



Auditing algorithms

Shuning Lu (Ph.D.)
Assistant Professor
Department of Communication
North Dakota State University

Email: shuning.lu@ndsu.edu
Twitter: [@shuning_lu](https://twitter.com/shuning_lu)



Today's agenda

Morning:

- Logistics
- Overview of algorithms auditing
- Algorithms auditing tutorial

Lunch break

Afternoon:

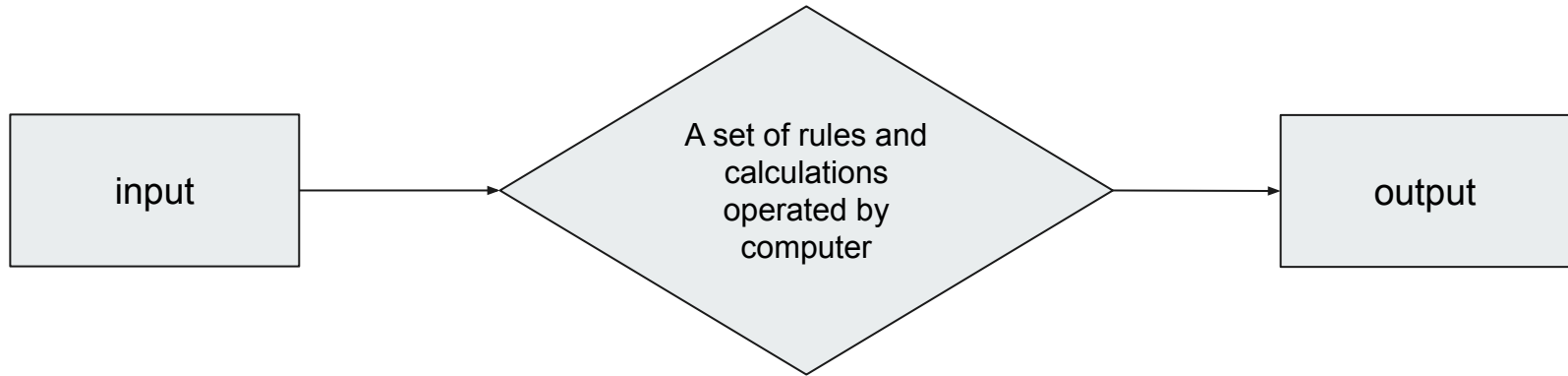
- Group activity & report



What is algorithm?



What is algorithm?





Meta (Facebook) algorithm

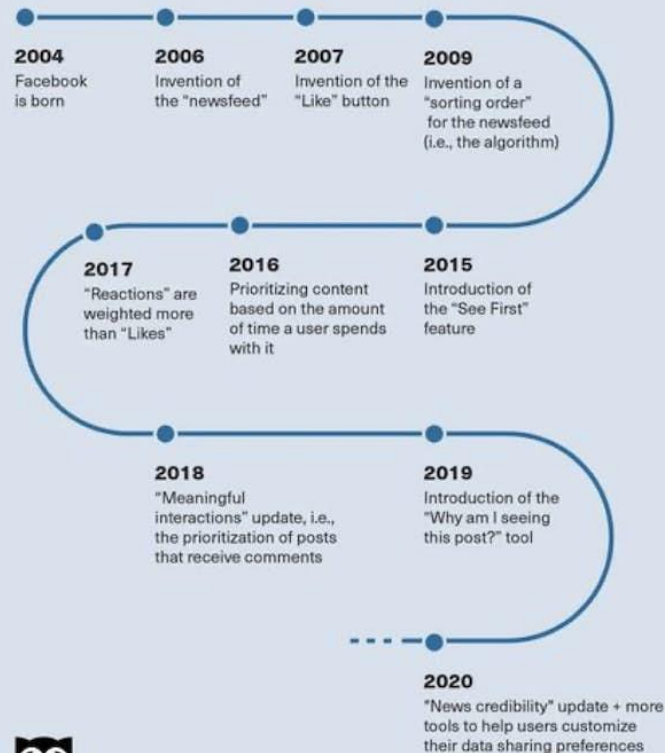
User

+

curation

Algorithmic

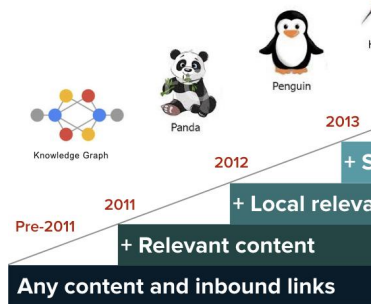
Key Moments in the History of the Facebook Algorithm



Google algorithm

Keeping Pace

(Rare)

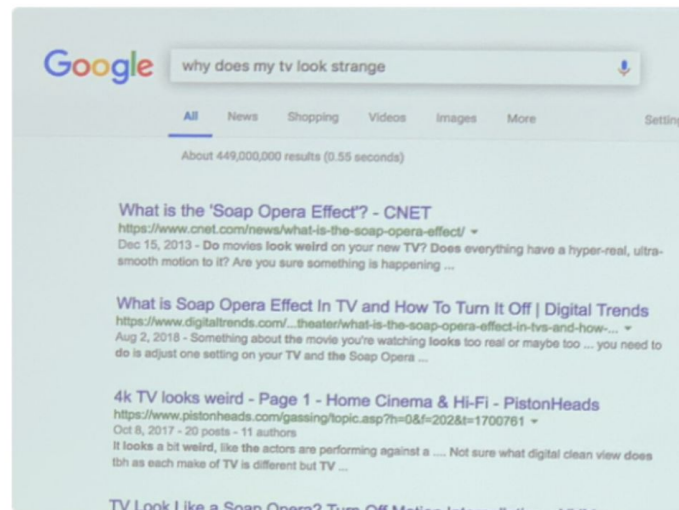


Danny Sullivan ✓

@dannysullivan

Follow

Last few months, Google has been using neural matching, --AI method to better connect words to concepts. Super synonyms, in a way, and impacting 30% of queries. Don't know what "soap opera effect" is to search for it? We can better figure it out.



10:24 AM - 24 Sep 2018



Why auditing algorithms?

Public interest scrutiny

Hold (bad) algorithms accountable



Why auditing algorithms?

make this technology more explainable, transparent, predictable and controllable

- by citizens, public institutions and also companies,
- either before the development of the system, during its development or a posteriori.



Goals

The purpose of an audit is to

- Identify or anticipate errors, risks or threats (actual or potential).
- Outline a strategy for improving the algorithmic processes.



What is algorithm auditing?

Broadly defined: a range of approaches to review algorithmic processing systems.

Narrowly defined: a method of repeatedly querying an algorithm and observing its output in order to draw conclusions about the algorithm's opaque inner workings and possible external impact



Algorithm auditing as field experiment

Field experiment

Treatment

Outcome

Algorithm auditing

Query

Results

In natural, real-world settings.



Legal and ethical compliance

A platform's Term of Services

Computer Fraud and Abuse Act

Researcher's own professional organizations' stances (e.g., IRB)



Common domains of algorithm auditing

Employment (e.g., ads, ranking of candidates)

Media consumption (e.g., news recommendation & ranking, music streaming sites)

Sharing economy (e.g., Airbnb rental platform, Uber matching)

Healthcare (e.g., information seeking, diagnose and care for patients)

Online market price and product pricing

...

Auditing search engines: The case of Google

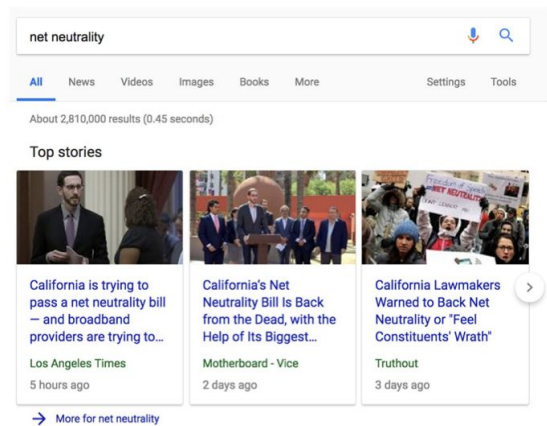


Figure 1: The Google Top Stories Component shown for a query of “net neutrality”.

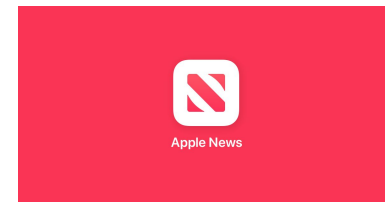


What they found:

- A high concentration of certain news sources
- a slight exaggeration in the ideological skew

Trielli, D., & Diakopoulos, N. (2019, May). Search as news curator: The role of Google in shaping attention to news information. In *Proceedings of the 2019 CHI Conference on human factors in computing systems* (pp. 1-15).

Auditing news recommendation systems: The case of Apple News



Trending story section on the app

Crowdsourced audit: minimal content personalization

Sock-puppet audit: no location-based personalization

Comparison between human-curated Top Stories section and the algorithmically-curated Trending Stories section.

- human performed algorithm in source diversity and evenness.
- Algorithm featured more “soft news”, while editorial curation featured more hard news

Bandy, J., & Diakopoulos, N. (2020, May). Auditing news curation systems: A case study examining algorithmic and editorial logic in Apple News. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 14, pp. 36-47).

Auditing news recommendation systems: The case of YouTube



Keyword based search Pro-vaccine videos from credible sources

Pro-vaccine video → related videos

Anti-vaccine video → related videos

Expert seed videos Anti-vaccine videos (2~12 times)

Pro-vaccine video → related videos

Anti-vaccine video → related videos

Tang, L., Fujimoto, K., Amith, M., Cunningham, R., Costantini, R. A., York, F., ... & Tao, C. (2021). "Down the rabbit hole" of vaccine misinformation on YouTube: Network exposure study. *Journal of Medical Internet Research*, 23(1), e23262.

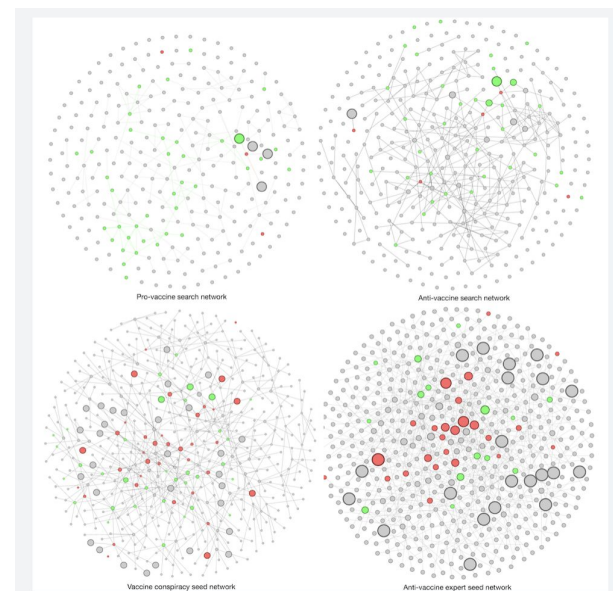


Figure 2. Visualization of the four YouTube video networks, where each node represents a video and an edge represents a related link. Green represents a provaccine video and red represents an antivaccine video. Gray nodes are nonvaccine-related videos. The size of the nodes depicts the exposure value to antivaccination videos. Visualization was generated with Gephi.

Auditing chatbots: the case of ChatGPT

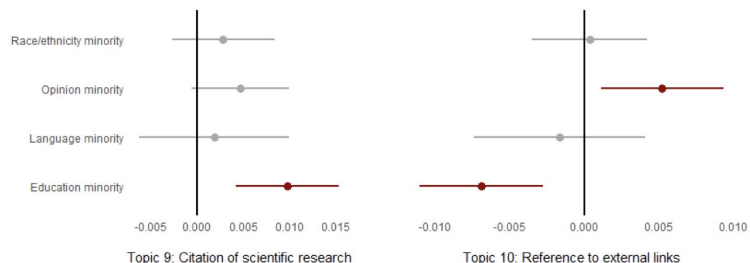


evaluating equity in human-AI dialogues

examine how GPT-3 responded to different subpopulations on climate change and the Black Lives Matter (BLM) movement.

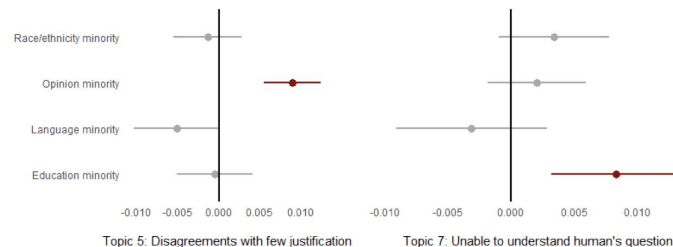
Panel A: Climate Change

(1). STM Analyses



Panel B: BLM

(1). STM Analyses



Chen, K., Shao, A., Burapachee, J., & Li, Y. (2022). A critical appraisal of equity in conversational AI: Evidence from auditing GPT-3's dialogues with different publics on climate change and Black Lives Matter. *arXiv preprint arXiv:2209.13627*.



How to conduct algorithm auditing?

Selecting a research topic

Selecting an algorithmic/platform to audit

Data collection

Data analysis

Communicating findings

Metaxa, D., Park, J. S., Robertson, R. E., Karahalios, K., Wilson, C., Hancock, J., & Sandvig, C. (2021). Auditing algorithms: Understanding algorithmic systems from the outside in. *Foundations and Trends® in Human–Computer Interaction*, 14(4), 272-344.



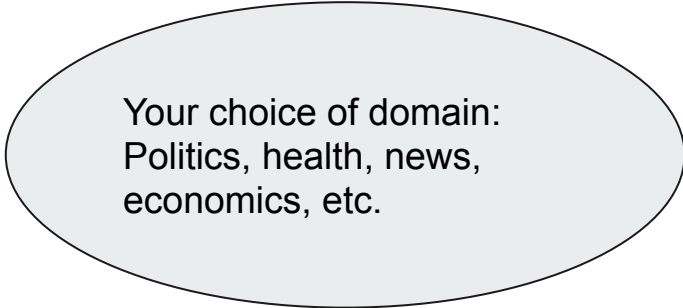
Selecting a research topic

What do algorithms represent in general?

- the baseline/real world pattern
- the normative standard (diversity)
- other ranking standard (time, location, prominence...)

How do algorithms treat different users?

- output across users
- user experience



Your choice of domain:
Politics, health, news,
economics, etc.



Selecting an algorithm/platform to audit

Search engines

Social media news feed

Conversational AI

Image retouch app

...

Reviewing the field



Research design

Single platform/algorithm

Cross-platform comparison

International differences

Temporal considerations

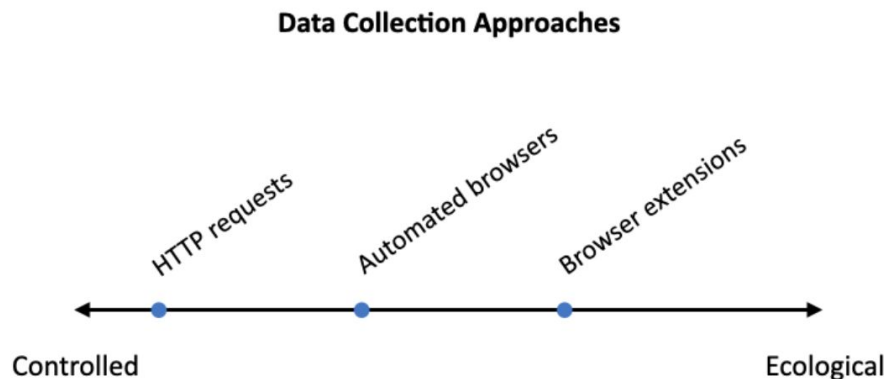
Data collection

Human-based approach:

- researchers
- volunteers, crowdworkers

Automated approach:

- Automated browsers
- Socket puppet
- Web extension





Data analysis

Quantitative content analysis

Qualitative thematic analysis

Computational text analysis

Statistical modeling



Group activity Day 2

Summary

An open-ended group exercise to propose and pilot an algorithm auditing study. You need to select a topic and platform for auditing, create a plan, report initial findings, and suggest a direction for future research.

Activity

- Split into small groups and select person(s) to take notes and report group process/results.
- 13:40-13:50: brainstorm potential research ideas and select one to pursue
- 13:50-14:00: discuss data collection strategy (collect by yourself as pilot or use/combine advanced methods)
- 14:00-15:20: collect and analyze initial data to answer 1-2 key research questions
- 15:20-15:30: reflect on the strengths/limitations of what you have completed and ways to address
- 15:30-16:00: come back together as a large group and discuss projects at the end of the day