Inquiring Minds topic – 14 December 2018

John Moore, Moderator

"The Role of Corporations in Addressing AI's Ethical Dilemmas"

Darrell M. West

This a condensation of a much longer article that appeared on September 13 in the Brookings Institution's online edition. It uses five cases to illustrate the dilemmas posed by the development of artificial intelligence. Each of the cases deserves attention, but it is hoped that more general points about the ethics of AI applications will emerge in our discussion. To help, here are some broad questions to bear in mind.

- **Whose responsibility is it to pass judgment on ethical issues involving AI? Government? Business?**
- **Does AI raise any new or unique ethical questions? For example, does the potential of AI for military applications raise issues different from those connected to atomic energy?**
- **Artificial intelligence is developing rapidly, so rapidly that it is hard to stay abreast of change, fast enough that the machinery of government may not be able to keep pace. How, then, can society be confident that its capabilities will not be misused?**
- **As the Chinese development of the social credit system shows, AI can be a powerful tool for governments that wish to exert control over their citizens. Could this become a problem in democratic societies?**

POLITICAL POLARIZATION AND DUAL-USE TECHNOLOGIES

In looking across AI activities, there are several applications that have raised ethical concerns. It is one thing to support general goals, such as fairness and accountability, but another to apply those concepts in particular domains and under specific political conditions. One cannot isolate ethics discussions from the broader political climate in which technology is being deployed.

The current polarization around politics and policymaking complicates the tasks facing decisionmakers. Republicans and Democrats have very different views of U.S. officeholders, policy

options, and political developments. Ethical issues that might not be very controversial during a time of normal politics become much more divisive when people don't like or trust the officials making the decisions.

In addition, running through many ethical dilemmas is the problem of dual-use technologies. There are many algorithms and software applications that can be used for good or ill. Facial recognition can be deployed to find lost children or facilitate widespread civilian surveillance.[13] It is not the technology so much that dictates the moral dilemma as the human use case involved with the application. The very same algorithm can serve a variety of purposes, which makes the ethics of decision making very difficult.[14]

For this reason, companies [and society more generally] have to consider not just the ethical aspects of emerging technologies, but also their possible use cases. Indeed, the latter represents an interesting opportunity to explore AI ethics because it illustrates concrete aspects of ethical dilemmas. Having in-depth knowledge of those issues is important for AI development.

DILEMMA ONE: WEAPONS DEVELOPMENT AND MILITARY APPLICATIONS

One topic that has attracted considerable attention involves AI applications devoted to war or military activities. As technology innovation has accelerated, there have been discussions regarding whether AI should be used in war-related activities. In its code of ethics, for example, Google wrote that it will not design or deploy AI in: "weapons or other technologies designed to cause or directly facilitate injury to people; in technologies that gather or use information for surveillance violating internationally accepted norms; or technologies for any purpose that contravene widely accepted principles of international law and human rights." To clarify the situation, its document also added, "[For any AI applications] where there is a material risk of harm, we will proceed only where we believe that the benefits substantially outweigh the risks, and will incorporate appropriate safety constraints."[15]

Of course, many other firms have not adopted this position. For example, Palantir has garnered at least $1.2 billion in federal contracts since 2009 through products popular with defense, intelligence agencies, homeland security, and law enforcement. One of its primary applications known as Gotham imports "large reams of structured data (like spreadsheets) and unstructured data (like images) into one centralized database, where all of the information can be visualized and analyzed in one workspace."[16] The goal is to use technology to make military applications more efficient and effective, and help defense planners achieve their objectives in the field.

Indeed, military leaders long have recognized the need to upgrade capabilities and incorporate the latest advances in their arsenals. The U.S. Department of Defense has set up a Joint Artificial Intelligence Center designed to improve "large-scale AI projects."[17] Its plan is to work with private

companies and university researchers to make sure America takes advantage of the latest products for defense purposes.

This is consistent with the urgings of Brookings President John Allen and business executive Amir Husain. They argue the world is moving towards "hyperwar," in which advanced capabilities will combine into rapid-style engagements based on physical and digital encounters.[18] As such, it is important for the United States to have the means to defend itself against possible AI-based attacks from adversaries.

Many commentators have noted that countries, such as Russia, China, Iran, and North Korea, have AI capabilities and are not refraining from deployment of high-tech tools. During a period of considerable international turbulence and global threats, America has to be careful not to engage in unilateral disarmament when possible adversaries are moving full-speed ahead. Disputes over AI deployment demonstrate not all agree on an AI prohibition for national security purposes.

The American public understands this point. In an August 2018 survey undertaken by Brookings researchers, 30 percent of respondents believed the United States should develop AI technologies for warfare, 39 percent did not, and 31 percent were undecided. However, when told that adversaries already are developing AI for war-related purposes, 45 percent thought America should develop these kinds of weapons, 25 percent did not, and 30 percent were undecided.[19]

There are substantial demographic differences in these attitudes. Men (51 percent) were much more likely than women (39 percent) to support AI for warfare if adversaries develop these kinds of weapons. The same was true for senior citizens (53 percent) compared to those aged 18 to 34 (38 percent).

DILEMMA TWO: LAW AND BORDER ENFORCEMENT

In the domestic policy area, there are similar concerns regarding the militarization of policing practices and shootings of unarmed black men in communities across the U.S. Those tendencies have led some to decry AI applications in law enforcement. Critics worry that emerging technologies, such as facial recognition software, unfairly target minorities and lead to biased or discriminatory enforcement, sometimes with tragic consequences.

Some business leaders have been quite outspoken on this topic. For example, Brian Brackeen, the chief executive officer of facial recognition firm Kairos, argues that its usage "opens the door for gross misconduct by the morally corrupt." He discusses the history of law enforcement against American minorities and concludes, "There is no place in America for facial recognition that supports false arrests and murder." Speaking on behalf of his company, Brackeen says his firm will not work with government agencies and says, "Any company in this space that willingly hands this

software over to a government, be it America or another nation's, is willfully endangering people's lives."[20]

The same logic applies to border enforcement under the current administration. With President Donald Trump's crackdown on undocumented arrivals, employees at some firms have complained about contracts with the Immigration and Customs Enforcement agency that is charged with enforcing administration decisions.[21] They object to Trump's policies and argue technology firms should not enable that crackdown by providing technologies for border enforcement. McKinsey & Company already have announced it no longer will work with Immigration and Customs Enforcement and Customs and Border Protection due to employee objections to enforcement actions at those agencies.[22]

DILEMMA THREE: GOVERNMENT SURVEILLANCE

Government surveillance is a challenge in many places. A number of countries have turned toward authoritarianism in recent years. They have shut down the internet, attacked dissidents, imprisoned reporters or NGO advocates, and attacked judges. All of these activities have fueled concerns regarding government use of technology to surveil or imprison innocent people.

As a result, some companies have disavowed any interest in selling to government agencies. As an illustration, CEO Rana el Kaliouby of Affectiva, an AI firm that works on image recognition, has turned down such opportunities. "We're not interested in applications where you're spying on people," he announced. This includes security agencies, airport authorities, or lie detection contracts.[23]

In addition, Microsoft has argued facial recognition is to "be left up to tech companies." Company President Brad Smith says this software "raises issues that go to the heart of fundamental human rights protections like privacy and freedom of expression." As a result, he supports "a government initiative to regulate the proper use of facial recognition technology, informed first by a bipartisan and expert commission."[24]

Other companies, however, have not taken this stance. Amazon sells its Rekognition facial recognition software to police agencies and other kinds of government units, even though some of its employees object to the practice.[25] It has the view that government authorities should have access to the latest technologies. But the firm has announced that "it will suspend … customer's right to use … services [like Rekognition] if it determines those services are being abused."[26]

In China, there is growing use of facial recognition combined with video cameras and AI to keep track of its own population. There, law enforcement scans people at train stations to find wanted people or identifies jaywalkers who cross the street illegally. It is estimated that the country has deployed 200 million video cameras, which makes possible surveillance on an unprecedented

scale.[27] When combined with AI analysis that matches images with personal identities, the capacity for in-depth population control is enormous.
In his analysis of the ethics of facial recognition software, Brookings scholar William Galston points out there should be "a reasonable expectation of anonymity." Government authorities should not deploy such technology unless there is "a justification weighty enough to override the presumption against doing so," and that "this process should be regulated by law … [and] the equivalent of a search warrant."[28] In his view, having clear legal standards is vital in order to prevent likely abuses.

DILEMMA FOUR: RACIAL BIAS

There is considerable evidence of racial biases in facial recognition software. Some systems have "misidentified darker-skinned women as often as 35 percent of the time and darker-skinned men 12 percent of the time," much higher than the rates for Caucasians.[29]

Most systems operate by comparing a person's face to a range of images in a large database. As pointed out by Joy Buolamwini of the Algorithmic Justice League, "If your facial recognition data contains mostly Caucasian faces, that's what your program will learn to recognize."[30] Unless the databases have access to diverse data, these programs perform poorly when attempting to recognize African-American or Asian-American features.

Many historical data sets reflect traditional values, which may or may not represent the preferences wanted in a current system. As Buolamwini notes, such an approach risks repeating inequities of the past.

The rise of automation and the increased reliance on algorithms for high-stakes decisions such as whether someone gets insurance or not, your likelihood to default on a loan or somebody's risk of recidivism means this is something that needs to be addressed. Even admissions decisions are increasingly automated—what school our children go to and what opportunities they have. We don't have to bring the structural inequalities of the past into the future we create.

This is one of the reasons why it is important to increase data openness so AI developers have access to large data sets for training purposes. They need unbiased information in order to instruct AI systems properly on how to recognize certain patterns and make reasonable decisions. Governments can help in this regard by promoting greater access to their information.[31]They have some of the largest data sets, and this information can be a valuable resource for training AI and overcoming past problems.

In addition, in sensitive areas, such as criminal justice—where inaccuracies can lead to higher incarceration rates—there need to be minimum standards of accuracy for facial recognition software to be deployed. Systems should certify what their rates are so officials understand what possible biases come with AI deployment. Jennifer Lynch of the Electronic Frontier Foundation

argues that “an inaccurate system will implicate people for crimes they didn’t commit and shift the burden to innocent defendants to show they are not who the system says they are.”[32]

DILEMMA FIVE: SOCIAL CREDIT SYSTEMS

China is expanding its use of social credit systems for daily life. It compiles data on people’s social media activities, personal infractions, and paying taxes on time, and uses the resulting score to rate people for credit-worthiness, travel, school enrollment, and government positions.[33] Those with high scores are accorded special discounts and privileges, while those who fare more poorly can be banned from travel, refused enrollment at favored schools, or restricted from government employment.

The problem with these systems depends on their opacity. As noted by Jack Karsten and me in a blog post, “It is not clear what factors affect someone’s score, and so those with a low score may face exclusion without knowing why.”[34] In addition, given inequitable access to activities that promote higher scores, such systems can increase disparities based on socio-economic background, ethnic category, or education level. Authoritarian regimes may turn to AI to support their interest in controlling the population.