



## RETAIL

OCTOBER 10, 2018 / 7:04 PM / UPDATED 4 YEARS AGO

## Amazon scraps secret AI recruiting tool that showed bias against women

By Jeffrey Dastin



SAN FRANCISCO (Reuters) - Amazon.com Inc's [AMZN.O](#) 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.

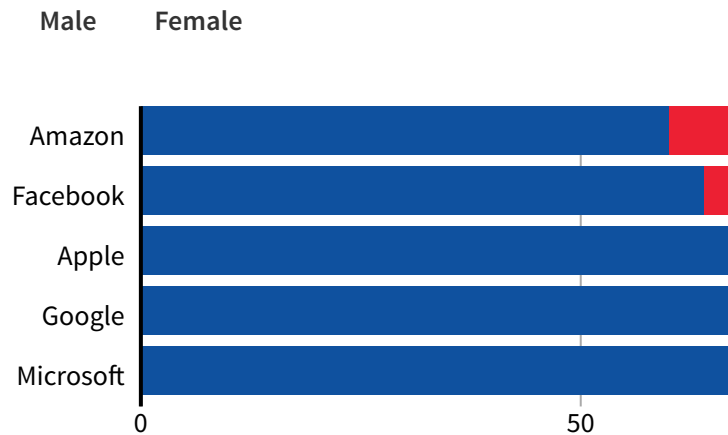
That is because Amazon's computer models were trained to vet applicants by observing patterns in resumes submitted to the company over a 10-year period. Most came from men, a

reflection of male dominance across the tech industry.

## Dominated by men

Top U.S. tech companies have yet to close the gender gap where men far outnumber women. Amazon's experiment with the word "women's" until the company discovered the

### GLOBAL HEADCOUNT



### EMPLOYEES IN TECHNICAL ROLES



Note: Amazon does not disclose the gender breakdown of its technical employees.

Source: Latest data available from the companies, since 2017.

By Han Huang | REUTERS GRAPHICS

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 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 technology's challenges, but said the tool "was never used by Amazon recruiters to evaluate candidates." The company did not elaborate further. It did not dispute that recruiters looked at the recommendations generated by the recruiting engine.

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 Holdings Inc [HLT.N](#) and Goldman Sachs Group Inc [GS.N](#) 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.

Slideshow ( 6 images )

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 PLC [ULVR.L](#) and Hilton.

Slideshow ( 6 images )

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 Corp’s [MSFT.O](#) 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.

Reporting By Jeffrey Dastin in San Francisco; Editing by Jonathan Weber and Marla Dickerson

*Our Standards: [The Thomson Reuters Trust Principles.](#)*

---

[Apps](#) [Newsletters](#) [Advertise with Us](#) [Advertising Guidelines](#) [Cookies](#) [Terms of Use](#) [Privacy](#)  
[Do Not Sell My Personal Information](#)



All quotes delayed a minimum of 15 minutes. See [here](#) for a complete list of exchanges and delays.

