

Great promise but potential for peril

Ethical concerns mount as AI takes bigger decision-making role in more industries

Christina Pazzanese
Harvard Staff Writer
October 26, 2020

Second in a four-part series that taps the expertise of the Harvard community to examine the promise and potential pitfalls of the rising age of artificial intelligence and machine learning, and how to humanize them.

For decades, artificial intelligence, or AI, was the engine of high-level STEM research. Most consumers became aware of the technology's power and potential through internet platforms like Google and Facebook, and retailer Amazon. Today, AI is essential across a vast array of industries, including health care, banking, retail, and manufacturing.

But its game-changing promise to do things like improve efficiency, bring down costs, and accelerate research and development has been tempered of late with worries that these complex, opaque systems may do more societal harm than economic good. With virtually no U.S. government oversight, private companies use AI software to make determinations about health and medicine, employment, creditworthiness, and even criminal justice without having to answer for how they're ensuring that programs aren't encoded, consciously or unconsciously, with structural biases.

Its growing appeal and utility are undeniable. Worldwide business spending on AI is expected to hit \$50 billion this year and \$110 billion annually by 2024, even after the global economic slump caused by the COVID-19 pandemic, according to a [forecast](#) released in August by technology research firm IDC. Retail and banking industries spent the most this year, at more than \$5 billion each. The company expects the media industry and federal and central governments will invest most heavily between 2018 and 2023 and predicts that AI will be "the disrupting influence changing entire industries over the next decade."

"Virtually every big company now has multiple AI systems and counts the deployment of AI as integral to their strategy," said [Joseph Fuller](#), professor of management practice at Harvard Business School, who co-leads [Managing the Future of Work](#), a research project that studies, in part, the development and

implementation of AI, including machine learning, robotics, sensors, and industrial automation, in business and the work world.

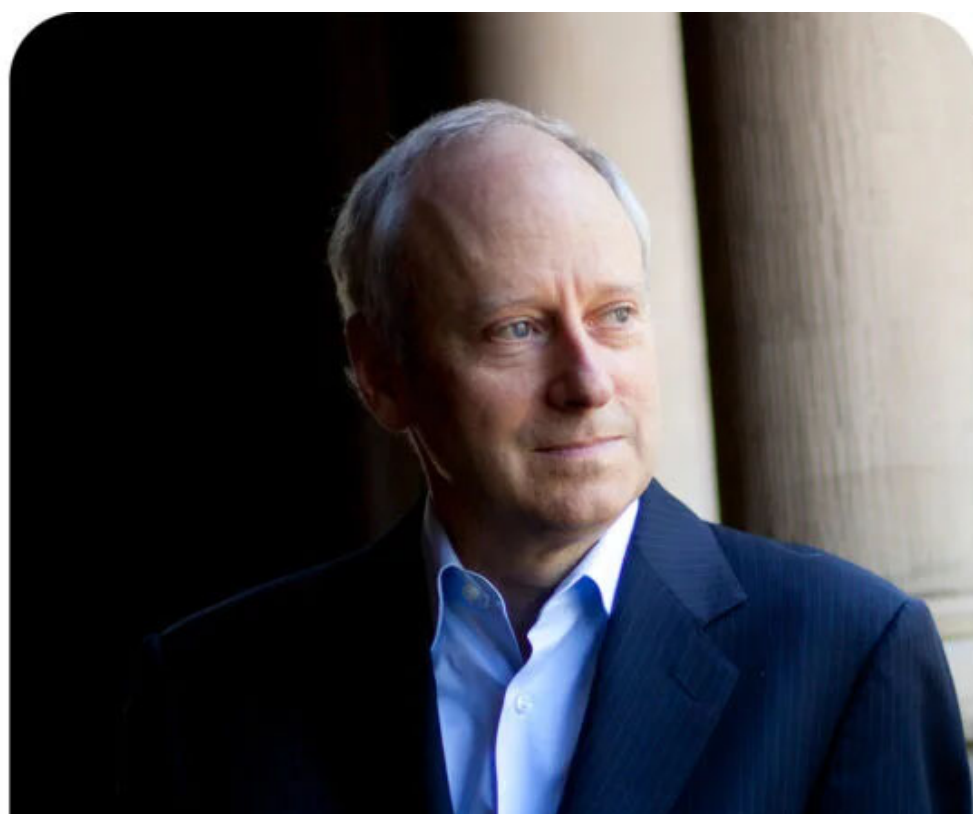
Early on, it was popularly assumed that the future of AI would involve the automation of simple repetitive tasks requiring low-level decision-making. But AI has rapidly grown in sophistication, owing to more powerful computers and the compilation of huge data sets. One branch, machine learning, notable for its ability to sort and analyze massive amounts of data and to learn over time, has transformed countless fields, including education.

Firms now use AI to manage sourcing of materials and products from suppliers and to integrate vast troves of information to aid in strategic decision-making, and because of its capacity to process data so quickly, AI tools are helping to minimize time in the pricey trial-and-error of product development — a critical advance for an industry like pharmaceuticals, where it costs \$1 billion to bring a new pill to market, Fuller said.

Health care experts see many possible uses for AI, including with billing and processing necessary paperwork. And medical professionals expect that the biggest, most immediate impact will be in analysis of data, imaging, and diagnosis. Imagine, they say, having the ability to bring all of the medical knowledge available on a disease to any given treatment decision.

In employment, AI software culls and processes resumes and analyzes job interviewees' voice and facial expressions in hiring and driving the growth of what's known as "hybrid" jobs. Rather than replacing employees, AI takes on important technical tasks of their work, like routing for package delivery trucks, which potentially frees workers to focus on other responsibilities, making them more productive and therefore more valuable to employers.

"It's allowing them to do more stuff better, or to make fewer errors, or to capture their expertise and disseminate it more effectively in the organization," said Fuller, who has studied the effects and attitudes of workers who have lost or are likeliest to lose their jobs to AI.





Stephanie Mitchell/Harvard file photo

“Can smart machines outthink us, or are certain elements of human judgment indispensable in deciding some of the most important things in life?”

— Michael Sandel, political philosopher and Anne T. and Robert M. Bass Professor of Government

Though automation is here to stay, the elimination of entire job categories, like highway toll-takers who were replaced by sensors because of AI’s proliferation, is not likely, according to Fuller.

“What we’re going to see is jobs that require human interaction, empathy, that require applying judgment to what the machine is creating [will] have robustness,” he said.

While big business already has a huge head start, small businesses could also potentially be transformed by AI, says Karen Mills ’75, M.B.A. ’77, who ran the U.S. Small Business Administration from 2009 to 2013. With half the country employed by small businesses before the COVID-19 pandemic, that could have major implications for the national economy over the long haul.

Rather than hamper small businesses, the technology could give their owners detailed new insights into sales trends, cash flow, ordering, and other important financial information in real time so they can better understand how the business is doing and where problem areas might loom without having to hire anyone, become a financial expert, or spend hours laboring over the books every week, Mills said.

One area where AI could “completely change the game” is lending, where access to capital is difficult in part because banks often struggle to get an accurate picture of a small business’s viability and creditworthiness.

“It’s much harder to look inside a business operation and know what’s going on” than it is to assess an individual, she said.

Information opacity makes the lending process laborious and expensive for both would-be borrowers and lenders, and applications are designed to analyze larger companies or those who’ve already borrowed, a built-in disadvantage for certain types of businesses and for historically underserved borrowers, like women and minority business owners, said Mills, a senior fellow at HBS.

But with AI-powered software pulling information from a business’s bank account, taxes, and online bookkeeping records and comparing it with data from thousands of similar businesses, even small community banks will be able to make informed assessments in minutes, without the agony of paperwork and delays, and, like blind auditions for musicians, without fear that any inequity crept into the decision-making.

“All of that goes away,” she said.

A VENEER OF OBJECTIVITY

Not everyone sees blue skies on the horizon, however. Many worry whether the coming age of AI will bring new, faster, and frictionless ways to discriminate and divide at scale.

“Part of the appeal of algorithmic decision-making is that it seems to offer an objective way of overcoming human subjectivity, bias, and prejudice,” said political philosopher Michael Sandel, Anne T. and Robert M. Bass Professor of Government. “But we are discovering that many of the algorithms that decide who should get parole, for example, or who should be presented with employment opportunities or housing ... replicate and embed the biases that already exist in our society.”





Jon Chase/Harvard file photo

“If we’re not thoughtful and careful, we’re going to end up with redlining again.”

– Karen Mills, senior fellow at the Business School and head of the U.S. Small Business Administration from 2009 to 2013

AI presents three major areas of ethical concern for society: privacy and surveillance, bias and discrimination, and perhaps the deepest, most difficult philosophical question of the era, the role of human judgment, said Sandel, who teaches a course in the moral, social, and political implications of new technologies.

“Debates about privacy safeguards and about how to overcome bias in algorithmic decision-making in sentencing, parole, and employment practices are by now familiar,” said Sandel, referring to conscious and unconscious prejudices of program developers and those built into datasets used to train the software. “But we’ve not yet wrapped our minds around the hardest question: Can smart machines outthink us, or are certain elements of human judgment indispensable in deciding some of the most important things in life?”

Panic over AI suddenly injecting bias into everyday life *en masse* is overstated, says Fuller. First, the business world and the workplace, rife with human decision-making, have always been riddled with “all sorts” of biases that prevent people from making deals or landing contracts and jobs.

When calibrated carefully and deployed thoughtfully, resume-screening software allows a wider pool of applicants to be considered than could be done otherwise, and should minimize the potential for favoritism

that comes with human gatekeepers, Fuller said.

Sandel disagrees. “AI not only replicates human biases, it confers on these biases a kind of scientific credibility. It makes it seem that these predictions and judgments have an objective status,” he said.

In the world of lending, algorithm-driven decisions do have a potential “dark side,” Mills said. As machines learn from data sets they’re fed, chances are “pretty high” they may replicate many of the banking industry’s past failings that resulted in systematic disparate treatment of African Americans and other marginalized consumers.

“If we’re not thoughtful and careful, we’re going to end up with redlining again,” she said.

A highly regulated industry, banks are legally on the hook if the algorithms they use to evaluate loan applications end up inappropriately discriminating against classes of consumers, so those “at the top levels” in the field are “very focused” right now on this issue, said Mills, who closely studies the rapid changes in financial technology, or “fintech.”

“They really don’t want to discriminate. They want to get access to capital to the most creditworthy borrowers,” she said. “That’s good business for them, too.”

OVERSIGHT OVERWHELMED

Given its power and expected ubiquity, some argue that the use of AI should be tightly regulated. But there’s little consensus on how that should be done and who should make the rules.

Thus far, companies that develop or use AI systems largely self-police, relying on existing laws and market forces, like negative reactions from consumers and shareholders or the demands of highly-prized AI technical talent to keep them in line.

“There’s no businessperson on the planet at an enterprise of any size that isn’t concerned about this and trying to reflect on what’s going to be politically, legally, regulatorily, [or] ethically acceptable,” said Fuller.

Firms already consider their own potential liability from misuse before a product launch, but it’s not realistic to expect companies to anticipate and prevent every possible unintended consequence of their product, he said.

Few think the federal government is up to the job, or will ever be.

“The regulatory bodies are not equipped with the expertise in artificial intelligence to engage in [oversight] without some real focus and investment,” said Fuller, noting the rapid rate of technological change means even the most informed legislators can’t keep pace. Requiring every new product using AI to be prescreened for potential social harms is not only impractical, but would create a huge drag on innovation.





Rose Lincoln/Harvard Staff Photographer

“I wouldn’t have a central AI group that has a division that does cars, I would have the car people have a division of people who are really good at AI.”

— Jason Furman, a professor of the practice of economic policy at the Kennedy School and a former top economic adviser to President Barack Obama

Jason Furman, a professor of the practice of economic policy at Harvard Kennedy School, agrees that government regulators need “a much better technical understanding of artificial intelligence to do that job well,” but says they could do it.

Existing bodies like the National Highway Transportation Safety Association, which oversees vehicle safety, for example, could handle potential AI issues in autonomous vehicles rather than a single watchdog agency, he said.

“I wouldn’t have a central AI group that has a division that does cars, I would have the car people have a division of people who are really good at AI,” said Furman, a former top economic adviser to President Barack Obama.

Though keeping AI regulation within industries does leave open the possibility of co-opted enforcement, Furman said industry-specific panels would be far more knowledgeable about the overarching technology of which AI is simply one piece, making for more thorough oversight.

While the European Union already has rigorous data-privacy laws and the European Commission is considering a formal regulatory framework for ethical use of AI, the U.S. government has historically been late when it comes to tech regulation.

“I think we should’ve started three decades ago, but better late than never,” said Furman, who thinks there needs to be a “greater sense of urgency” to make lawmakers act.

Business leaders “can’t have it both ways,” refusing responsibility for AI’s harmful consequences while also fighting government oversight, Sandel maintains.

“The problem is these big tech companies are neither self-regulating, nor subject to adequate government regulation. I think there needs to be more of both,” he said, later adding: “We can’t assume that market forces by themselves will sort it out. That’s a mistake, as we’ve seen with Facebook and other tech giants.”

Last fall, Sandel taught “Tech Ethics,” a popular new Gen Ed course with Doug Melton, co-director of Harvard’s Stem Cell Institute. As in his legendary “Justice” course, students consider and debate the big questions about new technologies, everything from gene editing and robots to privacy and surveillance.

“Companies have to think seriously about the ethical dimensions of what they’re doing and we, as democratic citizens, have to educate ourselves about tech and its social and ethical implications — not only to decide what the regulations should be, but also to decide what role we want big tech and social media to play in our lives,” said Sandel.

Doing that will require a major educational intervention, both at Harvard and in higher education more broadly, he said.

“We have to enable all students to learn enough about tech and about the ethical implications of new technologies so that when they are running companies or when they are acting as democratic citizens, they will be able to ensure that technology serves human purposes rather than undermines a decent civic life.”

Next: The AI revolution in medicine may lift personalized treatment, fill gaps in access to care, and cut red tape. Yet risks abound.