

AI's lack of transparency triggers a debate over ethics



Alexa may be up to more tricks than just unnerving us with her creepy [laugh](#).

The artificial intelligence technology that drives Amazon's digital assistant and others like it is being used in an ever-growing list of applications. Some, like a Netflix recommendation on what to binge-watch next, are relatively benign. Others, however, hold high stakes: sizing up job applicants, reviewing creditworthiness, even making risk assessments in court sentencing decisions.

And that has generated controversy, because it is not at all clear that AI is free of the prejudices that its human creators are heir to, writes SAGE Business Researcher freelance correspondent Hannah H. Kim in her report on [AI and Ethics](#). When AI is taken out of the lab and put to real-world uses, such as facial recognition systems, there is growing evidence of bias, Kim writes.

For example, a study earlier this year by MIT and Microsoft researchers found that three commercial systems designed to identify gender through facial recognition had much higher error rates for darker-skinned women than for lighter-skinned men – 34.7 percent for the former and 0.8 percent for the latter. The problem, the researchers concluded, was unintentional but systemic: the data sets used to “train” the systems were heavily weighted toward lighter-skinned examples, so the software simply had far less opportunity to learn to recognise differences among darker-skinned faces. IBM, which makes one of the systems, says it's now using different training data.

Two years ago, the nonprofit news organization ProPublica examined an algorithmic AI system used in courts around the United States to assess the likelihood that a defendant would commit future crimes. The system was more likely to label white defendants as low-risk than black ones, ProPublica concluded. The maker of the system, Northpointe, disputed the news organization's conclusions.

AI systems used in initial screening of job candidates often rely on historical data about past applicants who have gone on to become high- or low-performing employees, Kim writes. But because such data are based on manager ratings, “what the machine learning model is learning to predict is not who is going to perform well at the job but ... managers' evaluations of these people,” Solon Barocas, an assistant professor of information science at Cornell University, told Business Researcher. If those evaluations contained bias, so will the AI systems' assessments.

One of the difficulties in wringing out such bias is that AI systems are “black boxes” that lack transparency in how they make decisions. The algorithms that power the systems are often so complex that even their developers don't fully understand how they function, according to a 2016 report by a White House inter-agency working group on AI.

Since that report, researchers have increased their efforts to penetrate AI's opacity and understand how bias can be introduced. But they have achieved only limited success thus far, and some express concern that lack of transparency may erode public acceptance of AI. "Imagine that you're rejected from a job, and the answer is, 'Well, we know that the model performs well, but we don't really understand why you were rejected,'" Barocas told Kim. "Or, you apply for credit, and the rejection decision's explanation is, 'the model said so.'"

As businesses explore new uses for AI and increase their investment in it, they are also becoming more aware of such ethical concerns. Two years ago, companies including Microsoft, Apple, Google, IBM, Facebook, Amazon and the British software firm DeepMind came together to create a consortium called the Partnership on AI to Benefit People and Society. Its mission is to develop a set of industry best practices, including how to make AI systems more accountable.

Some industry executives say this is an idea whose time has come. "AI ethics is the new 'Green,'" wrote Rana el Kaliouby, CEO of the emotion measurement technology company Affectiva. "On the path to ubiquity of AI, there will be many ethics-related decisions that we, as AI leaders, need to make. We have a responsibility to drive those decisions, not only because it is the right thing to do for society but because it is the smart business decision."



Notes:

- *This blog post is based on the report [AI and Ethics](#), on [SAGE Business Researcher](#), by Hanna H. Kim.*
- *The post gives the views of its authors, not the position of LSE Business Review or the London School of Economics.*
- *Featured image credit: [Photo](#) by Andres Urena from Pexels*
- *When you leave a comment, you're agreeing to our [Comment Policy](#).*



Ken Fireman is managing editor for SAGE Business Researcher, which delivers deep dives on contemporary business issues to students and faculty three times a month. He was previously a senior editor for economics and politics at Bloomberg News and a White House correspondent, national political reporter and Moscow bureau chief for Newsday. @kfireman1