Auditing Algorithms

Understanding Algorithmic Systems from the Outside In

Other titles in Foundations and Trends® in Human-Computer Interaction

Design Framework for Interactive Highlighting Techniques
Joshua Leung and Andy Cockburn
ISBN: 978-1-68083-924-1

Post-WIMP Interaction for Information Visualization
Bongshin Lee, Arjun Srinivasan, Petra Isenberg and John Stasko
ISBN: 978-1-68083-826-8

Improving HCI with Brain Input: Review, Trends, and Outlook
Erin T. Solovey and Felix Putzer
ISBN: 978-1-68083-814-5

Patient-Generated Health Data: Dimensions, Challenges, and Open Questions
Mayara Costa Figueiredo and Yunan Cheng
ISBN: 978-1-68083-668-4

Patterns and Themes in Designing with Children
Jessica Korte
ISBN: 978-1-68083-660-8

Gender-Inclusive HCI Research and Design: A Conceptual Review
Simone Stumpf, Anicia Peters, Shaowen Bardzell, Margaret Burnett, Daniela Busse, Jessica Cauchard, Elizabeth Churchill
ISBN: 978-1-68083-656-1

Auditing Algorithms

Understanding Algorithmic Systems from the Outside In

Danaë Metaxa
Stanford University

Joon Sung Park
Stanford University

Ronald E. Robertson
Northeastern University

Karrie Karahalios
University of Illinois at Urbana-Champaign

Christo Wilson
Northeastern University

Jeff Hancock
Stanford University

Christian Sandvig
University of Michigan

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Human-Computer Interaction

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

D. Metaxa et al.. *Auditing Algorithms*. Foundations and Trends® in Human-Computer Interaction, vol. 14, no. 4, pp. 272–344, 2021.

ISBN: 978-1-68083-917-3

Foundations and Trends® in Human-Computer Interaction
Volume 14, Issue 4, 2021

Editorial Board

Editorial Scope

Topics

Foundations and Trends® in Human-Computer Interaction publishes survey and tutorial articles in the following topics:

- History of the research community
- Theory
- Technology
- Computer Supported Cooperative Work
- Interdisciplinary influence
- Advanced topics and trends

Information for Librarians

Foundations and Trends® in Human-Computer Interaction, 2021, Volume 14, 4 issues. ISSN paper version 1551-3955. ISSN online version 1551-3963. Also available as a combined paper and online subscription.

Contents

Auditing Algorithms

Danaë Metaxa[1], Joon Sung Park[2], Ronald E. Robertson[3], Karrie Karahalios[4], Christo Wilson[5], Jeff Hancock[6] and Christian Sandvig[7]

[1] *Stanford University; metaxa@cs.stanford.edu*
[2] *Stanford University; joonspk@stanford.edu*
[3] *Northeastern University; robertson.ron@northeastern.edu*
[4] *University of Illinois at Urbana-Champaign; kkarahal@illinois.edu*
[5] *Northeastern University; cbw@ccs.neu.edu*
[6] *Stanford University; hancockj@stanford.edu*
[7] *University of Michigan; csandvig@umich.edu*

ABSTRACT

Algorithms are ubiquitous and critical sources of information online, increasingly acting as gatekeepers for users accessing or sharing information about virtually any topic, including their personal lives and those of friends and family, news and politics, entertainment, and even information about health and well-being. As a result, algorithmically-curated content is drawing increased attention and scrutiny from users, the media, and lawmakers alike. However, studying such content poses considerable challenges, as it is both dynamic and ephemeral: these algorithms are constantly changing, and frequently changing silently, with no record of the content to which users have been exposed over time. One strategy that has proven effective is the *algorithm audit*: a method of repeatedly querying an algorithm and observing its output in order to draw conclusions about the algorithm's opaque

Danaë Metaxa, Joon Sung Park, Ronald E. Robertson, Karrie Karahalios, Christo Wilson, Jeff Hancock and Christian Sandvig (2021), "Auditing Algorithms", Foundations and Trends® in Human-Computer Interaction: Vol. 14, No. 4, pp 272–344. DOI: 10.1561/1100000083.

inner workings and possible external impact. In this work, we present an overview of the algorithm audit methodology, including the history of audit studies in the social sciences from which this method is derived; a summary of key algorithm audits over the last two decades in a variety of domains, including health, politics, discrimination, and others; and a set of best practices for conducting algorithm audits today, contextualizing these practices using search engine audits as a case study. Finally, we conclude by discussing the social, ethical, and political dimensions of auditing algorithms, and propose normative standards for the use of this method.

1

An Introduction to Auditing

In 2012, Harvard professor Latanya Sweeney and her colleague found themselves Googling Dr. Sweeney's name, searching the web for a copy of a paper she had written. Instead, at the top of the search page they found an advertisement with the headline, "Latanya Sweeney. Arrested?" (Sweeney, 2013a). With no arrest record to speak of, Dr. Sweeney was shocked. After paying a fee to access the company's supposed information, she confirmed that the company's records did not contain any criminal information under her name. Investigating further, Dr. Sweeney and her colleague searched for his name, and found an advertisement from the same company—but this one simply offering information about people with that name, with no mention of an arrest record or anything of the sort. Searching for more and more names, Dr. Sweeney and her colleague were forced to conclude that it seemed like the advertisements Google was serving were racially biased, suggestive of arrest records more often for Black-sounding names like Dr. Sweeney's than white-sounding names like her colleague's. Well-equipped to study this phenomenon rigorously, Dr. Sweeney undertook a study collecting the ads served by Google for over 2,000 names of real people, using one set of names likely to belong to someone Black and another likely to belong to someone white. She found that Google's advertisements were up to 25% more likely to suggest an arrest record for a Black name than a white one, a discrepancy that was statistically significant and large enough that, were an employer disparately treating employees by race to this degree, the employer could potentially be charged with violating U.S. labor discrimination laws.

The reason for this racist discrepancy in ads being shown by Google is hard to identify conclusively; at worst, companies buying advertisements from Google could be purposefully targeting minority-sounding names. But the same outcome could result if companies provided Google with several versions of ad copy for the algorithm to automatically choose to maximize clicks, and people searching for Black-sounding names were for some reason more likely to click ads mentioning arrest, while people searching for white-sounding names were more likely to click on neutrally-worded ads. In any case, the implications are obviously serious. Imagine your potential employers, university admissions officers, or even your new partner's parents searching for your name on Google and finding ads that suggest an arrest record. The negative impact of such ads could be severe and immediate, and in this case, as Dr. Sweeney showed, it disproportionately affected Black people.

This kind of discrimination, apparent only in aggregate, is especially challenging to study in the context of computer systems whose exact workings are opaque to an outside observer. Sweeney's strategy, systematically querying the Google Search algorithm with a wide range of inputs and statistically comparing the results, is one of the most effective ways to study bias in algorithms. It is known as the *algorithm audit*. In this monograph, we present an overview of this powerful method including what it is, how it is used, and why it matters. We discuss the history of the audit method, its use in algorithm contexts, and best practices for researchers conducting algorithm audits in their own work. Our team of researchers has extensive experience conducting algorithm audits, and in this work we seek to answer such questions by drawing from the history of auditing in the social sciences as well as exemplary work auditing sociotechnical systems in recent decades.

1.1 What is an Audit?

Algorithm audits, our focus in most of this monograph, are a specific sub-type of a broader method, the audit study. Before we delve into the specifics of what makes a good audit and how auditing is applied to different social and sociotechnical contexts, we must define this method. Developed originally as a type of experiment used by social scientists, auditing is a methodology used to deploy randomized controlled experiments in a field setting (i.e., outside the lab) (Gaddis, 2018). Auditors conducting such a study must probe a process (e.g., a company's hiring process; a professor's process of responding to student emails; an algorithm providing users search results) by providing it with one or more inputs, while changing some attributes of that input, such as e.g., the race of the applicant (Bertrand and Mullainathan, 2004); the gender of the student (Milkman *et al.*, 2012); or the search history or date of search (Robertson *et al.*, 2018b; Metaxa *et al.*, 2019). Many governments, including that of the United States, conduct audits routinely, as a part of civic infrastructure. In the U.S., for instance, the Government Accountability Office conducts audits at the specific request of Congress or as mandated by law, and investigates the allocation of federal funds, allegations of illegal activity, the success of policies enacted, and other aspects of government function *U.S. Government Accountability Office (U.S. GAO)* 2021.

Bertrand and Mullainathan (2004) is a classic example of a (non-algorithmic) audit, one that inspired Latanya Sweeney's later work online. In that study, the authors sought to test whether there was racial bias in hiring, specifically in the resume reviewing stage, across a wide range of companies and industries. To do so, they constructed and sent fictitious resumes with white-sounding or Black-sounding names in response to job postings, and measured the rate at which those fictional job applicants got callbacks for interviews. They found that overall, applicants with white-sounding names received 50% more callbacks than those with Black-sounding names, and that the amount of discrimination was uniform across the industries they studied, concluding that racial discrimination was still widely prevalent in the labor market.

Algorithm audits are a specific subset of audit studies focused on studying algorithmic systems and content (Sandvig *et al.*, 2014). Rather than studying racial bias in human resume reviewing, then, an algorithm audit might investigate potential bias in an automated, algorithmically-powered resume screening process. Challenges specific to studying algorithms also lead algorithm audits to use different strategies and techniques—while Dr. Sweeney was able to manually search for Black- and white-sounding names and examine the search results displayed, algorithm auditors often need to build a software apparatus to amass large quantities of data from their platform of interest.

1.2 Differentiating Algorithm Audits from Other Testing

As evidenced by the examples we have already discussed, audit studies often—but not always—have an end goal of determining whether a system is biased or discriminatory. What all algorithm audits do have in common is an aim to test whether some deficiency (discrimination, bias, or something else) exists in an algorithmic system or not, without direct access to the internals of that system. In pursuit of this goal, there are several key features of audits that differentiate them from other types of testing, including the focus of study, scope of the conclusions drawn, and the position of the investigator while auditing.

Unlike other forms of testing such as A/B tests, the audit's subject of study is the system itself, not any particular component or a user's response to it. In an A/B test, for instance, the subject of study is the user, with the investigator seeking to understand the user's change in behavior while interacting with a system. Auditors may also be interested in a system's effect on people, but the angle of an audit is different, focused on the system itself. For auditors, studying the user is neither necessary nor sufficient; while some audit studies may include a component of user testing, audits more often measure the raw output of a system and rely on theory to infer what these outputs mean for a system's users. In the rare case that an audit does experiment on users, they are usually paid and consenting participants, rather than unknowing users of a system. This is often the case because measuring user behavior would be impossible (as when auditing a system one to which one does not have internal direct access), or unethical (we further discuss the ethics of auditing in Section 4).

Algorithm audits are also differentiated from other types of system testing by their scope. Most other forms of testing, including test suites, result in binary pass/fail conclusions at the level of individual test cases. An audit, on the other hand, has a broader scope and, it follows, must be systematic. It results in a declaration about the system as a whole; while auditors may conduct tests as part of their auditing, the overall finding of an audit is not merely to conclude that a given system is "right" or "wrong"—the results can only be discerned in aggregate. In this sense, an audit is a method of inspection or analysis more than of testing.

Finally, a third key difference is the role and position of the investigator conducting an audit study. A distinguishing feature of an audit study, unlike other forms of testing, is that an audit may be conducted with varying levels of participation or consent from the entity being audited—including partial or none at all. Audits are purposefully intended to be external evaluations, based only on outward-facing aspects, not insider knowledge on the process being studied. Most other testing is conducted internally, at the explicit direction of the proprietors of the system. This point raises interesting questions around the cost accrued when conducting an audit (for example, in system resources). Sending fake resumes to job postings costs companies employee time; auditing ads served by the Google search engine by repeatedly querying it uses Google's servers' resources. While most other forms of testing are conducted internally by a willing entity who bears the full cost, audits are conducted externally on an entity that is not necessarily willing or even informed of the ongoing audit, but the cost of the audit is shared between the investigators and the entity itself.

Before returning to algorithm-specific audits, in the the next section we will delve into the history of audit studies in the social sciences, establishing how the method was developed, what kinds of social systems it has been used to study, and what impacts these studies have had on the world.

1.3 Positionality Statement

As academic researchers in the United States with experience conducting search audits, we write primarily for fellow researchers interested in conducting them, with a secondary goal of speaking to an audience of academics, journalists, and others interested in interpreting and evaluating such research. Our team of authors has combined experience performing over 35 audits, covering areas including web search, social media, ridesharing, online marketplaces, online dating, and advertising.

As social computing researchers, in relation to the positionality of this work, we find it important to draw attention to the way the artifacts we study are usually specific to a time and place, rather than being universal or permanent. This influences our work in three important ways.

First, our own experience is necessarily limited by the contexts in which we have gained that experience. While we seek to provide a broad range of examples in this work, we focus many of those examples in Sections 3 and 4 on audits of search engines, where we have a particular depth of expertise. Further, many of the articles we reference come from the U.S. context; auditing itself is a broadly applicable practice, but the systems being audited and legal contexts surrounding audits vary widely, and the U.S. context is the one with which we are most familiar.

Second, the context dependence of social computing research impacts the goals of this article and its contributions. Since we expect these systems to develop and change over time, we seek to strike a balance between providing enough concrete details that other researchers in this domain can draw practical guidance from this work, while also focusing at a sufficiently high-level such that future researchers can understand the current moment from which we write—the motivations and considerations currently entailed in studying search after the specific details are deprecated.

Finally, as social computing researchers we also wish to draw attention to the potential for algorithm auditing to have significant political implications, a position we elaborate upon in Section 5. The algorithms that researchers such as ourselves audit are neither inevitable nor unchanging; rather, they are constantly in flux, and both constructed and used by people, and our work as auditors has the potential to change them, and in doing so to change the society in which they exist. As has been argued by scholars from the related field of Science and Technology Studies, ownership over algorithmic tools and data, along with the ability to monitor and understand them, increasingly yields power in our society (Milan and Van Der Velden, 2016; Chun, 2011). The possibility for direct change precipitated by an audit presents great opportunity as well as risk, and we hope this work will help researchers consider the politically weighty and socially important aspect of the work at hand as deeply as the technical advice we can provide.

1.4 Road Map

In the sections that follow, we aim to provide readers with an understanding of the algorithm auditing method, including its history and best practices. To do so, in Section 2 we begin by describing the auditing method's roots in the social sciences, prior to its use in the digital realm. Next, in Section 3, we move our focus to algorithm auditing, describing the method itself and summarizing key domains in which it is applied along with notable algorithm audits. In Section 4, we decompose algorithm audits into nine key dimensions, describing the choices available to auditors and providing recommended best practices within each. Before concluding, in Section 5, we further discuss the social implications of conducting audits and advocate for auditors to view this work through the lens of its broader social impacts.

References

Abebe, R., S. Hill, J. W. Vaughan, P. M. Small, and H. A. Schwartz. (2019). "Using Search Queries to Understand Health Information Needs in Africa." *Proceedings of the International AAAI Conference on Web and Social Media*. 13(01): 3–14. URL: https://ojs.aaai.org/index.php/ICWSM/article/view/3360.

Abid, A., M. Farooqi, and J. Zou. (2021). "Large language models associate Muslims with violence." *Nature Machine Intelligence*. 3(6): 461–463.

ACLU. (2020). "Federal Court Rules 'Big Data' Discrimination Studies Do Not Violate Federal Anti-Hacking Law." Mar. URL: https://www.aclu.org/press-releases/federal-court-rules-big-data-discrimination-studies-do-not-violate-federal-anti (accessed on 08/26/2020).

Allam, A., P. J. Schulz, and K. Nakamoto. (2014). "The Impact of Search Engine Selection and Sorting Criteria on Vaccination Beliefs and Attitudes: Two Experiments Manipulating Google Output." *Journal of Medical Internet Research*. 16(4): e100. DOI: 10.2196/jmir.2642.

Angwin, J., J. Larson, S. Mattu, and L. Kirchner. (2016). "Machine Bias." *ProPublica*. May. URL: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (accessed on 01/05/2021).

Angwin, J. and S. Mattu. (2016). "Amazon Says It Puts Customers First. But Its Pricing Algorithm Doesn't." *ProPublica*. Sept. URL: https://www.propublica.org/article/amazon-says-it-puts-customers-first-but-its-pricing-algorithm-doesnt (accessed on 01/05/2021).

Asplund, J., M. Eslami, H. Sundaram, C. Sandvig, and K. Karahalios. (2020a). "Auditing Race and Gender Discrimination in Online Housing Markets." In: *Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2020)*. AAAI. 10.

Asplund, J., M. Eslami, H. Sundaram, C. Sandvig, and K. Karahalios. (2020b). "Auditing race and gender discrimination in online housing markets." In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 14. 24–35.

Ballatore, A. (2015). "Google Chemtrails: A Methodology to Analyze Topic Representation in Search Engine Results." *First Monday*. 20(7). DOI: 10.5210/fm.v20i7.5597.

Ballatore, A., M. Graham, and S. Sen. (2017). "Digital Hegemonies: The Localness of Search Engine Results." *Annals of the American Association of Geographers*. 107(5): 1194–1215. DOI: 10.1080/24694452.2017.1308240.

Bandy, J. (2021). "Problematic Machine Behavior: A Systematic Literature Review of Algorithm Audits." arXiv: 2102.04256 [cs.CY].

Bandy, J. and N. Diakopoulos. (2020). "Auditing News Curation Systems: A Case Study Examining Algorithmic and Editorial Logic in Apple News." *Proceedings of the International AAAI Conference on Web and Social Media*. 14(1): 36–47. URL: https://ojs.aaai.org/index.php/ICWSM/article/view/7277.

Barocas, S., M. Hardt, and A. Narayanan. (2017). "Fairness in machine learning." *Nips tutorial*. 1: 2017.

Bechmann, A. and K. L. Nielbo. (2018). "Are We Exposed to the Same "News" in the News Feed?" *Digital Journalism*. 6(8): 990–1002. DOI: 10.1080/21670811.2018.1510741. eprint: https://doi.org/10.1080/21670811.2018.1510741.

Bernstein, M. S., M. Levi, D. Magnus, B. Rajala, D. Satz, and C. Waeiss. (2021). "ESR: Ethics and Society Review of Artificial Intelligence Research." arXiv: 2106.11521 [cs.CY].

Bertrand, M. and S. Mullainathan. (2004). "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination." *American economic review*. 94(4): 991–1013.

Bobrowsky, M. (2021). "Facebook Disables Access for NYU Research Into Political-Ad Targeting." *The Wall Street Journal*. Aug.

Brown, S., J. Davidovic, and A. Hasan. (2021). "The algorithm audit: Scoring the algorithms that score us." *Big Data & Society*. 8(1): 2053951720983865. DOI: 10.1177/2053951720983865. eprint: https://doi.org/10.1177/2053951720983865.

Buolamwini, J. (2019). "Hearing on Facial Recognition Technology (Part 1): Its Impact on our Civil Rights and Liberties." URL: https://www.congress.gov/116/meeting/house/109521/witnesses/HHRG-116-GO00-Wstate-BuolamwiniJ-20190522.pdf.

Buolamwini, J. and T. Gebru. (2018). "Gender shades: Intersectional accuracy disparities in commercial gender classification." In: *Conference on fairness, accountability and transparency*. PMLR. 77–91.

Cartright, M.-A., R. W. White, and E. Horvitz. (2011). "Intentions and Attention in Exploratory Health Search." In: *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. ACM. 65–74.

Chen, L., R. Ma, A. Hannák, and C. Wilson. (2018). "Investigating the Impact of Gender on Rank in Resume Search Engines." In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. Montreal QC, Canada: ACM Press. 1–14. DOI: 10.1145/3173574.3174225.

Chen, L., A. Mislove, and C. Wilson. (2015). "Peeking beneath the hood of uber." In: *Proceedings of the 2015 internet measurement conference*. 495–508.

Chowdhury, R. and J. Williams. (2021). "Introducing Twitter's first algorithmic bias bounty challenge." URL: https://blog.twitter.com/engineering/en_us/topics/insights/2021/algorithmic-bias-bounty-challenge.

Chun, W. H. K. (2011). *Programmed visions: Software and memory*. MIT Press.

Cobbe, J., M. S. A. Lee, and J. Singh. (2021). "Reviewable Automated Decision-Making: A Framework for Accountable Algorithmic Systems." In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. FAccT '21*. Virtual Event, Canada: Association for Computing Machinery. 598–609. DOI: 10.1145/3442188.3445921.

Coders' Rights Project. (2020). URL: https://www.eff.org/issues/coders (accessed on 10/27/2020).

Computing Machinery, A. for. (1992). "ACM Code of Ethics and Professional Conduct." *Code of Ethics*.

Correll, S. J., S. Benard, and I. Paik. (2007). "Getting a job: Is there a motherhood penalty?" *American journal of sociology*. 112(5): 1297–1338.

Crenshaw, K. (1990). "Mapping the margins: Intersectionality, identity politics, and violence against women of color." *Stan. L. Rev.* 43: 1241.

Daniel, W. W. (1968). *Racial discrimination in England: based on the PEP report*. Vol. 1084. Penguin.

Dave, K. (2019). "Systemic Algorithmic Harms." URL: https://points.datasociety.net/systemic-algorithmic-harms-e00f99e72c42 (accessed on 08/01/2021).

De Choudhury, M., M. R. Morris, and R. W. White. (2014). "Seeking and Sharing Health Information Online: Comparing Search Engines and Social Media." In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM. 1365–1376.

Diakopoulos, N., D. Trielli, J. Stark, and S. Mussenden. (2018). "I vote for—how search informs our choice of candidate." *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple, M. Moore and D. Tambini (Eds.)* 22.

Díaz, M., I. Johnson, A. Lazar, A. M. Piper, and D. Gergle. (2018). "Addressing age-related bias in sentiment analysis." In: *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–14.

DuckDuckGo. (2018). "Measuring the "Filter Bubble": How Google Is Influencing What You Click." URL: https://spreadprivacy.com/google-filter-bubble-study/.

Edelman, B., M. Luca, and D. Svirsky. (2017). "Racial discrimination in the sharing economy: Evidence from a field experiment." *American Economic Journal: Applied Economics*. 9(2): 1–22.

Edelman, B. G. and M. Luca. (2014). "Digital discrimination: The case of Airbnb.com." *Harvard Business School NOM Unit Working Paper*. (14-054).

Engler, A. C. (2021). "Independent auditors are struggling to hold AI companies accountable." *Fast Company*. Jan. URL: https://www.fastcompany.com/90597594/ai-algorithm-auditing-hirevue (accessed on 01/05/2021).

Epstein, R. and R. E. Robertson. (2015). "The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections." *Proceedings of the National Academy of Sciences*. 112(33): E4512–E4521. DOI: 10.1073/pnas.1419828112.

Epstein, R., R. E. Robertson, D. Lazer, and C. Wilson. (2017). "Suppressing the Search Engine Manipulation Effect (SEME)." *Proceedings of the ACM on Human-Computer Interaction*. 1(CSCW): 1–22. DOI: 10.1145/3134677.

Eriksson, M. C. and A. Johansson. (2017). "Tracking gendered streams." *Culture Unbound*. 9(2): 163–183.

Fortunato, S., A. Flammini, F. Menczer, and A. Vespignani. (2006). "Topical Interests and the Mitigation of Search Engine Bias." *Proceedings of the National Academy of Sciences*. 103(34): 12684–12689. DOI: 10.1073/pnas.0605525103.

Gaddis, S. M. (2018). *Audit studies: Behind the scenes with theory, method, and nuance*. Vol. 14. Springer.

GAO), U. G. A. O. ((2021). "Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities." URL: https://www.gao.gov/products/gao-21-519sp.

Ginsberg, J., M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant. (2009). "Detecting Influenza Epidemics Using Search Engine Query Data." *Nature*. 457(7232): 1012.

Gneezy, U. and J. List. (2004). "Are the disabled discriminated against in product markets? Evidence from field experiments." In: *American Economic Association Annual Meeting*.

Goldman, E. (2008). "Search engine bias and the demise of search engine utopianism." In: *Web Search*. Springer. 121–133.

Google Terms of Service. (2017). URL: https://policies.google.com/terms?hl=en-US (accessed on 08/22/2019).

"Google China." (2021). https://en.wikipedia.org/wiki/Google_China. (Accessed on 08/01/2021).

Granka, L. A. (2010). "The Politics of Search: A Decade Retrospective." *The Information Society*. 26(5): 364–374. DOI: 10.1080/01972243.2010.511560.

Gross, T. (2017). "A 'Forgotten History' Of How The U.S. Government Segregated America." *NPR*. May. URL: https://www.npr.org/2017/05/03/526655831/a-forgotten-history-of-how-the-u-s-government-segregated-america (accessed on 01/05/2021).

Hannak, A., P. Sapiezynski, A. Molavi Kakhki, B. Krishnamurthy, D. Lazer, A. Mislove, and C. Wilson. (2013). "Measuring Personalization of Web Search." In: *Proceedings of the 22nd International Conference on World Wide Web - WWW '13*. Rio de Janeiro, Brazil: ACM Press. 527–538. DOI: 10.1145/2488388.2488435.

Hannak, A., G. Soeller, D. Lazer, A. Mislove, and C. Wilson. (2014). "Measuring price discrimination and steering on e-commerce web sites." In: *Proceedings of the 2014 conference on internet measurement conference*. 305–318.

Hannak, A., C. Wagner, D. Garcia, A. Mislove, M. Strohmaier, and C. Wilson. (2017). "Bias in Online Freelance Marketplaces: Evidence from TaskRabbit and Fiverr." In: *20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*. Portland, OR.

Hu, D., S. Jiang, R. E. Robertson, and C. Wilson. (2019). "Auditing the Partisanship of Google Search Snippets." In: *The World Wide Web Conference on - WWW '19*. San Francisco, CA, USA: ACM Press. 693–704. DOI: 10.1145/3308558.3313654.

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

Hutchinson, B. and M. Mitchell. (2019). "50 Years of Test (Un)Fairness: Lessons for Machine Learning." In: *Proceedings of the Conference on Fairness, Accountability, and Transparency. FAT* '19*. Atlanta, GA, USA: Association for Computing Machinery. 49–58. DOI: 10.1145/3287560.3287600.

Introna, L. D. and H. Nissenbaum. (2000). "Shaping the Web: Why the Politics of Search Engines Matters." *The Information Society*. 16(3): 169–185. DOI: 10.1080/01972240050133634.

Joachims, T., L. Granka, B. Pan, H. Hembrooke, F. Radlinski, and G. Gay. (2007). "Evaluating the Accuracy of Implicit Feedback from Clicks and Query Reformulations in Web Search." *ACM Transactions on Information Systems*. 25(2): 7–es. DOI: 10.1145/1229179.1229181.

Juneja, P. and T. Mitra. (2021). "Auditing E-Commerce Platforms for Algorithmically Curated Vaccine Misinformation." In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI '21*. Yokohama, Japan: Association for Computing Machinery. DOI: 10.1145/3411764.3445250.

Karahalios, K. (2018). "Discrimination Audits: Challenges to Discrimination Studies." URL: http://social.cs.illinois.edu/presentations/law-data-summit/karahalios.pdf.

Kawakami, A., K. Umarova, and E. Mustafaraj. (2020). "The Media Coverage of the 2020 US Presidential Election Candidates through the Lens of Google's Top Stories." In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 14. 868–877.

Kay, M., C. Matuszek, and S. A. Munson. (2015). "Unequal Representation and Gender Stereotypes in Image Search Results for Occupations." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. Seoul, Republic of Korea: ACM Press. 3819–3828. DOI: 10.1145/2702123.2702520.

Keyes, O., J. Hutson, and M. Durbin. (2019). "A Mulching Proposal: Analysing and Improving an Algorithmic System for Turning the Elderly into High-Nutrient Slurry." In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. CHI EA '19*. Glasgow, Scotland Uk: Association for Computing Machinery. 1–11. DOI: 10.1145/3290607.3310433.

Kliman-Silver, C., A. Hannak, D. Lazer, C. Wilson, and A. Mislove. (2015). "Location, Location, Location: The Impact of Geolocation on Web Search Personalization." In: *Proceedings of the 2015 ACM Conference on Internet Measurement Conference - IMC '15*. Tokyo, Japan: ACM Press. 121–127. DOI: 10.1145/2815675.2815714.

Koshiyama, A., E. Kazim, P. Treleaven, P. Rai, L. Szpruch, G. Pavey, G. Ahamat, F. Leutner, R. Goebel, A. Knight, *et al.* (2021). "Towards Algorithm Auditing: A Survey on Managing Legal, Ethical and Technological Risks of AI, ML and Associated Algorithms."

Kugelmass, H. (2016). ""Sorry, I'm Not Accepting New Patients" an audit study of access to mental health care." *Journal of Health and Social Behavior*. 57(2): 168–183.

Kulshrestha, J., M. Eslami, J. Messias, M. B. Zafar, S. Ghosh, K. P. Gummadi, and K. Karahalios. (2019). "Search Bias Quantification: Investigating Political Bias in Social Media and Web Search." *Information Retrieval Journal*. 22(1-2): 188–227. DOI: 10.1007/s10791-018-9341-2.

Lampos, V., A. C. Miller, S. Crossan, and C. Stefansen. (2015). "Advances in Nowcasting Influenza-like Illness Rates Using Search Query Logs." *Scientific reports*. 5: 12760.

Lawrence, S. and C. L. Giles. (1999). "Accessibility of Information on the Web." *Nature*. 400(6740): 107–107. DOI: 10.1038/21987.

Lazer, D., R. Kennedy, G. King, and A. Vespignani. (2014). "The Parable of Google Flu: Traps in Big Data Analysis." *Science*. 343(6176): 1203–1205.

Le, H., R. Maragh, B. Ekdale, A. High, T. Havens, and Z. Shafiq. (2019). "Measuring Political Personalization of Google News Search." In: *The World Wide Web Conference on - WWW '19*. San Francisco, CA, USA: ACM Press. 2957–2963. DOI: 10.1145/3308558.3313682.

Lurie, E. and E. Mustafaraj. (2018). "Investigating the Effects of Google's Search Engine Result Page in Evaluating the Credibility of Online News Sources." In: *Proceedings of the 10th ACM Conference on Web Science - WebSci '18*. Amsterdam, Netherlands: ACM Press. 107–116. DOI: 10.1145/3201064.3201095.

Makhortykh, M., A. Urman, and R. Ulloa. (2020). "How Search Engines Disseminate Information about COVID-19 and Why They Should Do Better." *Harvard Kennedy School Misinformation Review*. May. DOI: 10.37016/mr-2020-017.

Matias, J. N. (2020). "Why We Need Industry-Independent Research on Tech & Society." URL: https://citizensandtech.org/2020/01/industry-independent-research/.

McCown, F. and M. L. Nelson. (2007). "Agreeing to Disagree: Search Engines and Their Public Interfaces." In: *Proceedings of the 2007 Conference on Digital Libraries - JCDL '07*. Vancouver, BC, Canada: ACM Press. 309. DOI: 10.1145/1255175.1255237.

McCullagh, D. (2006). "AOL's Disturbing Glimpse into Users' Lives." CNET. URL: https://www.cnet.com/news/aols-disturbing-glimpse-into-users-lives/.

McMahon, C., I. Johnson, and B. Hecht. (2017). "The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship between Peer Production Communities and Information Technologies." In: *Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2018)*. AAAI. 10.

Mehrotra, R., A. Anderson, F. Diaz, A. Sharma, H. Wallach, and E. Yilmaz. (2017). "Auditing Search Engines for Differential Satisfaction Across Demographics." In: *Proceedings of the 26th International Conference on World Wide Web Companion - WWW '17 Companion*. Perth, Australia: ACM Press. 626–633. DOI: 10.1145/3041021.3054197.

Metaxa, D., J. S. Park, J. A. Landay, and J. Hancock. (2019). "Search Media and Elections: A Longitudinal Investigation of Political Search Results in the 2018 U.S. Elections." In: *Proceedings of the 22nd ACM Conference on Computer-Supported Cooperative Work and Social Computing*. ACM.

Metaxas, P. T. and E. Mustafaraj. (2009). "The Battle for the 2008 US Congressional Elections on the Web." In: *Proceedings of the 2009 WebScience: Society On-Line Conference*.

Metaxas, P. T. and Y. Pruksachatkun. (2017). "Manipulation of Search Engine Results during the 2016 US Congressional Elections." In: *Proceedings of the ICIW 2017*. 6.

Mikians, J., L. Gyarmati, V. Erramilli, and N. Laoutaris. (2012). "Detecting price and search discrimination on the internet." In: *Proceedings of the 11th ACM workshop on hot topics in networks*. 79–84.

Milan, S. and L. Van Der Velden. (2016). "The alternative epistemologies of data activism." *Digital Culture & Society*. 2(2): 57–74.

Milkman, K. L., M. Akinola, and D. Chugh. (2012). "Temporal Distance and Discrimination: An Audit Study in Academia." *Psychological Science*. 23(7): 710–717. DOI: 10.1177/0956797611434539. eprint: https://doi.org/10.1177/0956797611434539.

Mowshowitz, A. and A. Kawaguchi. (2002a). "Assessing Bias in Search Engines." *Information Processing & Management*. 38(1): 141–156. DOI: 10.1016/S0306-4573(01)00020-6.

Mowshowitz, A. and A. Kawaguchi. (2002b). "Bias on the Web." *Communications of the ACM*. 45(9). DOI: 10.1145/567498.567527.

Mowshowitz, A. and A. Kawaguchi. (2005). "Measuring Search Engine Bias." *Information Processing & Management*. 41(5): 1193–1205. DOI: 10.1016/j.ipm.2004.05.005.

Mullainathan, S., M. Noeth, and A. Schoar. (2012). "The market for financial advice: An audit study." *Tech. rep.* National Bureau of Economic Research.

Mustafaraj, E., E. Lurie, and C. Devine. (2020). "The Case for Voter-Centered Audits of Search Engines During Political Elections." In: *Conference on Fairness, Accountability, and Transparency (FAT* '20)*. 11.

Noble, S. U. (2013). "Google search: Hyper-visibility as a means of rendering black women and girls invisible." *InVisible Culture*. (19).

Noble, S. U. (2018a). *Algorithms of oppression*. New York University Press.

Noble, S. U. (2018b). "Google Has a Striking History of Bias Against Black Girls." *Time Magazine*.

Nosek, B. A. and T. M. Errington. (2020). "What is replication?" *PLoS biology*. 18(3): e3000691.

Obermeyer, Z., B. Powers, C. Vogeli, and S. Mullainathan. (2019). "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science*. 366(6464): 447–453. DOI: 10.1126/science.aax2342.

Oreopoulos, P. (2011). "Why do skilled immigrants struggle in the labor market? A field experiment with thirteen thousand resumes." *American Economic Journal: Economic Policy*. 3(4): 148–71.

Pan, B., H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, and L. Granka. (2007). "In Google We Trust: Users' Decisions on Rank, Position, and Relevance." *Journal of Computer-Mediated Communication*. 12(3): 801–823. DOI: 10.1111/j.1083-6101.2007.00351.x.

Pitoura, E., P. Tsaparas, G. Flouris, I. Fundulaki, P. Papadakos, S. Abiteboul, and G. Weikum. (2018). "On Measuring Bias in Online Information." *ACM SIGMOD Record*. 46(4): 16–21. DOI: 10.1145/3186549.3186553.

President, E. O. of the. (2016). "Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights." URL: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf.

Prosperi, M. and J. Bian. (2019). "Is it time to rethink institutional review boards for the era of big data?" *Nature Machine Intelligence*. 1(6): 260–260.

Purcell, K. and J. Brenner. (2012). "Search Engine Use 2012." URL: http://www.pewinternet.org/2012/03/09/search-engine-use-2012/.

Raji, I. D. and J. Buolamwini. (2019). "Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products." In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. 429–435.

Robertson, A. (2019). "Scraping public data from a website probably isn't hacking, says court." The Verge. URL: https://www.theverge.com/2019/9/10/20859399/linkedin-hiq-data-scraping-cfaa-lawsuit-ninth-circuit-ruling.

Robertson, R. E., S. Jiang, K. Joseph, L. Friedland, D. Lazer, and C. Wilson. (2018a). "Auditing Partisan Audience Bias within Google Search." *Proceedings of the ACM on Human-Computer Interaction*. 2(CSCW): 1–22. DOI: 10.1145/3274417.

Robertson, R. E., S. Jiang, D. Lazer, and C. Wilson. (2019). "Auditing Autocomplete: Suggestion Networks and Recursive Algorithm Interrogation." In: *Proceedings of the 10th ACM Conference on Web Science - WebSci '19*. Boston, Massachusetts, USA: ACM Press. 235–244. DOI: 10.1145/3292522.3326047.

Robertson, R. E., D. Lazer, and C. Wilson. (2018b). "Auditing the Personalization and Composition of Politically-Related Search Engine Results Pages." In: *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*. Lyon, France: ACM Press. 955–965. DOI: 10.1145/3178876.3186143.

Rothshild, A., E. Lurie, and E. Mustafaraj. (2019). "How the Interplay of Google and Wikipedia Affects Perceptions of Online News Sources." In: *Computation+ Journalism Symposium*.

Sandvig, C., K. Hamilton, K. Karahalios, and C. Langbort. (2014). "Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms." *Data and discrimination: converting critical concerns into productive inquiry*. 22.

Sapiezynski, P., W. Zeng, R. E. Robertson, A. Mislove, and C. Wilson. (2019). "Quantifying the Impact of User Attention on Fair Group Representation in Ranked Lists." In: *Companion Proceedings of The 2019 World Wide Web Conference on - WWW '19*. San Francisco, USA: ACM Press. 553–562. DOI: 10.1145/3308560.3317595.

Sharma, R., A. Mitra, and M. Stano. (2015). "Insurance, race/ethnicity, and sex in the search for a new physician." *Economics Letters*. 137: 150–153.

Shen, H., A. DeVos, M. Eslami, and K. Holstein. (2021). "Everyday algorithm auditing: Understanding the power of everyday users in surfacing harmful algorithmic behaviors." *arXiv preprint arXiv:2105.02980*.

Siegelman, P. and J. Heckman. (1993). "The Urban Institute audit studies: Their methods and findings." *Clear and Convincing Evidence: Measurement of Discrimination in America, Washington*. 187: 258.

Sloane, M. (2021). "The Algorithmic Auditing Trap." URL: https://onezero.medium.com/the-algorithmic-auditing-trap-9a6f2d4d461d.

Speicher, T., M. Ali, G. Venkatadri, F. Ribeiro, G. Arvanitakis, F. Benevenuto, K. Gummadi, P. Loiseau, and A. Mislove. (2018). "Potential for discrimination in online targeted advertising." In: *Conference on Fairness, Accountability, and Transparency (FAT '20*.

"Search Engine Market Share China." (2019). URL: https://gs.statcounter.com/search-engine-market-share/all/china (accessed on 08/22/2019).

"Search Engine Market Share Worldwide." (2019). URL: https://gs.statcounter.com/search-engine-market-share (accessed on 08/22/2019).

Sweeney, L. (2013a). "Discrimination in Online Ad Delivery." URL: https://dataprivacylab.org/projects/onlineads/index.html (accessed on 10/26/2020).

Sweeney, L. (2013b). "Discrimination in Online Ad Delivery." *Queue*. 11(3): 10:10–10:29. DOI: 10.1145/2460276.2460278.

The Economist. (2019). "Google's Algorithm – Seek and You Shall Find." *The Economist*. June: 81 (US).

Thomas, L. (2021). "Supreme Court ruling that limits hacking law supports U-M researcher."

Tober, M., J. Grundmann, and A. Thakur. (2016). "Universal & Extended Search 2016: Facts, Trends and Optimization Tips." *Tech. rep.* searchmetrics.

Trielli, D. and N. Diakopoulos. (2019). "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 - CHI '19*. Glasgow, Scotland Uk: ACM Press. 1–15. DOI: 10.1145/3290605.3300683.

Trielli, D. and N. Diakopoulos. (2020). "Partisan Search Behavior and Google Results in the 2018 U.S. Midterm Elections." *Information, Communication & Society*. May: 1–17. DOI: 10.1080/1369118X.2020.1764605.

Tripodi, F. (2018). "Searching for Alternative Facts." *Tech. rep.* Data&Society. 64.

"U.S. Government Accountability Office (U.S. GAO)." URL: https://www.gao.gov (accessed on 08/01/2021).

Union, A. C. L. (2019). "Sandvig v. Barr – Challenge to CFAA Prohibition on Uncovering Racial Discrimination Online." May. URL: https://www.aclu.org/cases/sandvig-v-barr-challenge-cfaa-prohibition-uncovering-racial-discrimination-online (accessed on 01/05/2021).

Vaidhyanathan, S. (2011). *The Googlization of everything*. University of California Press.

Vincent, N. and B. Hecht. (2021). "A Deeper Investigation of the Importance of Wikipedia Links to Search Engine Results." *Proc. ACM Hum.-Comput. Interact.* 5(CSCW1). DOI: 10.1145/3449078.

Vincent, N., I. Johnson, P. Sheehan, and B. Hecht. (2019). "Measuring the Importance of User-Generated Content to Search Engines." In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 13.

White, R. W. and E. Horvitz. (2009). "Cyberchondria: Studies of the Escalation of Medical Concerns in Web Search." *ACM Transactions on Information Systems (TOIS)*. 27(4): 23.

Wilson, C., A. Ghosh, S. Jiang, A. Mislove, L. Baker, J. Szary, K. Trindel, and F. Polli. (2021). "Building and Auditing Fair Algorithms: A Case Study in Candidate Screening." In: *Conference on Fairness, Accountability, and Transparency (FAccT '21)*. 12.

Xing, X., W. Meng, D. Doozan, N. Feamster, W. Lee, and A. C. Snoeren. (2014). "Exposing Inconsistent Web Search Results with Bobble." In: *Passive and Active Measurement*. Ed. by M. Faloutsos and A. Kuzmanovic. *Lecture Notes in Computer Science*. Cham: Springer International Publishing. 131–140. DOI: 10.1007/978-3-319-04918-2_13.

Yinger, J. (1998). "Evidence on discrimination in consumer markets." *Journal of Economic perspectives*. 12(2): 23–40.

Zetter, K. (2010). "Wiseguys Plead Guilty in Ticketmaster Captcha Case." Wired. URL: https://www.wired.com/2010/11/wiseguys-plead-guilty/.

Zetter, K. (2015). "Experts Say Myspace Suicide Indictment Sets 'Scary' Legal Precedent." Wired. URL: https://www.wired.com/2008/05/myspace-indictm/.

Zetter, K. (2016). "Researchers Sue the Government Over Computer Hacking Law." Wired. URL: https://www.wired.com/2016/06/researchers-sue-government-computer-hacking-law/.