Critical Algorithm Studies: a Reading List

This list is an attempt to collect and categorize a growing critical literature on algorithms as social concerns. The work included spans sociology, anthropology, science and technology studies, geography, communication, media studies, and legal studies, among others. Our interest in assembling this list was to catalog the emergence of “algorithms” as objects of interest for disciplines beyond mathematics, computer science, and software engineering.

As a result, our list does not contain much writing by computer scientists, nor does it cover potentially relevant work on topics such as quantification, rationalization, automation, software more generally, or big data, although these interests are well-represented in these works’ reference sections of the essays themselves.

This area is growing in size and popularity so quickly that many contributions are popping up without reference to work from disciplinary neighbors. One goal for this list is to help nascent scholars of algorithms to identify broader conversations across disciplines and to avoid reinventing the wheel or falling into analytic traps that other scholars have already identified. We also thought it would be useful, especially for those teaching these materials, to try to loosely categorize it. The organization of the list is meant merely as a first-pass, provisional sense-making effort. Within categories the entries are offered in chronological order, to help make sense of these rapid developments.

In light of all of those limitations, we encourage you to see it as an unfinished document, and we welcome comments. These could be recommendations of other work to include, suggestions on how to reclassify a particular entry, or ideas for reorganizing the categories themselves. Please use the comment space at the bottom of the page to offer suggestions and criticism; we will try to update the list in light of these suggestions.

<https://socialmediacollective.org/reading-lists/critical-algorithm-studies/>

# Critical Data and Algorithm Studies

Progressing digitalization, pervasive mobile computing, and omnipresent social media platforms with billions of users have led to the rise of big data analytics and the implementation of data-driven research approaches in a wide variety of scientific fields. Recently, the critical study of digital practices related to data and algorithms as well as the examination of implications of data science for privacy and surveillance have attracted considerable attention, not just since recent data scandals that severely harmed public trust in social media and web research.   
  
In the emerging fields of Critical Data Studies (1,2,3) and Critical Algorithm Studies (4,5,6) scholars systematically study and tackle legal, ethical and social challenges of data science. Data scientists themselves have long documented their critical engagement with the creation, collection, storage and analysis of big data and research integrity (7,8,9). However, the focus on important constraints and limitations of computational methods, such as blackboxed methods and opaque access to data, the resulting biases, issues of sampling and representation, external validity and evaluation – just to name a few – is discussed only at the margins of data science. Furthermore, on the solution-oriented, computational end of the spectrum, we often lack the productive intersection of cultures of critique with those of practice.  
  
Therefore, this special issue is dedicated to bringing together critical expertise of scientists in data-driven research areas, who reflect their daily routines, their methods, data sources and the social impact of their research. We would also like to give space to those experiences coming from newly established collaborations of computer scientists with social scientists and humanities’ scholars, moreover with policy makers, activists, or in citizen science projects. The focus is on the critical reflection of scientific methods, data sources, modeling, validation, replication, and review procedures including questions of their impact regarding social behaviour, power relations, ethics, and accountability, thus the performative and normative aspects of data science practices.  
  
The special issue welcomes contributions that utilize data science to engage with methodological, theoretical, practical, and ethical issues of core scientific practices and discuss complementary or alternative routes to robust scientific insights. The special issue seeks to describe and discuss the following topics from a critical data studies perspective:  
  
1. Data collection, modeling, data mining, data sources, and data processing in general  
2. Methodological issues, such as bias, limits, scaling, etc. for data mining, machine learning, social network analysis, text analysis  
3. Visual analytics and information visualization  
4. Reproducibility, archiving, stewardship and digital preservation  
5. Data security and protection as well as privacy and data regulation  
6. Education and training in critical data science  
  
The editors encourage interdisciplinary co-authorships, e.g. from computer science and critical data studies, as well as reflections on projects on newly developed or prototyped solutions to overcome some of the limitations and issues identified by critical analysis.  
  
Authors can submit abstracts and manuscripts via this website by clicking on "Submit your abstract" and "Submit your manuscript", respectively.  
For abstract submission: Title (500 chars max) and text (1000 words max).  
For manuscript submission, please follow the author guidelines: https://www.frontiersin.org/about/author-guidelines.

<https://www.frontiersin.org/research-topics/9570/critical-data-and-algorithm-studies>

# Critical Algorithm Lab (CALL)

### New methods in the meeting between qualitative and quantitative social data

Among other things, the Critical Algorithm Lab (CALL) research group studies researchers who work in a world of websites, social media and large online databases, and identifies the role that social big data can play in new forms of interdisciplinary social science.

At the same time, the group experiments with using social big data in new ways, since new digital data formats, tools and methods call for a completely different approach to describing, analysing and interpreting data in social science research.

When researchers work with social big data that can be downloaded from digital platforms, they need to be very aware of the political and ethical aspects of working with these data. In the Critical Algorithm Lab project, researchers from the fields of anthropology, sociology, and science and technology studies work on combining qualitative and quantitative data and methods in social science research.

In practice this means that the CALL researchers both study how other researchers handle new forms of digital data based on ethical and practical considerations, and at the same time take part in new interdisciplinary collaborations.

This is because digital data has a number of characteristics that challenge traditional distinctions in social science, not least the classic dividing line between quantitative and qualitative research methods.

Some researchers speak of new “quali-quantitative” methods, in which the qualitative focus on specific local situations and the bigger, quantitative picture of social patterns formed on the basis of the new enormous digital databases are no longer mutually exclusive.

But these new quali-quantitative research methods are still in their infancy, and there is a need for an ongoing methodological development and a critical evaluation of the researchers’ approach to the data material and their analytical tools. The ethical issues that arise when researchers gain access to data that originates from private individuals, companies and public authorities also need to be discussed in more detail.

It is therefore important to experiment with new ways to link data that may at first sight seem radically different. This applies, for example, to the way new digital data from mobile phones can be combined with more established methods of sociological interviews or anthropological fieldwork.

These are research methods that may provide complementary insights into our social life and the way we form friendships and networks. Through practical experiments with these types of “data stitching”, the researchers in the Critical Algorithm Lab project will contribute to a better understanding of the potentials and pitfalls that arise when researchers and others swoop onto the new kinds of big social data.

This research group is headed by Professor of Anthropology [Morten Axel Pedersen](https://anthropology.ku.dk/staff/professor-and-associate-professor/?pure=en/persons/255694) and Associate Professor of Sociology [Anders Blok](https://www.sociology.ku.dk/staff/professor-and-associate-professor/?pure=en/persons/182002).

<https://sodas.ku.dk/projects/call/>

# Amazon scraps a secret A.I. recruiting tool that showed bias against women

**PUBLISHED WED, OCT 10 2018 • 6:15 AM EDT  UPDATED THU, OCT 11 2018 • 2:25 PM EDT**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |

**KEY POINTS**

* Amazon.com’s machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.
* The team had been building computer programs since 2014 to review job applicants’ resumes with the aim of mechanizing the search for top talent, five people familiar with the effort told Reuters.
* The company’s experimental hiring tool used artificial intelligence to give job candidates scores ranging from one to five stars — much like shoppers rate products on Amazon, some of the people said.

**An Amazon Prime package**

*Andrew Harrer | Bloomberg | Contributor*

[Amazon.com’s](https://www.cnbc.com/quotes/?symbol=AMZN) machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

The team had been building computer programs since 2014 to review job applicants’ resumes with the aim of mechanizing the search for top talent, five people familiar with the effort told Reuters.

Automation has been key to Amazon’s e-commerce dominance, be it inside warehouses or driving pricing decisions. The company’s experimental hiring tool used artificial intelligence to give job candidates scores ranging from one to five stars — much like shoppers rate products on Amazon, some of the people said.

“Everyone wanted this holy grail,” one of the people said. “They literally wanted it to be an engine where I’m going to give you 100 resumes, it will spit out the top five, and we’ll hire those.”

But by 2015, the company realized its new system was not rating candidates for software developer jobs and other technical posts in a gender-neutral way.

That is because Amazon’s computer models were trained to vet applicants by observing patterns in resumes submitted to the company over a 10-year period. Most came from men, a reflection of male dominance across the tech industry.

In effect, Amazon’s system taught itself that male candidates were preferable. It penalized resumes that included the word “women’s,” as in “women’s chess club captain.” And it downgraded graduates of two all-women’s colleges, according to people familiar with the matter. They did not specify the names of the schools.

Amazon edited the programs to make them neutral to these particular terms. But that was no guarantee that the machines would not devise other ways of sorting candidates that could prove discriminatory, the people said.

The [Seattle](https://www.cnbc.com/seattle/) company ultimately disbanded the team by the start of last year because executives lost hope for the project, according to the people, who spoke on condition of anonymity. Amazon’s recruiters looked at the recommendations generated by the tool when searching for new hires, but never relied solely on those rankings, they said.

Amazon declined to comment on the recruiting engine or its challenges, but the company says it is committed to workplace diversity and equality.

The company’s experiment, which Reuters is first to report, offers a case study in the limitations of machine learning. It also serves as a lesson to the growing list of large companies including [Hilton Worldwide](https://www.cnbc.com/quotes/?symbol=HLT) and [Goldman Sachs](https://www.cnbc.com/quotes/?symbol=GS) that are looking to automate portions of the hiring process.

Some 55 percent of U.S. human resources managers said artificial intelligence, or AI, would be a regular part of their work within the next five years, according to a 2017 survey by talent software firm CareerBuilder.

Employers have long dreamed of harnessing technology to widen the hiring net and reduce reliance on subjective opinions of human recruiters. But computer scientists such as Nihar Shah, who teaches machine learning at Carnegie Mellon University, say there is still much work to do.

“How to ensure that the algorithm is fair, how to make sure the algorithm is really interpretable and explainable — that’s still quite far off,” he said.

**Masculine language**

Amazon’s experiment began at a pivotal moment for the world’s largest online retailer. Machine learning was gaining traction in the technology world, thanks to a surge in low-cost computing power. And Amazon’s Human Resources department was about to embark on a hiring spree: Since June 2015, the company’s global headcount has more than tripled to 575,700 workers, regulatory filings show.

So it set up a team in Amazon’s Edinburgh engineering hub that grew to around a dozen people. Their goal was to develop AI that could rapidly crawl the web and spot candidates worth recruiting, the people familiar with the matter said.

The group created 500 computer models focused on specific job functions and locations. They taught each to recognize some 50,000 terms that showed up on past candidates’ resumes. The algorithms learned to assign little significance to skills that were common across IT applicants, such as the ability to write various computer codes, the people said.

Instead, the technology favored candidates who described themselves using verbs more commonly found on male engineers’ resumes, such as “executed” and “captured,” one person said.

Gender bias was not the only issue. Problems with the data that underpinned the models’ judgments meant that unqualified candidates were often recommended for all manner of jobs, the people said. With the technology returning results almost at random, Amazon shut down the project, they said.

**The problem, or the cure?**

Other companies are forging ahead, underscoring the eagerness of employers to harness AI for hiring.

Kevin Parker, chief executive of HireVue, a startup near Salt Lake City, said automation is helping firms look beyond the same recruiting networks upon which they have long relied. His firm analyzes candidates’ speech and facial expressions in video interviews to reduce reliance on resumes.

“You weren’t going back to the same old places; you weren’t going back to just Ivy League schools,” Parker said. His company’s customers include [Unilever](https://www.cnbc.com/quotes/?symbol=ULVR-GB" \o "" \t ")and Hilton.

Goldman Sachs has created its own resume analysis tool that tries to match candidates with the division where they would be the “best fit,” the company said.

[Microsoft’s](https://www.cnbc.com/quotes/?symbol=MSFT) LinkedIn, the world’s largest professional network, has gone further. It offers employers algorithmic rankings of candidates based on their fit for job postings on its site.

Still, John Jersin, vice president of LinkedIn Talent Solutions, said the service is not a replacement for traditional recruiters.

“I certainly would not trust any AI system today to make a hiring decision on its own,” he said. “The technology is just not ready yet.”

Some activists say they are concerned about transparency in AI. The American Civil Liberties Union is currently challenging a law that allows criminal prosecution of researchers and journalists who test hiring websites’ algorithms for discrimination.

“We are increasingly focusing on algorithmic fairness as an issue,” said Rachel Goodman, a staff attorney with the Racial Justice Program at the ACLU.

Still, Goodman and other critics of AI acknowledged it could be exceedingly difficult to sue an employer over automated hiring: Job candidates might never know it was being used.

As for Amazon, the company managed to salvage some of what it learned from its failed AI experiment. It now uses a “much-watered down version” of the recruiting engine to help with some rudimentary chores, including culling duplicate candidate profiles from databases, one of the people familiar with the project said.

Another said a new team in Edinburgh has been formed to give automated employment screening another try, this time with a focus on diversity.

# Surveillance Capitalism

<https://www.youtube.com/watch?v=uJwf6oLvc2Q>

<https://www.theguardian.com/technology/2019/jan/20/shoshana-zuboff-age-of-surveillance-capitalism-google-facebook>