

BROOKINGS

RESEARCH

How artificial intelligence is transforming the world

Darrell M. West and John R. Allen

April 24, 2018

Artificial intelligence (AI) is a wide-ranging tool that enables people to rethink how we integrate information, analyze data, and use the resulting insights to improve decision making—and already it is transforming every walk of life. In this report, Darrell West and John Allen discuss AI's application across a variety of sectors, address issues in its development, and offer recommendations for getting the most out of AI while still protecting important human values.

Table of Contents

- I. Qualities of artificial intelligence
- II. Applications in diverse sectors
- III. Policy, regulatory, and ethical issues
- IV. Recommendations
- V. Conclusion

Sections

 Chapter I

 Chapter II

 Chapter III

 Chapter IV

 Chapter V

Contact

Jessica Harris

jbharris@brookings.edu

202.238.3507

 Print

Most people are not very familiar with the concept of artificial intelligence (AI). As an illustration, when 1,500 senior business leaders in the United States in 2017 were asked about AI, only 17 percent said they were familiar with it.¹ ([#_edn1](#)) A number of them were not sure what it was or how it would affect their particular companies. They understood there was considerable potential for altering business processes, but were not clear how AI could be deployed within their own organizations.

Despite its widespread lack of familiarity, AI is a technology that is transforming every walk of life. It is a wide-ranging tool that enables people to rethink how we integrate information, analyze data, and use the resulting insights to improve decisionmaking. Our hope through this comprehensive

overview is to explain AI to an audience of policymakers, opinion leaders, and interested observers, and demonstrate how AI already is altering the world and raising important questions for society, the economy, and governance.

In this paper, we discuss novel applications in finance, national security, health care, criminal justice, transportation, and smart cities, and address issues such as data access problems, algorithmic bias, AI ethics and transparency, and legal liability for AI decisions. We contrast the regulatory approaches of the U.S. and European Union, and close by making a number of recommendations for getting the most out of AI while still protecting important human values.²

In order to maximize AI benefits, we recommend nine steps for going forward:

- Encourage greater data access for researchers without compromising users' personal privacy,
- invest more government funding in unclassified AI research,
- promote new models of digital education and AI workforce development so employees have the skills needed in the 21st-century economy,
- create a federal AI advisory committee to make policy recommendations,
- engage with state and local officials so they enact effective policies,
- regulate broad AI principles rather than specific algorithms,
- take bias complaints seriously so AI does not replicate historic injustice, unfairness, or discrimination in data or algorithms,
- maintain mechanisms for human oversight and control, and

- penalize malicious AI behavior and promote cybersecurity.
-

CHAPTER I

Qualities of artificial intelligence

Although there is no uniformly agreed upon definition, AI generally is thought to refer to “machines that respond to stimulation consistent with traditional responses from humans, given the human capacity for contemplation, judgment and intention.”³ According to researchers Shubhendu and Vijay, these software systems “make decisions which normally require [a] human level of expertise” and help people anticipate problems or deal with issues as they come up.⁴ As such, they operate in an intentional, intelligent, and adaptive manner.

Intentionality

Artificial intelligence algorithms are designed to make decisions, often using real-time data. They are unlike passive machines that are capable only of mechanical or predetermined responses. Using sensors, digital data, or remote inputs, they combine information from a variety of different sources, analyze the material instantly, and act on the insights derived from those data. With massive improvements in storage systems, processing speeds, and analytic techniques, they are capable of tremendous sophistication in analysis and decisionmaking.

Artificial intelligence is already altering the world and raising important

questions for society, the economy, and governance.

Intelligence

AI generally is undertaken in conjunction with machine learning and data analytics.⁵ Machine learning takes data and looks for underlying trends. If it spots something that is relevant for a practical problem, software designers can take that knowledge and use it to analyze specific issues. All that is required are data that are sufficiently robust that algorithms can discern useful patterns. Data can come in the form of digital information, satellite imagery, visual information, text, or unstructured data.

Adaptability

AI systems have the ability to learn and adapt as they make decisions. In the transportation area, for example, semi-autonomous vehicles have tools that let drivers and vehicles know about upcoming congestion, potholes, highway construction, or other possible traffic impediments. Vehicles can take advantage of the experience of other vehicles on the road, without human involvement, and the entire corpus of their achieved “experience” is immediately and fully transferable to other similarly configured vehicles. Their advanced algorithms, sensors, and cameras incorporate experience in current operations, and use dashboards and visual displays to present information in real time so human drivers are able to make sense of ongoing traffic and vehicular conditions. And in the case of fully autonomous vehicles, advanced systems can completely control the car or truck, and make all the navigational decisions.

Applications in diverse sectors

AI is not a futuristic vision, but rather something that is here today and being integrated with and deployed into a variety of sectors. This includes fields such as finance, national security, health care, criminal justice, transportation, and smart cities. There are numerous examples where AI already is making an impact on the world and augmenting human capabilities in significant ways.⁶

One of the reasons for the growing role of AI is the tremendous opportunities for economic development that it presents. A project undertaken by PriceWaterhouseCoopers estimated that “artificial intelligence technologies could increase global GDP by \$15.7 trillion, a full 14%, by 2030.”⁷ That includes advances of \$7 trillion in China, \$3.7 trillion in North America, \$1.8 trillion in Northern Europe, \$1.2 trillion for Africa and Oceania, \$0.9 trillion in the rest of Asia outside of China, \$0.7 trillion in Southern Europe, and \$0.5 trillion in Latin America. China is making rapid strides because it has set a national goal of investing \$150 billion in AI and becoming the global leader in this area by 2030.

Meanwhile, a McKinsey Global Institute study of China found that “AI-led automation can give the Chinese economy a productivity injection that would add 0.8 to 1.4 percentage points to GDP growth annually, depending on the speed of adoption.”⁸ Although its authors found that China currently lags the United States and the United Kingdom in AI deployment, the sheer size of its AI market gives that country tremendous opportunities for pilot testing and future development.

Investments in financial AI in the United States tripled between 2013 and 2014 to a total of \$12.2 billion.⁹ According to observers in that sector, “Decisions about loans are now being made by software that can take into account a variety of finely parsed data about a borrower, rather than just a credit score and a background check.”¹⁰ In addition, there are so-called robo-advisers that “create personalized investment portfolios, obviating the need for stockbrokers and financial advisers.”¹¹ These advances are designed to take the emotion out of investing and undertake decisions based on analytical considerations, and make these choices in a matter of minutes.

A prominent example of this is taking place in stock exchanges, where high-frequency trading by machines has replaced much of human decisionmaking. People submit buy and sell orders, and computers match them in the blink of an eye without human intervention. Machines can spot trading inefficiencies or market differentials on a very small scale and execute trades that make money according to investor instructions.¹² Powered in some places by advanced computing, these tools have much greater capacities for storing information because of their emphasis not on a zero or a one, but on “quantum bits” that can store multiple values in each location.¹³ That dramatically increases storage capacity and decreases processing times.

Fraud detection represents another way AI is helpful in financial systems. It is sometimes difficult to discern fraudulent activities in large organizations, but AI can identify abnormalities, outliers, or deviant cases requiring additional investigation. That helps managers find problems early in the cycle, before they reach dangerous levels.¹⁴

National security

AI plays a substantial role in national defense. Through its Project Maven, the American military is deploying AI “to sift through the massive troves of data and video captured by surveillance and then alert human analysts of patterns or when there is abnormal or suspicious activity.”¹⁵ According to Deputy Secretary of Defense Patrick Shanahan, the goal of emerging technologies in this area is “to meet our warfighters’ needs and to increase [the] speed and agility [of] technology development and procurement.”¹⁶

Artificial intelligence will accelerate the traditional process of warfare so rapidly that a new term has been coined: hyperwar.

The big data analytics associated with AI will profoundly affect intelligence analysis, as massive amounts of data are sifted in near real time—if not eventually in real time—thereby providing commanders and their staffs a level of intelligence analysis and productivity heretofore unseen. Command and control will similarly be affected as human commanders delegate certain routine, and in special circumstances, key decisions to AI platforms, reducing dramatically the time associated with the decision and subsequent action. In the end, warfare is a time competitive process, where the side able to decide the fastest and move most quickly to execution will generally prevail. Indeed, artificially intelligent intelligence systems, tied to AI-assisted command and control systems, can move decision support and decisionmaking to a speed vastly superior to the speeds of the traditional means of waging war. So fast will be this process, especially if coupled to automatic decisions to launch artificially intelligent autonomous weapons systems capable of lethal outcomes, that a new term has been coined specifically to embrace the speed at which war will be waged: hyperwar.

While the ethical and legal debate is raging over whether America will ever wage war with artificially intelligent autonomous lethal systems, the Chinese and Russians are not nearly so mired in this debate, and we should

anticipate our need to defend against these systems operating at hyperwar speeds. The challenge in the West of where to position “humans in the loop” in a hyperwar scenario will ultimately dictate the West’s capacity to be competitive in this new form of conflict.¹⁷

Just as AI will profoundly affect the speed of warfare, the proliferation of zero day or zero second cyber threats as well as polymorphic malware will challenge even the most sophisticated signature-based cyber protection. This forces significant improvement to existing cyber defenses. Increasingly, vulnerable systems are migrating, and will need to shift to a layered approach to cybersecurity with cloud-based, cognitive AI platforms. This approach moves the community toward a “thinking” defensive capability that can defend networks through constant training on known threats. This capability includes DNA-level analysis of heretofore unknown code, with the possibility of recognizing and stopping inbound malicious code by recognizing a string component of the file. This is how certain key U.S.-based systems stopped the debilitating “WannaCry” and “Petya” viruses.

Preparing for hyperwar and defending critical cyber networks must become a high priority because China, Russia, North Korea, and other countries are putting substantial resources into AI. In 2017, China’s State Council issued a plan for the country to “build a domestic industry worth almost \$150 billion” by 2030.¹⁸ As an example of the possibilities, the Chinese search firm Baidu has pioneered a facial recognition application that finds missing people. In addition, cities such as Shenzhen are providing up to \$1 million to support AI labs. That country hopes AI will provide security, combat terrorism, and improve speech recognition programs.¹⁹ The dual-use nature of many AI algorithms will mean AI research focused on one sector of society can be rapidly modified for use in the security sector as well.²⁰

Health care

AI tools are helping designers improve computational sophistication in health care. For example, Merantix is a German company that applies deep learning to medical issues. It has an application in medical imaging that “detects lymph nodes in the human body in Computer Tomography (CT) images.”²¹ According to its developers, the key is labeling the nodes and identifying small lesions or growths that could be problematic. Humans can do this, but radiologists charge \$100 per hour and may be able to carefully read only four images an hour. If there were 10,000 images, the cost of this process would be \$250,000, which is prohibitively expensive if done by humans.

What deep learning can do in this situation is train computers on data sets to learn what a normal-looking versus an irregular-appearing lymph node is. After doing that through imaging exercises and honing the accuracy of the labeling, radiological imaging specialists can apply this knowledge to actual patients and determine the extent to which someone is at risk of cancerous lymph nodes. Since only a few are likely to test positive, it is a matter of identifying the unhealthy versus healthy node.

AI has been applied to congestive heart failure as well, an illness that afflicts 10 percent of senior citizens and costs \$35 billion each year in the United States. AI tools are helpful because they “predict in advance potential challenges ahead and allocate resources to patient education, sensing, and proactive interventions that keep patients out of the hospital.”²²

Criminal justice

AI is being deployed in the criminal justice area. The city of Chicago has developed an AI-driven “Strategic Subject List” that analyzes people who have been arrested for their risk of becoming future perpetrators. It ranks 400,000 people on a scale of 0 to 500, using items such as age, criminal

activity, victimization, drug arrest records, and gang affiliation. In looking at the data, analysts found that youth is a strong predictor of violence, being a shooting victim is associated with becoming a future perpetrator, gang affiliation has little predictive value, and drug arrests are not significantly associated with future criminal activity.²³

Judicial experts claim AI programs reduce human bias in law enforcement and leads to a fairer sentencing system. R Street Institute Associate Caleb Watney writes:

Empirically grounded questions of predictive risk analysis play to the strengths of machine learning, automated reasoning and other forms of AI. One machine-learning policy simulation concluded that such programs could be used to cut crime up to 24.8 percent with no change in jailing rates, or reduce jail populations by up to 42 percent with no increase in crime rates.²⁴

However, critics worry that AI algorithms represent “a secret system to punish citizens for crimes they haven’t yet committed. The risk scores have been used numerous times to guide large-scale roundups.”²⁵ The fear is that such tools target people of color unfairly and have not helped Chicago reduce the murder wave that has plagued it in recent years.

Despite these concerns, other countries are moving ahead with rapid deployment in this area. In China, for example, companies already have “considerable resources and access to voices, faces and other biometric data in vast quantities, which would help them develop their technologies.”²⁶ New technologies make it possible to match images and voices with other types of information, and to use AI on these combined data sets to improve law enforcement and national security. Through its “Sharp Eyes”

program, Chinese law enforcement is matching video images, social media activity, online purchases, travel records, and personal identity into a “police cloud.” This integrated database enables authorities to keep track of criminals, potential law-breakers, and terrorists.²⁷ Put differently, China has become the world’s leading AI-powered surveillance state.

Transportation

Transportation represents an area where AI and machine learning are producing major innovations. Research by Cameron Kerry and Jack Karsten of the Brookings Institution has found that over \$80 billion was invested in autonomous vehicle technology between August 2014 and June 2017. Those investments include applications both for autonomous driving and the core technologies vital to that sector.²⁸

Autonomous vehicles—cars, trucks, buses, and drone delivery systems—use advanced technological capabilities. Those features include automated vehicle guidance and braking, lane-changing systems, the use of cameras and sensors for collision avoidance, the use of AI to analyze information in real time, and the use of high-performance computing and deep learning systems to adapt to new circumstances through detailed maps.²⁹

Light detection and ranging systems (LIDARs) and AI are key to navigation and collision avoidance. LIDAR systems combine light and radar instruments. They are mounted on the top of vehicles that use imaging in a 360-degree environment from a radar and light beams to measure the speed and distance of surrounding objects. Along with sensors placed on the front, sides, and back of the vehicle, these instruments provide information that keeps fast-moving cars and trucks in their own lane, helps them avoid other vehicles, applies brakes and steering when needed, and does so instantly so

as to avoid accidents.

Advanced software enables cars to learn from the experiences of other vehicles on the road and adjust their guidance systems as weather, driving, or road conditions change. This means that software is the key—not the physical car or truck itself.

Since these cameras and sensors compile a huge amount of information and need to process it instantly to avoid the car in the next lane, autonomous vehicles require high-performance computing, advanced algorithms, and deep learning systems to adapt to new scenarios. This means that software is the key, not the physical car or truck itself.³⁰ Advanced software enables cars to learn from the experiences of other vehicles on the road and adjust their guidance systems as weather, driving, or road conditions change.³¹

Ride-sharing companies are very interested in autonomous vehicles. They see advantages in terms of customer service and labor productivity. All of the major ride-sharing companies are exploring driverless cars. The surge of car-sharing and taxi services—such as Uber and Lyft in the United States, Daimler's Mytaxi and Hailo service in Great Britain, and Didi Chuxing in China—demonstrate the opportunities of this transportation option. Uber recently signed an agreement to purchase 24,000 autonomous cars from Volvo for its ride-sharing service.³²

However, the ride-sharing firm suffered a setback in March 2018 when one of its autonomous vehicles in Arizona hit and killed a pedestrian. Uber and several auto manufacturers immediately suspended testing and launched investigations into what went wrong and how the fatality could have occurred.³³ Both industry and consumers want reassurance that the technology is safe and able to deliver on its stated promises. Unless there

are persuasive answers, this accident could slow AI advancements in the transportation sector.

Smart cities

Metropolitan governments are using AI to improve urban service delivery. For example, according to Kevin Desouza, Rashmi Krishnamurthy, and Gregory Dawson:

The Cincinnati Fire Department is using data analytics to optimize medical emergency responses. The new analytics system recommends to the dispatcher an appropriate response to a medical emergency call—whether a patient can be treated on-site or needs to be taken to the hospital—by taking into account several factors, such as the type of call, location, weather, and similar calls.³⁴

Since it fields 80,000 requests each year, Cincinnati officials are deploying this technology to prioritize responses and determine the best ways to handle emergencies. They see AI as a way to deal with large volumes of data and figure out efficient ways of responding to public requests. Rather than address service issues in an ad hoc manner, authorities are trying to be proactive in how they provide urban services.

Cincinnati is not alone. A number of metropolitan areas are adopting smart city applications that use AI to improve service delivery, environmental planning, resource management, energy utilization, and crime prevention, among other things. For its smart cities index, the magazine Fast Company ranked American locales and found Seattle, Boston, San Francisco, Washington, D.C., and New York City as the top adopters. Seattle, for

example, has embraced sustainability and is using AI to manage energy usage and resource management. Boston has launched a “City Hall To Go” that makes sure underserved communities receive needed public services. It also has deployed “cameras and inductive loops to manage traffic and acoustic sensors to identify gun shots.” San Francisco has certified 203 buildings as meeting LEED sustainability standards.³⁵

Through these and other means, metropolitan areas are leading the country in the deployment of AI solutions. Indeed, according to a National League of Cities report, 66 percent of American cities are investing in smart city technology. Among the top applications noted in the report are “smart meters for utilities, intelligent traffic signals, e-governance applications, Wi-Fi kiosks, and radio frequency identification sensors in pavement.”³⁶

CHAPTER III

Policy, regulatory, and ethical issues

These examples from a variety of sectors demonstrate how AI is transforming many walks of human existence. The increasing penetration of AI and autonomous devices into many aspects of life is altering basic operations and decisionmaking within organizations, and improving efficiency and response times.

At the same time, though, these developments raise important policy, regulatory, and ethical issues. For example, how should we promote data access? How do we guard against biased or unfair data used in algorithms? What types of ethical principles are introduced through software

programming, and how transparent should designers be about their choices? What about questions of legal liability in cases where algorithms cause harm?³⁷

The increasing penetration of AI into many aspects of life is altering decisionmaking within organizations and improving efficiency. At the same time, though, these developments raise important policy, regulatory, and ethical issues.

Data access problems

The key to getting the most out of AI is having a “data-friendly ecosystem with unified standards and cross-platform sharing.” AI depends on data that can be analyzed in real time and brought to bear on concrete problems. Having data that are “accessible for exploration” in the research community is a prerequisite for successful AI development.³⁸

According to a McKinsey Global Institute study, nations that promote open data sources and data sharing are the ones most likely to see AI advances. In this regard, the United States has a substantial advantage over China. Global ratings on data openness show that U.S. ranks eighth overall in the world, compared to 93 for China.³⁹

But right now, the United States does not have a coherent national data strategy. There are few protocols for promoting research access or platforms that make it possible to gain new insights from proprietary data. It is not always clear who owns data or how much belongs in the public sphere. These uncertainties limit the innovation economy and act as a drag on academic research. In the following section, we outline ways to improve data access for researchers.

Biases in data and algorithms

In some instances, certain AI systems are thought to have enabled discriminatory or biased practices.⁴⁰ For example, Airbnb has been accused of having homeowners on its platform who discriminate against racial minorities. A research project undertaken by the Harvard Business School found that “Airbnb users with distinctly African American names were roughly 16 percent less likely to be accepted as guests than those with distinctly white names.”⁴¹

Racial issues also come up with facial recognition software. Most such systems operate by comparing a person’s face to a range of faces 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.”⁴² 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 get 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.⁴³

AI ethics and transparency

Algorithms embed ethical considerations and value choices into program decisions. As such, these systems raise questions concerning the criteria used in automated decisionmaking. Some people want to have a better understanding of how algorithms function and what choices are being made.⁴⁴

In the United States, many urban schools use algorithms for enrollment decisions based on a variety of considerations, such as parent preferences, neighborhood qualities, income level, and demographic background. According to Brookings researcher Jon Valant, the New Orleans–based Bricolage Academy “gives priority to economically disadvantaged applicants for up to 33 percent of available seats. In practice, though, most cities have opted for categories that prioritize siblings of current students, children of school employees, and families that live in school’s broad geographic area.”⁴⁵ Enrollment choices can be expected to be very different when considerations of this sort come into play.

Depending on how AI systems are set up, they can facilitate the redlining of mortgage applications, help people discriminate against individuals they don’t like, or help screen or build rosters of individuals based on unfair criteria. The types of considerations that go into programming decisions matter a lot in terms of how the systems operate and how they affect customers.⁴⁶

For these reasons, the EU is implementing the General Data Protection Regulation (GDPR) in May 2018. The rules specify that people have “the right to opt out of personally tailored ads” and “can contest ‘legal or similarly significant’ decisions made by algorithms and appeal for human intervention”

in the form of an explanation of how the algorithm generated a particular outcome. Each guideline is designed to ensure the protection of personal data and provide individuals with information on how the “black box” operates.⁴⁷

Legal liability

There are questions concerning the legal liability of AI systems. If there are harms or infractions (or fatalities in the case of driverless cars), the operators of the algorithm likely will fall under product liability rules. A body of case law has shown that the situation’s facts and circumstances determine liability and influence the kind of penalties that are imposed. Those can range from civil fines to imprisonment for major harms.⁴⁸ The Uber-related fatality in Arizona will be an important test case for legal liability. The state actively recruited Uber to test its autonomous vehicles and gave the company considerable latitude in terms of road testing. It remains to be seen if there will be lawsuits in this case and who is sued: the human backup driver, the state of Arizona, the Phoenix suburb where the accident took place, Uber, software developers, or the auto manufacturer. Given the multiple people and organizations involved in the road testing, there are many legal questions to be resolved.

In non-transportation areas, digital platforms often have limited liability for what happens on their sites. For example, in the case of Airbnb, the firm “requires that people agree to waive their right to sue, or to join in any class-action lawsuit or class-action arbitration, to use the service.” By demanding that its users sacrifice basic rights, the company limits consumer protections and therefore curtails the ability of people to fight discrimination arising from unfair algorithms.⁴⁹ But whether the principle of neutral networks holds up in many sectors is yet to be determined on a widespread basis.

CHAPTER IV

Recommendations

In order to balance innovation with basic human values, we propose a number of recommendations for moving forward with AI. This includes improving data access, increasing government investment in AI, promoting AI workforce development, creating a federal advisory committee, engaging with state and local officials to ensure they enact effective policies, regulating broad objectives as opposed to specific algorithms, taking bias seriously as an AI issue, maintaining mechanisms for human control and oversight, and penalizing malicious behavior and promoting cybersecurity.

Improving data access

The United States should develop a data strategy that promotes innovation and consumer protection. Right now, there are no uniform standards in terms of data access, data sharing, or data protection. Almost all the data are proprietary in nature and not shared very broadly with the research community, and this limits innovation and system design. AI requires data to test and improve its learning capacity.⁵⁰ Without structured and unstructured data sets, it will be nearly impossible to gain the full benefits of artificial intelligence.

In general, the research community needs better access to government and business data, although with appropriate safeguards to make sure researchers do not misuse data in the way Cambridge Analytica did with Facebook information. There is a variety of ways researchers could gain

data access. One is through voluntary agreements with companies holding proprietary data. Facebook, for example, recently announced a partnership with Stanford economist Raj Chetty to use its social media data to explore inequality.⁵¹ As part of the arrangement, researchers were required to undergo background checks and could only access data from secured sites in order to protect user privacy and security.

In the U.S., there are no uniform standards in terms of data access, data sharing, or data protection. Almost all the data are proprietary in nature and not shared very broadly with the research community, and this limits innovation and system design.

Google long has made available search results in aggregated form for researchers and the general public. Through its “Trends” site, scholars can analyze topics such as interest in Trump, views about democracy, and perspectives on the overall economy.⁵² [\(#_edn52\)](#) That helps people track movements in public interest and identify topics that galvanize the general public.

Twitter makes much of its tweets available to researchers through application programming interfaces, commonly referred to as APIs. These tools help people outside the company build application software and make use of data from its social media platform. They can study patterns of social media communications and see how people are commenting on or reacting to current events.

In some sectors where there is a discernible public benefit, governments can facilitate collaboration by building infrastructure that shares data. For example, the National Cancer Institute has pioneered a data-sharing protocol where certified researchers can query health data it has using de-

identified information drawn from clinical data, claims information, and drug therapies. That enables researchers to evaluate efficacy and effectiveness, and make recommendations regarding the best medical approaches, without compromising the privacy of individual patients.

There could be public-private data partnerships that combine government and business data sets to improve system performance. For example, cities could integrate information from ride-sharing services with its own material on social service locations, bus lines, mass transit, and highway congestion to improve transportation. That would help metropolitan areas deal with traffic tie-ups and assist in highway and mass transit planning.

Some combination of these approaches would improve data access for researchers, the government, and the business community, without impinging on personal privacy. As noted by Ian Buck, the vice president of NVIDIA, “Data is the fuel that drives the AI engine. The federal government has access to vast sources of information. Opening access to that data will help us get insights that will transform the U.S. economy.”⁵³ Through its Data.gov portal, the federal government already has put over 230,000 data sets into the public domain, and this has propelled innovation and aided improvements in AI and data analytic technologies.⁵⁴ The private sector also needs to facilitate research data access so that society can achieve the full benefits of artificial intelligence.

Increase government investment in AI

According to Greg Brockman, the co-founder of OpenAI, the U.S. federal government invests only \$1.1 billion in non-classified AI technology.⁵⁵ That is far lower than the amount being spent by China or other leading nations in this area of research. That shortfall is noteworthy because the economic

payoffs of AI are substantial. In order to boost economic development and social innovation, federal officials need to increase investment in artificial intelligence and data analytics. Higher investment is likely to pay for itself many times over in economic and social benefits.⁵⁶

Promote digital education and workforce development

As AI applications accelerate across many sectors, it is vital that we reimagine our educational institutions for a world where AI will be ubiquitous and students need a different kind of training than they currently receive. Right now, many students do not receive instruction in the kinds of skills that will be needed in an AI-dominated landscape. For example, there currently are shortages of data scientists, computer scientists, engineers, coders, and platform developers. These are skills that are in short supply; unless our educational system generates more people with these capabilities, it will limit AI development.

For these reasons, both state and federal governments have been investing in AI human capital. For example, in 2017, the National Science Foundation funded over 6,500 graduate students in computer-related fields and has launched several new initiatives designed to encourage data and computer science at all levels from pre-K to higher and continuing education.⁵⁷ The goal is to build a larger pipeline of AI and data analytic personnel so that the United States can reap the full advantages of the knowledge revolution.

But there also needs to be substantial changes in the process of learning itself. It is not just technical skills that are needed in an AI world but skills of critical reasoning, collaboration, design, visual display of information, and independent thinking, among others. AI will reconfigure how society and the economy operate, and there needs to be “big picture” thinking on what this

will mean for ethics, governance, and societal impact. People will need the ability to think broadly about many questions and integrate knowledge from a number of different areas.

One example of new ways to prepare students for a digital future is IBM's Teacher Advisor program, utilizing Watson's free online tools to help teachers bring the latest knowledge into the classroom. They enable instructors to develop new lesson plans in STEM and non-STEM fields, find relevant instructional videos, and help students get the most out of the classroom.⁵⁸ As such, they are precursors of new educational environments that need to be created.

Create a federal AI advisory committee

Federal officials need to think about how they deal with artificial intelligence. As noted previously, there are many issues ranging from the need for improved data access to addressing issues of bias and discrimination. It is vital that these and other concerns be considered so we gain the full benefits of this emerging technology.

In order to move forward in this area, several members of Congress have introduced the "Future of Artificial Intelligence Act," a bill designed to establish broad policy and legal principles for AI. It proposes the secretary of commerce create a federal advisory committee on the development and implementation of artificial intelligence. The legislation provides a mechanism for the federal government to get advice on ways to promote a "climate of investment and innovation to ensure the global competitiveness of the United States," "optimize the development of artificial intelligence to address the potential growth, restructuring, or other changes in the United States workforce," "support the unbiased development and application of

artificial intelligence,” and “protect the privacy rights of individuals.”⁵⁹

Among the specific questions the committee is asked to address include the following: competitiveness, workforce impact, education, ethics training, data sharing, international cooperation, accountability, machine learning bias, rural impact, government efficiency, investment climate, job impact, bias, and consumer impact. The committee is directed to submit a report to Congress and the administration 540 days after enactment regarding any legislative or administrative action needed on AI.

This legislation is a step in the right direction, although the field is moving so rapidly that we would recommend shortening the reporting timeline from 540 days to 180 days. Waiting nearly two years for a committee report will certainly result in missed opportunities and a lack of action on important issues. Given rapid advances in the field, having a much quicker turnaround time on the committee analysis would be quite beneficial.

Engage with state and local officials

States and localities also are taking action on AI. For example, the New York City Council unanimously passed a bill that directed the mayor to form a taskforce that would “monitor the fairness and validity of algorithms used by municipal agencies.”⁶⁰ The city employs algorithms to “determine if a lower bail will be assigned to an indigent defendant, where firehouses are established, student placement for public schools, assessing teacher performance, identifying Medicaid fraud and determine where crime will happen next.”⁶¹

According to the legislation’s developers, city officials want to know how these algorithms work and make sure there is sufficient AI transparency and

accountability. In addition, there is concern regarding the fairness and biases of AI algorithms, so the taskforce has been directed to analyze these issues and make recommendations regarding future usage. It is scheduled to report back to the mayor on a range of AI policy, legal, and regulatory issues by late 2019.

Some observers already are worrying that the taskforce won't go far enough in holding algorithms accountable. For example, Julia Powles of Cornell Tech and New York University argues that the bill originally required companies to make the AI source code available to the public for inspection, and that there be simulations of its decisionmaking using actual data. After criticism of those provisions, however, former Councilman James Vacca dropped the requirements in favor of a task force studying these issues. He and other city officials were concerned that publication of proprietary information on algorithms would slow innovation and make it difficult to find AI vendors who would work with the city.⁶² It remains to be seen how this local task force will balance issues of innovation, privacy, and transparency.

Regulate broad objectives more than specific algorithms

The European Union has taken a restrictive stance on these issues of data collection and analysis.⁶³ It has rules limiting the ability of companies from collecting data on road conditions and mapping street views. Because many of these countries worry that people's personal information in unencrypted Wi-Fi networks are swept up in overall data collection, the EU has fined technology firms, demanded copies of data, and placed limits on the material collected.⁶⁴ This has made it more difficult for technology companies operating there to develop the high-definition maps required for autonomous vehicles.

The GDPR being implemented in Europe place severe restrictions on the use of artificial intelligence and machine learning. According to published guidelines, “Regulations prohibit any automated decision that ‘significantly affects’ EU citizens. This includes techniques that evaluates a person’s ‘performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location, or movements.’”⁶⁵ In addition, these new rules give citizens the right to review how digital services made specific algorithmic choices affecting people.

By taking a restrictive stance on issues of data collection and analysis, the European Union is putting its manufacturers and software designers at a significant disadvantage to the rest of the world.

If interpreted stringently, these rules will make it difficult for European software designers (and American designers who work with European counterparts) to incorporate artificial intelligence and high-definition mapping in autonomous vehicles. Central to navigation in these cars and trucks is tracking location and movements. Without high-definition maps containing geo-coded data and the deep learning that makes use of this information, fully autonomous driving will stagnate in Europe. Through this and other data protection actions, the European Union is putting its manufacturers and software designers at a significant disadvantage to the rest of the world.

It makes more sense to think about the broad objectives desired in AI and enact policies that advance them, as opposed to governments trying to crack open the “black boxes” and see exactly how specific algorithms operate. Regulating individual algorithms will limit innovation and make it difficult for companies to make use of artificial intelligence.

Take biases seriously

Bias and discrimination are serious issues for AI. There already have been a number of cases of unfair treatment linked to historic data, and steps need to be undertaken to make sure that does not become prevalent in artificial intelligence. Existing statutes governing discrimination in the physical economy need to be extended to digital platforms. That will help protect consumers and build confidence in these systems as a whole.

For these advances to be widely adopted, more transparency is needed in how AI systems operate. Andrew Burt of Immuta argues, “The key problem confronting predictive analytics is really transparency. We’re in a world where data science operations are taking on increasingly important tasks, and the only thing holding them back is going to be how well the data scientists who train the models can explain what it is their models are doing.”
66

Maintaining mechanisms for human oversight and control

Some individuals have argued that there needs to be avenues for humans to exercise oversight and control of AI systems. For example, Allen Institute for Artificial Intelligence CEO Oren Etzioni argues there should be rules for regulating these systems. First, he says, AI must be governed by all the laws that already have been developed for human behavior, including regulations concerning “cyberbullying, stock manipulation or terrorist threats,” as well as “entrap[ping] people into committing crimes.” Second, he believes that these systems should disclose they are automated systems and not human beings. Third, he states, “An A.I. system cannot retain or disclose confidential information without explicit approval from the source of that information.”⁶⁷
His rationale is that these tools store so much data that people have to be

cognizant of the privacy risks posed by AI.

In the same vein, the IEEE Global Initiative has ethical guidelines for AI and autonomous systems. Its experts suggest that these models be programmed with consideration for widely accepted human norms and rules for behavior. AI algorithms need to take into effect the importance of these norms, how norm conflict can be resolved, and ways these systems can be transparent about norm resolution. Software designs should be programmed for “nondeception” and “honesty,” according to ethics experts. When failures occur, there must be mitigation mechanisms to deal with the consequences. In particular, AI must be sensitive to problems such as bias, discrimination, and fairness.⁶⁸

A group of machine learning experts claim it is possible to automate ethical decisionmaking. Using the trolley problem as a moral dilemma, they ask the following question: If an autonomous car goes out of control, should it be programmed to kill its own passengers or the pedestrians who are crossing the street? They devised a “voting-based system” that asked 1.3 million people to assess alternative scenarios, summarized the overall choices, and applied the overall perspective of these individuals to a range of vehicular possibilities. That allowed them to automate ethical decisionmaking in AI algorithms, taking public preferences into account.⁶⁹ This procedure, of course, does not reduce the tragedy involved in any kind of fatality, such as seen in the Uber case, but it provides a mechanism to help AI developers incorporate ethical considerations in their planning.

Penalize malicious behavior and promote cybersecurity

As with any emerging technology, it is important to discourage malicious treatment designed to trick software or use it for undesirable ends.⁷⁰ This is

especially important given the dual-use aspects of AI, where the same tool can be used for beneficial or malicious purposes. The malevolent use of AI exposes individuals and organizations to unnecessary risks and undermines the virtues of the emerging technology. This includes behaviors such as hacking, manipulating algorithms, compromising privacy and confidentiality, or stealing identities. Efforts to hijack AI in order to solicit confidential information should be seriously penalized as a way to deter such actions.⁷¹

In a rapidly changing world with many entities having advanced computing capabilities, there needs to be serious attention devoted to cybersecurity. Countries have to be careful to safeguard their own systems and keep other nations from damaging their security.⁷² According to the U.S. Department of Homeland Security, a major American bank receives around 11 million calls a week at its service center. In order to protect its telephony from denial of service attacks, it uses a “machine learning-based policy engine [that] blocks more than 120,000 calls per month based on voice firewall policies including harassing callers, robocalls and potential fraudulent calls.”⁷³ This represents a way in which machine learning can help defend technology systems from malevolent attacks.

CHAPTER V

Conclusion

To summarize, the world is on the cusp of revolutionizing many sectors through artificial intelligence and data analytics. There already are significant deployments in finance, national security, health care, criminal justice, transportation, and smart cities that have altered decisionmaking, business models, risk mitigation, and system performance. These developments are

generating substantial economic and social benefits.

The world is on the cusp of revolutionizing many sectors through artificial intelligence, but the way AI systems are developed need to be better understood due to the major implications these technologies will have for society as a whole.

Yet the manner in which AI systems unfold has major implications for society as a whole. It matters how policy issues are addressed, ethical conflicts are reconciled, legal realities are resolved, and how much transparency is required in AI and data analytic solutions.⁷⁴ Human choices about software development affect the way in which decisions are made and the manner in which they are integrated into organizational routines. Exactly how these processes are executed need to be better understood because they will have substantial impact on the general public soon, and for the foreseeable future. AI may well be a revolution in human affairs, and become the single most influential human innovation in history.

Note: We appreciate the research assistance of Grace Gilberg, Jack Karsten, Hillary Schaub, and Kristjan Tomasson on this project.

The Brookings Institution is a nonprofit organization devoted to independent research and policy solutions. Its mission is to conduct high-quality, independent research and, based on that research, to provide innovative, practical recommendations for policymakers and the public. The conclusions and recommendations of any Brookings publication are solely those of its author(s), and do not reflect the views of the Institution, its management, or its other scholars.

Support for this publication was generously provided by Amazon. Brookings recognizes that the value it provides is in its absolute commitment to quality, independence, and impact. Activities supported by its donors reflect this commitment.

John R. Allen is a member of the Board of Advisors of Amida Technology and on the Board of Directors of Spark Cognition. Both companies work in fields discussed in this piece.

[\(#_ednref1\)](#)

AUTHORS



Darrell M. West Senior Fellow - Center for Technology Innovation,
Douglas Dillon Chair in Governmental Studies



John R. Allen

Footnotes

1. Thomas Davenport, Jeff Loucks, and David Schatsky, "Bullish on the Business Value of

Cognitive” (Deloitte, 2017), p. 3 (www2.deloitte.com/us/en/pages/deloitte-analytics/articles/cognitive-technology-adoption-survey.html).

2. Luke Dormehl, *Thinking Machines: The Quest for Artificial Intelligence—and Where It’s Taking Us Next* (New York: Penguin–TarcherPerigee, 2017).
3. Shubhendu and Vijay, “Applicability of Artificial Intelligence in Different Fields of Life.”
4. Ibid.
5. Andrew McAfee and Erik Brynjolfsson, *Machine Platform Crowd: Harnessing Our Digital Future* (New York: Norton, 2017).
6. Portions of this paper draw on Darrell M. West, *The Future of Work: Robots, AI, and Automation*, Brookings Institution Press, 2018.
7. PriceWaterhouseCoopers, “Sizing the Prize: What’s the Real Value of AI for Your Business and How Can You Capitalise?” 2017.
8. Dominic Barton, Jonathan Woetzel, Jeongmin Seong, and Qinzhen Tian, “Artificial Intelligence: Implications for China” (New York: McKinsey Global Institute, April 2017), p. 1.
9. Nathaniel Popper, “Stocks and Bots,” *New York Times Magazine*, February 28, 2016.
10. Ibid.
11. Ibid.
12. Michael Lewis, *Flash Boys: A Wall Street Revolt* (New York: Norton, 2015).
13. Cade Metz, “In Quantum Computing Race, Yale Professors Battle Tech Giants,” *New York Times*, November 14, 2017, p. B3.
14. Executive Office of the President, “Artificial Intelligence, Automation, and the Economy,” December 2016, pp. 27-28.
15. Christian Davenport, “Future Wars May Depend as Much on Algorithms as on Ammunition, Report Says,” *Washington Post*, December 3, 2017.
16. Ibid.
17. John R. Allen and Amir Husain, “On Hyperwar,” *Naval Institute Proceedings*, July 17, 2017, pp. 30-36.
18. Paul Mozur, “China Sets Goal to Lead in Artificial Intelligence,” *New York Times*, July 21, 2017, p. B1.
19. Paul Mozur and John Markoff, “Is China Outsmarting American Artificial Intelligence?”

New York Times, May 28, 2017.

20. *Economist*, "America v China: The Battle for Digital Supremacy," March 15, 2018.
21. Rasmus Rothe, "Applying Deep Learning to Real-World Problems," *Medium*, May 23, 2017.
22. Eric Horvitz, "Reflections on the Status and Future of Artificial Intelligence," Testimony before the U.S. Senate Subcommittee on Space, Science, and Competitiveness, November 30, 2016, p. 5.
23. Jeff Asher and Rob Arthur, "Inside the Algorithm That Tries to Predict Gun Violence in Chicago," *New York Times Upshot*, June 13, 2017.
24. Caleb Watney, "It's Time for our Justice System to Embrace Artificial Intelligence," *TechTank* (blog), Brookings Institution, July 20, 2017.
25. Asher and Arthur, "Inside the Algorithm That Tries to Predict Gun Violence in Chicago."
26. Paul Mozur and Keith Bradsher, "China's A.I. Advances Help Its Tech Industry, and State Security," *New York Times*, December 3, 2017.
27. Simon Denyer, "China's Watchful Eye," *Washington Post*, January 7, 2018.
28. Cameron Kerry and Jack Karsten, "Gauging Investment in Self-Driving Cars," Brookings Institution, October 16, 2017.
29. Portions of this section are drawn from Darrell M. West, "Driverless Cars in China, Europe, Japan, Korea, and the United States," Brookings Institution, September 2016.
30. Ibid.
31. Yuming Ge, Xiaoman Liu, Libo Tang, and Darrell M. West, "Smart Transportation in China and the United States," Center for Technology Innovation, Brookings Institution, December 2017.
32. Peter Holley, "Uber Signs Deal to Buy 24,000 Autonomous Vehicles from Volvo," *Washington Post*, November 20, 2017.
33. Daisuke Wakabayashi, "Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam," *New York Times*, March 19, 2018.
34. Kevin Desouza, Rashmi Krishnamurthy, and Gregory Dawson, "Learning from Public Sector Experimentation with Artificial Intelligence," *TechTank* (blog), Brookings Institution, June 23, 2017.
35. Boyd Cohen, "The 10 Smartest Cities in North America," *Fast Company*, November 14, 2013.

36. Teena Maddox, "66% of US Cities Are Investing in Smart City Technology," *TechRepublic*, November 6, 2017.
37. Osonde Osoba and William Welser IV, "The Risks of Artificial Intelligence to Security and the Future of Work" (Santa Monica, Calif.: RAND Corp., December 2017) (www.rand.org/pubs/perspectives/PE237.html).
38. *Ibid.*, p. 7.
39. Dominic Barton, Jonathan Woetzel, Jeongmin Seong, and Qinzhen Tian, "Artificial Intelligence: Implications for China" (New York: McKinsey Global Institute, April 2017), p. 7.
40. Executive Office of the President, "Preparing for the Future of Artificial Intelligence," October 2016, pp. 30-31.
41. Elaine Glusac, "As Airbnb Grows, So Do Claims of Discrimination," *New York Times*, June 21, 2016.
42. "Joy Buolamwini," *Bloomberg Businessweek*, July 3, 2017, p. 80.
43. *Ibid.*
44. Mark Purdy and Paul Daugherty, "Why Artificial Intelligence is the Future of Growth," Accenture, 2016.
45. Jon Valant, "Integrating Charter Schools and Choice-Based Education Systems," *Brown Center Chalkboard blog*, Brookings Institution, June 23, 2017.
46. Tucker, "A White Mask Worked Better."
47. Cliff Kuang, "Can A.I. Be Taught to Explain Itself?" *New York Times Magazine*, November 21, 2017.
48. Yale Law School Information Society Project, "Governing Machine Learning," September 2017.
49. Katie Benner, "Airbnb Vows to Fight Racism, But Its Users Can't Sue to Prompt Fairness," *New York Times*, June 19, 2016.
50. Executive Office of the President, "Artificial Intelligence, Automation, and the Economy" and "Preparing for the Future of Artificial Intelligence."
51. Nancy Sclar, "Facebook's Next Project: American Inequality," *Politico*, February 19, 2018.
52. Darrell M. West, "What Internet Search Data Reveals about Donald Trump's First Year in Office," Brookings Institution policy report, January 17, 2018.

53. Ian Buck, "Testimony before the House Committee on Oversight and Government Reform Subcommittee on Information Technology," February 14, 2018.
54. Keith Nakasone, "Testimony before the House Committee on Oversight and Government Reform Subcommittee on Information Technology," March 7, 2018.
55. Greg Brockman, "The Dawn of Artificial Intelligence," Testimony before U.S. Senate Subcommittee on Space, Science, and Competitiveness, November 30, 2016.
56. Amir Khosrowshahi, "Testimony before the House Committee on Oversight and Government Reform Subcommittee on Information Technology," February 14, 2018.
57. James Kurose, "Testimony before the House Committee on Oversight and Government Reform Subcommittee on Information Technology," March 7, 2018.
58. Stephen Noonoo, "Teachers Can Now Use IBM's Watson to Search for Free Lesson Plans," *EdSurge*, September 13, 2017.
59. Congress.gov, "H.R. 4625 FUTURE of Artificial Intelligence Act of 2017," December 12, 2017.
60. Elizabeth Zima, "Could New York City's AI Transparency Bill Be a Model for the Country?" *Government Technology*, January 4, 2018.
61. Ibid.
62. Julia Powles, "New York City's Bold, Flawed Attempt to Make Algorithms Accountable," *New Yorker*, December 20, 2017.
63. Sheera Frenkel, "Tech Giants Brace for Europe's New Data Privacy Rules," *New York Times*, January 28, 2018.
64. Claire Miller and Kevin O'Brien, "Germany's Complicated Relationship with Google Street View," *New York Times*, April 23, 2013.
65. Cade Metz, "Artificial Intelligence is Setting Up the Internet for a Huge Clash with Europe," *Wired*, July 11, 2016.
66. Eric Siegel, "Predictive Analytics Interview Series: Andrew Burt," *Predictive Analytics Times*, June 14, 2017.
67. Oren Etzioni, "How to Regulate Artificial Intelligence," *New York Times*, September 1, 2017.
68. "Ethical Considerations in Artificial Intelligence and Autonomous Systems," unpublished paper. IEEE Global Initiative, 2018.
69. Ritesh Noothigattu, Snehal Kumar Gaikwad, Edmond Awad, Sohan Dsouza, Iyad

Rahwan, Pradeep Ravikumar, and Ariel Procaccia, "A Voting-Based System for Ethical Decision Making," *Computers and Society*, September 20, 2017 (www.media.mit.edu/publications/a-voting-based-system-for-ethical-decision-making/).

70. Miles Brundage, et al., "The Malicious Use of Artificial Intelligence," University of Oxford unpublished paper, February 2018.
 71. John Markoff, "As Artificial Intelligence Evolves, So Does Its Criminal Potential," *New York Times*, October 24, 2016, p. B3.
 72. *Economist*, "The Challenger: Technopolitics," March 17, 2018.
 73. Douglas Maughan, "Testimony before the House Committee on Oversight and Government Reform Subcommittee on Information Technology," March 7, 2018.
 74. Levi Tillemann and Colin McCormick, "Roadmapping a U.S.-German Agenda for Artificial Intelligence Policy," New American Foundation, March 2017.
-