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


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%)

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):
  - 1/3 women for “person”
  - 1/3 women for “intelligent person”

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

---


**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)
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)


Automatic labelling (Cabañas et al., 2018)

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)

Intransparent labelling (Datta et al., 2018)

Group A
visited substance abuse websites

RQ1: substance abuse Ads?

yes

Group B
did not visit any website

no

RQ2: differences in the Google Ads preferences?

no

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

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

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)

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


Why does all of this occur?
Machine learning algorithms are opaque

Data is biased

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)

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


27


Hinders research projects


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.

Conflicts of interest (Pelley, October 4th 2021)

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)

Discrimination  Distortion  Exploitation  Misjudgment

What has been investigated? (Bandy, 2021)

Discrimination  Distortion  Exploitation  Misjudgment

What has been investigated? (Bandy, 2021)

Silva et al., (2020)


What has been investigated? (Bandy, 2021)

<table>
<thead>
<tr>
<th>Discrimination</th>
<th>Distortion</th>
<th>Exploitation</th>
<th>Misjudgment</th>
</tr>
</thead>
</table>

What has been investigated? (Bandy, 2021)

<table>
<thead>
<tr>
<th>Platform</th>
<th>Discrimination</th>
<th>Distortion</th>
<th>Exploitation</th>
<th>Misjudgment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advertising</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pricing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vision</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What has been investigated? (Bandy, 2021)

<table>
<thead>
<tr>
<th>Platform</th>
<th>Discrimination</th>
<th>Distortion</th>
<th>Exploitation</th>
<th>Misjudgment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advertising</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pricing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vision</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criminal Justice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language Processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mapping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What has been investigated? (Bandy, 2021)

<table>
<thead>
<tr>
<th>Platforms</th>
<th>Discrimination</th>
<th>Distortion</th>
<th>Exploitation</th>
<th>Misjudgment</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search</td>
<td>5</td>
<td>18</td>
<td>2</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>Advertising</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Recommendation</td>
<td>1</td>
<td>7</td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Pricing</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Vision</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Criminal Justice</td>
<td>1</td>
<td></td>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Language Processing</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Mapping</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>21</strong></td>
<td><strong>29</strong></td>
<td><strong>5</strong></td>
<td><strong>7</strong></td>
<td><strong>62</strong></td>
</tr>
</tbody>
</table>

Fill the gap with more search engines

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)
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!
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**.
- Search GESIS 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 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 02  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