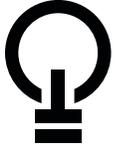


DATA AND DISCRIMINATION: COLLECTED ESSAYS



EDITED BY SEETA PEÑA GANGADHARAN
WITH VIRGINIA EUBANKS AND SOLON BAROCAS



OCTOBER 2014

An Algorithm Audit

CHRISTIAN SANDVIG

ASSOCIATE PROFESSOR, COMMUNICATION STUDIES AND SCHOOL OF INFORMATION, UNIVERSITY OF MICHIGAN

KEVIN HAMILTON

ASSOCIATE DEAN OF RESEARCH, COLLEGE OF FINE AND APPLIED ARTS AND ASSOCIATE PROFESSOR OF NEW MEDIA AND PAINTING, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

KARRIE KARAHALIOS

ASSOCIATE PROFESSOR, COMPUTER SCIENCE AND DIRECTOR, CENTER FOR PEOPLE & INFRASTRUCTURES, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

CEDRIC LANGBORT

ASSOCIATE PROFESSOR, AEROSPACE ENGINEERING AND CO-DIRECTOR, CENTER FOR PEOPLE AND INFRASTRUCTURES, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

AN ALGORITHM AUDIT

When it is time to buy a used car, many consumers turn to the advice of a trusted third-party like the Consumers Union, publisher of Consumer Reports. While we may not know anything about how cars work, Consumer Reports operates a test track where automotive experts run cars through their paces. Even better, to devise its public rating for a particular model Consumer Reports sends current owners a survey to draw conclusions from their past experiences. Finally, Consumer Reports is trustworthy because it is a non-profit advocacy organization collectively organized by consumers with no relationship to the auto industry.

We need a Consumer Reports for algorithms.

Invisible Algorithms Dominate Our Everyday Life

Computer algorithms now dominate our daily life, providing our communication with our family and friends, our search for housing, our media preferences, our driving directions, the advertisements that we see, the information we look up, encryption of our data for our privacy, and more.

Yet there is a tremendous gap between public understanding of algorithms and their prevalence and importance in our lives. For instance, the majority of Facebook users in a recent study did not even know that Facebook ever used an algorithm to filter the news stories that they saw.¹

Unfair Algorithms, Undetectable Without Help

Algorithms differ from earlier processes of harmful discrimination (such as redlining) in a number of crucial ways. First, algorithms that affect large number of people (e.g., the Google search algorithm) are complicated packages of computer code crafted jointly by a large team of engineers.

These algorithms represent trade secrets.

Second, the computer code for an algorithm does not make it interpretable. At the level of complexity that is typical for these systems, an algorithm cannot be interpreted by reading it. Even an expert in the area (or the algorithm's authors) may not be able to predict what results an algorithm would produce without plugging in some example data and looking at the results.

Third, algorithms also increasingly depend on unique personal data as inputs. As a result, the same programmatically generated Web page may never be generated twice.

Finally, we have little reason to believe the companies we depend on will act in our interest in the absence of regulatory oversight. Almost every major operator of an Internet platform, including Google, Twitter, Facebook, Microsoft, and



“Dislike Graffiti.” Photo by [zeevveez](#). CC-BY-2.0.

Apple, has already been investigated by the U.S. government for violations that include anti-competitive behavior, deceptive business practices, failing to protect the personal information of consumers, failing to honor promises made to consumers about their own data, and charging customers for purchases that they did not authorize.²

Testing the Platforms that Test Us

Luckily, a method exists for researchers to look inside these complicated, algorithmically driven computer decision systems: the “audit study”.³ This method, which serves as *the* most respected social scientific method for the detection of racial discrimination in employment and housing, uses fictitious correspondence. For instance, an audit study might submit fictitious resumes targeted at

a real employer or fictitious housing applications targeted at a real landlord. In these studies, researchers test the fairness of an employer or landlord by preparing two or more equivalent documents which reflect equal backgrounds, including levels of education and experience, but which only vary according to race. For example, researchers could manipulate the fictitious applicant’s race between the two conditions of “Emily” and “Lakisha” to signal “Caucasian” vs. “African-American” to a prospective employer. The difference in employer responses to two otherwise identical resumes therefore measures racism.

In the spirit of these real-life audits of employers and real estate agents performed by journalists and watchdog organizations, we propose that the



Lemons at a market. Photo by [MarcusObal](#). CC-BY-SA-3.0.

advantage of them; platform “lemon warnings” that can explain the operation of faulty or deceptive social media sites; and quality rankings which tell us when an algorithm is working for us or for someone else.

recent concerns about algorithms demand an audit of online platforms. In essence, this means Internet platforms powered by large amounts of data (e.g., YouTube, Google, Facebook, Netflix, and so on) that are operated via secret computer algorithms require testing by an impartial expert third party. These audits will ascertain whether algorithms result in harmful discrimination by class, race, gender, geography, or other important attributes.

Although the complexity of these algorithmic platforms makes them seem impossible to understand, audit studies can crack the code through trial and error: researchers can apply expert knowledge to the results of these audit tests. By closely monitoring these online platforms, we can discover interactions between algorithm and data. In short, auditing these algorithms demands a third party that can combine both expert and everyday evaluations, testing algorithms on the public’s behalf and investigating and reporting situations where algorithms may have gone wrong.

Lemon Warnings in a Data-Driven Society

We envision a future where Internet users can know in advance if a search box is planning to take

Index

1. Christian Sandvig, Karrie Karahalios, and Cedric Langbort, *Uncovering Algorithms: Looking Inside the Facebook News Feed*, In the Berkman Center Seminar Series: Berkman Center for Internet & Society, Harvard University (July 22, 2014): <http://cyber.law.harvard.edu/events/luncheon/2014/07/sandvigkarahalios>.
2. US Department of Justice, *United States v. Microsoft Corporation*, Civil Action No. 98-1232 (1999): http://www.justice.gov/atr/cases/ms_index.htm#other; Federal Trade Commission, *In the matter of Twitter, Inc., a corporation*, File number 092 3093 (2010): <http://www.ftc.gov/enforcement/cases-proceedings/092-3093/twitter-inc-corporation>; Federal Trade Commission, *In the matter of Google, Inc. a corporation*, File number 102 3136. (2011): <http://www.ftc.gov/enforcement/cases-proceedings/102-3136/google-inc-matter>; Federal Trade Commission, *In the matter of Facebook, Inc., a corporation*, File number 092 3184 (2011): <http://www.ftc.gov/enforcement/cases-proceedings/092-3184/facebook-inc>; Federal Trade Commission, *In the matter of Apple, Inc., a corporation*, File number 112 3108 (2014): <http://www.ftc.gov/enforcement/cases-proceedings/112-3108/apple-inc>.
3. Devah Pager, "The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future," *The Annals of the American Academy of Political and Social Science* 609, no. 1 (2007): 104-33.