for You** based on users’ profiles.

Two heteronormated clusters emerge clearly. The gender *None* was the more divergent in personalization.
Videos in Recommended Category for You based on users' profiles.

Not all the profiles had the Recommended Category for you section. Some clusters are still visible.
The lack of **awareness and empowerment** in internet users has a massive impact on their ability to truly benefit from digital innovation and often creates the **foundation for their exploitation**.
Those bias is not a technological problem, therefore the solution cannot be only technological. We have to deconstruct social stereotypes and their algorithmic expressions.
Tracking Exposed is just one of the **many steps** we need to take in order to create a way out of Surveillance Capitalism.
Thank you!

info@tracking.exposed