ONE | WHAT IS AI?

Artificial intelligence (AI) is a difficult concept to grasp. As an illustration, only 17 percent of 1,500 U.S. business leaders in 2017 said they were familiar with how AI would affect their companies.[1] These executives understood there was considerable potential for revolutionizing business processes but were not clear how AI would be deployed within their own organizations or how it would alter their industries.

Hollywood offers little help in improving people's understanding of advanced technologies. Many movies conflate AI with malevolent robots or hyperintelligent beings such as the Terminator or the evil HAL in Arthur C. Clarke's *2001: A Space Odyssey*. In film depictions, superpowered entities inevitably gain humanlike intelligence, go rogue, and inflict tremendous harm on the human population. The ominous message in these cinematic renditions is that AI is dangerous because it will eventually enslave humans.[2]

One of the difficulties in understanding AI is the lack of a uniform definition. People often intertwine many things in their conceptions and then imagine the worst possible outcomes. They assume advanced technologies will have omniscient capabilities, destructive motivations, and lim-

ited human oversight and will be impossible to control. For those reasons, it is important to clarify what we mean by artificial intelligence, provide understandable examples of how it is being used, and outline its major risks.

AI ORIGINS

Alan Turing generally is credited with conceptualizing the idea of AI in 1950, when he speculated about "thinking machines" that could reason at the level of a human being. His well-known "Turing Test" specifies that computers need to complete "reasoning puzzles" as effectively as humans to be considered autonomous.[3]

Turing was followed a few years later by visionary John McCarthy. In 1956, he was the first scientist to coin the term "artificial intelligence" to denote machines that could think and act independently. He defined AI as "getting a computer to do things which, when done by people, are said to involve intelligence."

Since that time, scientists have argued over what constitutes "thinking," "intelligence," and "autonomy" when it comes to hardware and software. Advanced computers such as IBM's Watson already have beaten humans at chess and poker, and are capable of instantly processing enormous amounts of information.[4] Deep Mind's AlphaGo and AlphaGo Master have advanced even further and used AI to defeat expert players of the board game Go.[5] Learning from a series of competitions, these AI-based systems figured out how to make complex moves and develop strategies that humans never had considered.

More recently, futurist Ray Kurzweil raised the AI notion to a new level when he predicted a "singularity," that is, a "machine-based super-intelligence [that is] greater than human intelligence."[6] With advances in AI, data analytics, and machine learning, it no longer seems far-fetched to foresee sentient machines that can take on advanced functions.[7] As computing power grows, Kurzweil believes there will be a convergence between humans and computers.

The wide scope in how people define artificial intelligence lends itself to a broad range of interpretations regarding its societal impact. Those who

are worried about a loss of humanity can point to plausible scenarios where robots make everything and Terminator-like war machines operate independently of human control. Conversely, people who are more optimistic can point to ways that automated machines could spread convenience, relieve humans of boring or dangerous tasks, and improve the overall quality of life. Being specific about what we mean by AI is important for how observers evaluate the possible array of opportunities and risks.

DEFINING AI

Engineers Shukla Shubhendu and Jaiswal Vijay define AI as "machines that respond to stimulation consistent with traditional responses from humans, given the human capacity for contemplation, judgment, and intention."[8] This interpretation is helpful because it suggests several qualities that separate AI from mechanical devices or traditional computer software: intentionality, intelligence, and adaptability. It is these features that enable AI algorithms to find patterns or associations through "neural networks" that group data based on common similarities (see Glossary of Key Terms). In addition, AI-based systems learn from insights gathered via "deep learning" techniques that use statistics to spot underlying trends or patterns in data and apply that knowledge to changing circumstances. And it is not just numbers that can be mined, as AI can employ "natural language processing" that analyzes textual information in order to make sense of its meaning. As we outline below, AI employs a number of sophisticated traits to bring powerful computational strengths to a wide range of endeavors.[9]

Intentionality

Artificial intelligence algorithms are designed to undertake actions or make decisions using real-time data. They are unlike passive machines that are only capable of static or mechanical responses. Using sensors, digital data, or remote inputs, advanced algorithms combine information from a variety of sources, analyze the material instantly, and act on the insights derived from those data.[10]

With massive improvements in storage systems, processing speeds, and analytic techniques, AI algorithms are capable of tremendous sophistication in analysis and decisionmaking.[11] For example, financial algorithms can spot minute differentials in stock valuations and undertake market transactions based on that information. Similarly, artificially intelligent sensors on dynamic systems such as electric generators or aircraft turbine engines can sense operating anomalies far earlier than previous instrumentation and thereby save millions of dollars by signaling the need for preventive maintenance. The same logic applies in environmental sustainability systems that use sensors to determine whether someone is in a room and automatically adjust heating, cooling, and lighting utilizing that sensory input.

As long as these systems conform to human-based values, there is little risk of AI endangering human beings. Sophisticated algorithms can be intentional while analyzing information in ways that augment human performance and understanding. However, if the software is poorly designed or based on incomplete or biased information, it can make decisions that are unfair or discriminatory and therefore create either intentional or unintentional harm.

Intelligence

AI is adept at finding statistical associations through deep learning techniques. Using sophisticated data analytics, software designers can develop algorithms that identify data patterns, and use that knowledge to make specific decisions.

As we discuss in chapter 3, there are AI systems for managing school enrollments. These systems compile information on neighborhood location, desired schools, and enrollment preferences, and assign pupils based on that compiled information. As long as there is little disagreement regarding basic criteria, these systems can work intelligently and effectively.[12]

Of course, that standard often is not the case in many policy areas. Reflecting the importance of education for life outcomes, parents, teachers, and school administrators fight over the weighting of different factors.

Should students be assigned to their neighborhood school or should other criteria, such as a desire to overcome residential segregation and diversify schools, override locational considerations? Those criteria are important because in a city with widespread racial segregation and economic inequalities, emphasizing neighborhood school assignments invariably exacerbates inequality and heightens racial segregation.

This is an important point because making those kinds of decisions increasingly falls to computer programmers, many of whom are inexperienced when it comes to resolving ethical disputes or values-based disagreements. Programmers are adept at building intelligent algorithms that compile information based on specified considerations, but figuring out how to reconcile conflicting values is not something for which most of them are trained. Computer programmers' expertise is in analyzing data and writing code, not resolving societal disagreements about fundamental objectives.

Adaptability

The last quality that marks AI systems is the ability to learn and adapt as they compile information and make decisions. Effective artificial intelligence makes adjustments as circumstances or conditions shift. This may involve alterations in financial situations, road conditions, environmental considerations, or military circumstances pertaining to the AI system. In each of these areas, AI has to integrate new information quickly in its algorithms and make shifts based on relevant data.

Adaptability represents a valuable characteristic of AI because conditions change quickly in many areas and organizations need the capacity to assess needs and requirements for adaptation and make relevant decisions that can be significantly assisted and even accelerated by trained algorithms. That agility is an important benefit of autonomous algorithms over earlier systems that merely project past trends into the future in a linear manner.

Transportation illustrates how this adaptability and agility can happen. Autonomous vehicles use machine-to-machine communications to alert other cars on the road about upcoming congestion, potholes,

highway construction, or traffic impediments. Vehicles can take advantage of other vehicles' road experience, without human involvement. Everything that is learned can be immediately and fully transferable to other connected vehicles. Their advanced algorithms, sensors, and cameras integrate real-time experiences and use dashboards and visual displays to help people make sense of ongoing traffic and vehicular conditions. This is made possible by advances in computing power.

Similar capabilities are incorporated into software for scheduling appointments. Digital personal assistants can ascertain a person's preferences and respond to email requests for appointments in a dynamic manner. Without any human intervention, the digital assistant can make appointments, adjust schedules, and communicate preferences to other individuals.

Increasingly, human interactions with computers are taking place through verbal conversation. Rather than the conventional technique of point and click based on mouse navigation, people can talk to computers and provide oral instructions of what they want to do.[13] Through "chatbots" built into devices such as Amazon Alexa and Apple Siri, people can order pizza, pay bills, or request music through conversational interfaces.

The simplicity of these tools dramatically expands the possibilities for AI and human-machine interactions. Voice-activated features can turn what used to be complex interfaces into easy-to-use consumer devices. People no longer have to be computer scientists or engineers to utilize complex systems or operate sophisticated products. They can employ AI without even knowing they are interacting with an advanced algorithm. The software code may have millions of programming lines built into it but only require verbal commands to operate.

AI DEPLOYMENT

With its growing capabilities, AI is being deployed in a number of sectors. Fields ranging from finance and shopping to smart cities, energy management, and criminal justice are incorporating AI, data analytics, and machine learning. Motivated by efforts to build efficiency and improve

effectiveness, algorithms have altered the way many leaders decide and organizations operate. Below are brief descriptions of some ways in which algorithms are being utilized.

Finance

The financial services industry employs many AI applications.[14] According to industry observers, "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."[15] In addition, there are so-called robo-advisers that "create personalized investment portfolios, obviating the need for stockbrokers and financial advisers."[16] These advances take the emotion out of investing and make decisions based on analytical considerations.[17]

A prominent illustration of this is stock exchanges, where automated, high-frequency trading has replaced much of human decisionmaking. Systems make buy and sell orders based on an analysis of trading inefficiencies or financial differentials.[18] Powered in some cases by advanced computing, these tools have much greater capacities for storing information and processing data.[19]

Fraud detection is another area where AI is helpful in financial systems. It sometimes is difficult to discern fraudulent activities in large organizations, but AI can identify early or telltale abnormalities, outliers, or deviant cases. These detection systems help human managers find problems before they reach dangerous levels.[20]

For some individuals, however, AI creates consumer financial protection problems. Digital tools alter the retail experience and can expose consumers to deceptive marketing strategies or fraudulent practices. High-tech appeals may target people with limited financial literacy and encourage them to buy products for which they are not well suited. According to Brookings Institution scholar Makada Henry-Nickie, "AI can lead to a surge of wicked, legacy problems: product steering, discriminatory pricing, unfair credit rationing, exclusionary filtering, and digital redlining."[21] All of those practices harm consumer well-being and exacerbate financial inclusion problems.

In addition, AI creates the means by which financial institutions can bypass conventional nondiscrimination rules. For example, Brookings Institution scholar Aaron Klein notes that banks are prohibited from incorporating marital status into their lending decisions. Yet, algorithms can gauge the potential for marriage problems through proxy measures that examine travel, hotel, gift, and restaurant bills and therefore estimate the odds of impending divorces that could affect the ability to repay loans. When AI has the ability to assess relationship problems through indirect means, Klein argues it should be limited in the same way decisions based on marital status would be in lending decisions.[22]

Shopping

Many retail establishments have pioneered technologies designed to improve the consumer experience. Amazon, for example, has opened convenience stores with no salesclerks. A consumer downloads the company's mobile app, enters the store through turnstiles, picks out desired items, and exits through turnstiles that charge the person's credit card or mobile payment system.[23]

Giant Food Stores has introduced robots powered by AI and equipped with cameras and voice-paging capabilities that can spot food spills and empty shelves. These devices page human employees to let them know of the problem so that debris can be cleaned up and items restocked. The robots can also do price checks and answer customer questions through audio interfaces.[24] Company executives hope emerging technologies will make it easier for shoppers to find what they want and speed the checkout process. Such systems furthermore offer the virtue of instantaneous updates on inventory control and reorder points, thereby facilitating shelf restocking.

Walmart is adding 3,900 advanced robots to its retail stores. These devices look for shelves that need restocking, trucks that can be unloaded, or floors that need to be cleaned, with the goal of improving business operations. The robots are being spread across the company's 4,700 stores to enhance customer shopping and make sure there are fewer barriers to people's purchases.[25]

Yet there are concerns these robots will take human jobs and invade personal privacy. For example, some companies use smartphones to track consumer movements and shopping preferences while inside a store. They can use this information to tailor ads for specific people or learn what products are of greatest interest to the shoppers. Those practices raise a number of questions regarding the ethics of ad targeting and shopper profiling.

Smart Cities

Metropolitan governments are using AI to improve urban service delivery. According to Brookings Institution writers 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."[26]

Since the Cincinnati Fire Department fields 80,000 requests each year, city 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 and systematic in how they provide urban services. So far, city officials have been pleased with the deployment and feel it has put them in a better position to manage emergency services.

A number of metropolitan areas are adopting smart city applications that use AI to improve 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 of smart city solutions. Seattle, for example, has embraced sustainability and is using AI to manage energy usage and resource management. Boston has launched "City Hall to Go," a mobile

app 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 gunshots."[27]

Through these means, metropolitan areas are leading the country in the deployment of AI solutions. According to a National League of Cities report, 66 percent of American cities are investing in smart city technology. Among the top applications are "smart meters for utilities, intelligent traffic signals, e-governance applications, Wi-Fi kiosks, and radio frequency identification sensors in pavement."[28]

But in some places, these innovations have raised concerns regarding the collection of personal information and how city officials are deploying technology. Even worse are fears that technology contributes to "urban fragility." As noted by Kevin Desouza and David Selby, digital innovations sometimes can disrupt financial safety nets, weaken societal connections, and thereby undermine the ability of communities to deal with problems.[29]

Energy and Climate Change

One of the important ways AI helps is through improved analysis of energy resources and climate change. Estimating consumer and business demands for energy has always been challenging, and that limitation creates major challenges for the sector. Advanced analytics make it possible to model supply and demand in a granular manner and thereby project needs and consequences more clearly.

The same is true in regard to climate change. There are many factors that contribute to climate shifts, and AI enables scientists to analyze these dynamics very effectively. For example, they can examine the root sources of emission levels and project possible impacts over the next several decades. It may be impossible to stop the warming of the planet, but AI puts humans in a better position to manage and to mitigate some of the negative ramifications of climate change.

According to David Victor of the University of California at San Diego, "AI helps make markets more efficient and easier for analysts and market participants to understand highly complex phenomena—from

the behavior of electrical power grids to climate change. But AI itself won't assure that outcome without clear policy incentives."[30] To take advantage of the possible benefits, it is necessary to provide consumers and businesses with relevant information and thereby enable them to make efficient energy decisions. Victor argues that relatively small policy "nudges" can generate significant cost and resource savings. Figuring out how to manage emissions is a key to mitigating their consequences.

Law Enforcement and Criminal Justice

In a number of places, AI is being deployed in law enforcement and criminal justice. For example, 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. Analysts have found that youth is a strong predictor of violence and being a shooting victim is associated with becoming a future perpetrator.[31]

Judicial experts claim AI and data analytic programs reduce human bias in law enforcement and lead to a fairer sentencing system. For example, R Street Institute associate Caleb Watney writes that "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."[32]

Others hope AI can improve courtroom outcomes through natural language processing of case material, evidence, legal opinions, and judicial verdicts. Judges could use the insights derived from big data analytics to determine flight risks, set reasonable bail, find legal precedents for similar crimes, assist in sentencing decisions, reduce racial bias, and provide broad guidelines for pronouncing verdicts.

However, critics charge 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."[33] Others worry about a "police industrial complex" that deploys tools that target people of color unfairly and do not help cities reduce crime rates.[34] Some analysts, such as Rashida Richardson, Jason Schultz, and Kate Crawford of the AI Now Institute, go even further and argue that unrepresentative crime data yield inaccurate predictions and ultimately lead to civil rights violations.[35]

GOVERNANCE, ETHICS, POLICY, LEGAL, AND GEOPOLITICAL ISSUES

AI's increasing penetration into many aspects of life creates tremendous opportunities for advancement and innovation. Yet, its algorithms also raise important governance, ethical, policy, legal, and geopolitical challenges. For example, who should make decisions about emerging technologies: software designers, corporate executives, policymakers, or consumers? What types of ethical problems are introduced through software programming, and how transparent should developers be about their coding choices? How can we guard against biased or unfair data used in algorithms? How can we protect personal privacy and security, and guard against offensive material? What about questions of legal liability in cases where algorithms cause harm? How does AI affect geopolitical relationships around the world and the ability to handle large-scale changes arising from pandemics, income inequality, and climate change? How do we maintain human control over advanced technologies?[36]

Governance

Emerging technologies raise governance questions about who should decide issues of innovation, deployment, and remediation in case of consumer harm.[37] In the old order, national governments were the relevant decisionmakers on basic questions about public policy and society. Leaders would pass laws and enact regulations designed to address how innovations from the telegraph and telephone to television and nuclear energy

were handled. Nation-states in a Westphalian system were the primary locus of public debate and deliberation.[38]

Today, however, that system is in decline. Many technology decisions have migrated from the world of government to private companies, which in some cases act as proto-states in and of themselves. As an illustration, Facebook now can be considered a digital sovereign nation with over 2 billion people. The same is true for large firms with dominant power over particular domains. Their coders, engineers, and computer scientists make decisions all the time that affect the way people communicate, what information is at their disposal, how they buy products, and the manner in which democracy functions. Few of their choices are subject to detailed government oversight, since, until recently, many countries have taken a hands-off stance regarding private-sector technology innovation.

However, there are now raging debates on whether countries should maintain a hands-off stance concerning new technologies. Should leading digital platforms be subject to government rules and regulations, similar to the way companies are treated in other sectors? Should private companies retain the considerable autonomy they currently hold? To what extent should public concerns constrain corporate practices?

As AI advances, these kinds of governance issues must be resolved. With the acceleration of technology innovation, people must determine who sets the societal rules and whether legislators, regulators, and consumers should take more affirmative stances on how technology affects individual people and societal well-being. As Microsoft president Brad Smith notes in his book with Carol Ann Browne, "rather than fret excessively about the dangers of overregulation, the tech sector would be better served by thinking about what shape intelligent regulation should take."[39]

Ethics and Transparency

Algorithms embed value choices into program design and therefore raise questions concerning the criteria used in automated decisionmaking. Susan Etlinger of the Altimeter Group has noted that "algorithms aren't neutral; they replicate and reinforce bias and misinformation."[40] Given

this situation, it is important to have a better understanding of how software functions and what choices are made.[41]

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, and allow credit ratings of individuals based on unfair criteria. The types of considerations that go into programming decisions matter a lot in terms of how organizations operate and the manner in which software affects customers.[42]

Having data that are "accessible for exploration" in the research community can identify problematic features and lead to software that is trustworthy and serves the public good. According to a McKinsey Global Institute study, nations that promote open data sources and data sharing are the ones most likely to see AI advancement. In this regard, the United States has a substantial advantage over other nations. Global ratings on data openness show that the United States ranks 8th overall in the world, compared with 93rd for China.[43]

Biases in Data and Algorithms

In some instances, certain AI systems have enabled discriminatory or biased practices.[44] For example, Airbnb has been accused of having homeowners on its platform who discriminate against racial minorities. A research project undertaken by scholars at 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."[45] That shows how bias can shape digital platforms and enable discriminatory practices on the part of online users.

Issues of racial discrimination also have come up regarding facial recognition software. Most such systems operate by comparing a person's face to images in a 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."[46] Unless the databases have access to diverse data, these programs perform poorly when attempting to recognize African American or Asian American features.

A research study undertaken by the U.S. National Institute of Standards and Technology found that a "majority of commercial facial-recognition systems exhibit bias." In looking at differences by race, gender, and age, analysts discovered that algorithms misidentified African Americans and Asian American faces "10 to 100 times more than Caucasian faces," that it had greater errors for women than for men, and that it "falsely identified older adults up to 10 times more than middle-aged adults."[47]

The inaccuracy problem reflects the fact that many historical data sets often do not represent current demographic diversity. As Buolamwini notes, such an approach risks repeating past inequities: "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."[48]

Privacy and Cybersecurity

As human activities move online, companies are accumulating extraordinary information about consumers, which raises important questions regarding how to maintain personal privacy. People worry that companies and outside firms have too much confidential information, the material is at risk of being hacked, and organizations use that data for their own purposes such as advertising, marketing, e-commerce, political persuasion, or public surveillance.

As an illustration, Axios published a series of articles in 2019 on what various companies know about their online users. Its reporters included commonly collected information from firms such as Amazon, Google, Facebook, Tesla, hospitals, and the internet in general. When shown across each of these domains, the scope of the data collection is quite extraordinary.

For Amazon, the list includes the following:

- "Everything you've bought [on Amazon.com], plus the things you have just put in your cart, or searched for, or added to a wish list. And all of your addresses, and the names and addresses of anyone you've ever sent stuff to.
- All the books you've read [through Kindle or Audible], plus how far into the book you got. Amazon also knows which books you have browsed or sampled, and what passages you've highlighted in Kindle.
- What you've watched, browsed, and searched for [on Prime Video]. What you've watched, browsed, and searched for [on Twitch].
- For customers with a paid recording plan [through Ring], Amazon stores videos for 30 to 120 days depending on location, or until a customer manually deletes the video.
- If you shop here [at Whole Foods], Amazon knows your grocery list.
- Alexa knows all the things you ask it."[49]

For Google, the information is quite extensive:

- "The terms you search for.
- The videos you watch.
- Voice and audio information when you use audio features.
- Purchase activity.
- People with whom you communicate or share content.
- Activity on third-party sites and apps that use Google services.
- The ads and content you view on Google's sites, as well as interactions with that content.
- Chrome browsing history you've synced with your Google Account.
- Location data, which Google can either gather directly via GPS data or infer from other sensors and data, including IP addresses, nearby Wi-Fi routers, and Bluetooth beacons."[50]

Facebook compiles considerable material as well:

- "When you create an account, Facebook asks for your name and birthdate, along with either a phone number or e-mail.
- Then there's all the information you give Facebook as you fill out your profile, potentially including schools, current and past occupations, relationship status, hometown and current city, as well as your physical address, birth name, website, and other social links.
- When you log in, how long you spend online and where you are logging in from—hence it can welcome you to new cities and suggest places to visit and eat (and also serve up local ads).
- Places you check in.
- The pages, accounts, and hashtags you connect with on Facebook—and not just who you are connected with, but how long you interact and for how long.
- Your contacts, if you choose to upload your phone book or call history.
- Things you buy directly from or through Facebook, but also things you may not think about, like the metadata from photos you upload.
- Your friends can tag you in posts and photos, which gives Facebook additional information.
- Messenger can collect information on whom you talk to, how often, and for how long, as well as phone history if users opt in.
- Facebook has tools that partner websites use to integrate with Facebook, including the inclusion of 'Like' and 'Share' buttons, as well as a tracking cookie known as Facebook Pixel.
- Facebook knows your location, even if you haven't directly given it permission to access your phone's GPS, by tracking the IP address of the phones, computers, and other devices you use to access its servers.

- Facebook also reserves the right to enhance its data trove by adding information from outside providers."[51]

Tesla (and other automotive companies) collects an enormous amount of material through its vehicles:

- "It knows your speed, your mileage, and where and when you charge the battery.
- It also monitors airbag deployments, braking, and acceleration, which helps in accident investigations.
- And it knows when Autopilot, Tesla's assisted-driving feature, is engaged or disengaged, and whether you have your hands on the wheel as you should.
- Teslas are constantly in record mode, using cameras and other sensors to log every detail about what they encounter while driving, even when Autopilot is turned off.
- This includes short video clips from the car's external cameras to learn how to recognize lane lines, street signs, and traffic light positions."[52]

Hospitals maintain detailed databases regarding medical symptoms, emotional well-being, treatments, insurance claims, and prescriptions.

- "Health records house more information than most people may realize.
- They contain all the obvious stuff: Height, weight, and age; every appointment, vital sign, allergy test, surgery, procedure, and scan; and any prescription drugs you take, or have taken in the past.
- Everything else you, your family, or your friends divulge to doctors also gets recorded. That could include describing your drinking habits, admitting responsibility in a car accident, sharing marital problems, or even sending a Christmas card.
- Whatever you tell them is fair game to go into your file."[53]

The internet contains a vast trove of information through search engines, data aggregators, and people-finder sites:

- "Home address
- Phone numbers
- Vehicle information
- Court records (like marriage, divorce, or arrest records)
- Income range
- Political affiliations
- Relatives (based on shared last names) and their info
- Roommates (based on shared addresses) and their info"[54]

The list of companies could go on to include banks, credit card companies, e-commerce sites, travel firms, hotels, restaurants, airlines, media sites, schools, universities, health providers, government agencies, and the like. But the point is clear: in a world dictated by AI and data analytics and with people engaging in so many online activities, the potential for privacy intrusions, data abuses, cybersecurity attacks, and outright bias or discrimination is quite high. Figuring out how to deal with these issues represents a major challenge for consumers, businesses, and government officials.

Offensive Applications

Discrimination and privacy intrusions are not the only risks in the digital world. There are "deepfakes," computer-created artificial videos or other digital material that show well-known people either saying or doing something that is highly offensive but that never actually happened. This could be prominent politicians seeming to endorse causes or people even though they have not done so, or business leaders in compromising situations that are actually fake or untrue.

An AI designer known as Alberto created grave consternation when he devised a "DeepNude" app that could "digitally undress women in photos." The software "used artificial intelligence to create the 'deepfake'

images, presenting realistic approximations of what a woman—it was not designed to work on men—might look like without her clothes."[55] The depictions were so realistic that they appeared genuine to many viewers. The backlash that ensued from the application was so strong that Alberto took down the app within a few days of starting to sell it for $50.[56]

These types of inventions show how far the technology has come and what the moral risks are. It is possible through current software to develop AI that is quite offensive and abusive. These systems go way beyond what nearly everyone would consider legitimate, but it is perfectly legal under most current jurisprudence.

One of the first states to recognize the problems of this kind of AI-generated images is Virginia. In 2019, it enacted a statute outlawing the distribution of nude "falsely created videographic or still image[s]" of actual people without their consent.[57] The hope was that clear penalties would discourage nefarious or exploitative behavior through virtual technology. Other jurisdictions are considering similar legislation to impose penalties on those transmitting these kinds of abusive digital images.

Legal Liability

In the newly emerging world of AI, there are questions concerning the legal liability of software applications. 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 uses a situation's facts and circumstances to determine liability, and these influence the kind of penalties that are imposed, ranging from civil fines to imprisonment for major harms.[58]

An Uber-related fatality in Arizona, for example, represents an important test case for legal liability. The state actively recruited Uber to test its autonomous vehicles and gave the company considerable latitude in road testing. Yet in the case of accidents, it is unclear who gets sued: the human backup driver, the state of Arizona, the Phoenix suburb where the accident took place, Uber, software developers, data analytics experts, or the auto manufacturer. Given the multiple people and organizations

involved in autonomous vehicles, there are many legal questions to be resolved in AI-directed harms.

In many areas, digital platforms have limited liability for what happens on their sites due to the provisions of the 1996 Telecommunications Act. This landmark legislation, enacted well before the growth of the internet, mobile phones, and digital platforms, placed major limits on consumers' ability to sue internet companies based on what happens on their platforms. The idea was that firms in a nascent sector should not be responsible for what others do on their sites. The hope was this freedom would encourage innovation and provide helpful products and services.

Since then, some digital companies have imposed additional restrictions on people's ability to sue. For example, Airbnb "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 legal rights, the company limits consumer protections and therefore curtails the ability of people to fight discrimination or other problems arising from unfair algorithms.[59]

Geopolitical Considerations

Debates concerning the future of AI are taking place in a rapidly changing landscape. Regional alignments are shifting at the very same time a number of countries are investing in AI, machine learning, and data analytics in hopes of gaining advantages in economic development, international competitiveness, and national security. There also are pandemics, high levels of inequality, climate change, demographic shifts, and a movement of people into urban areas, all of which affect how AI is utilized and viewed by the public.

Amid these large-scale developments, a research project undertaken by PricewaterhouseCoopers estimated that "artificial intelligence technologies could increase global GDP by $15.7 trillion, a full 14%, by 2030."[60] But these gains are not spread evenly around the world and likely will increase inequality. The report notes that the improvements include $7 trillion for 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.

To maximize the growth potential, several countries are making significant investments. Florence Parly, French Minister of the Armies, announced in 2019 her country was raising its AI defense investment to 1.5 billion euros (around US$1.83 billion) by 2022. It hopes to develop algorithms that will enable "electronic warfare and intelligence gathering; automatic image recognition; collaborative combat on land and in the air; autonomous navigation robots; cybersecurity; predictive equipment maintenance; anti-mine warfare; and decision and command support."[61]

Germany meanwhile is investing 3 billion euros (or US$3.66 billion) in AI that focuses on industrial applications. It seeks to use its strength in manufacturing and trade to develop AI-based systems that improve the efficiency and speed of operations. The country plans to loosen regulations and provide data that will help researchers and businesses develop new applications. It sees opportunities to boost smart manufacturing and automotive companies, among other sectors.[62]

The United Kingdom has established a 160-million-pound (about US$196 million) defense fund that would finance new applications in "sub-surface submarine threats, surveillance and reconnaissance capabilities, and command and control in land operations."[63] Its private sector also is investing heavily in health and medical technologies. According to the McKinsey Global Institute, the British are aiming for "better medical research and for improved diagnosis, prevention, and treatment of diseases."[64]

Israel has become a global leader in cybersecurity and surveillance equipment. It has developed AI solutions that integrate data, apply computer vision, integrate satellite imaging data, and assess this material in real time. In addition, Israel is pushing ahead in advanced technology for autonomous vehicles, health information technology, and environmental sustainability.[65]

India is prioritizing AI in hopes of furthering its economic development. A strategic plan developed by the National Institution for Transforming India (NITI) Aayog think tank concluded the country should focus on health, education, agriculture, smart cities, and mobility. Ac-

cording to planners, those areas had the potential for "high social impact" and were thereby worthy of national investment.[66]

Russia does not have much of a domestic AI sector, and researchers have it "ranked 20th in the world by number of AI startups," which limits its growth prospects.[67] However, the government aspires to overcome these shortfalls. In discussing technology, President Vladimir Putin noted "artificial intelligence is the future not only for Russia but of all of mankind. . . . Whoever becomes the leader in this sphere will become the ruler of the world."[68] To advance its development in this area, Russian officials have "released a 10-point AI agenda, which calls for the establishment of an AI and Big Data consortium, a Fund for Analytical Algorithms and Programs, a state-backed AI training and education program, a dedicated AI lab, and a National Center for Artificial Intelligence."

China meanwhile is investing enormous resources in AI. As noted by author Kai-Fu Lee, China is making rapid strides because it has set a national goal of becoming the global leader in this area by 2030.[69] With its large population and limited privacy restrictions, the country has the ability to train AI data on a massive scale. That allows it to scale-up innovations quickly and efficiently, both in domestic and military affairs. In recent years, China has boosted its financial support of AI research. In 2017, for example, China's State Council issued a plan for the country to build a domestic industry worth roughly US$150 billion by 2030.[70] Cities such as Shenzhen are providing up to US$1 million to support AI labs. China hopes AI will provide security, combat terrorism, and improve speech recognition programs.[71]

These are not idle moves on the part of the Chinese. The traditional interpretation that China lags American ingenuity needs to be updated in light of recent scientific advances. China has the second largest AI talent pool in the world. While the United States has 13.9 percent of AI researchers, China now has 8.9 percent. Chinese scientists also have made remarkable AI progress over the past decade. According to a 2018 Tsinghua University report, "China leads the world in AI papers and highly cited AI papers." And it has "more AI patents than [the] U.S. and Japan."[72]

A report by the Information Technology and Innovation Foundation documented the progress China has made. Researchers Rob Atkinson and Caleb Foote examined thirty-six indicators of science and technology innovation between China and the United States and found "on all the indicators China has closed the gap or, in some cases, extended its lead over the United States."[73] As an illustration, they looked at R&D investments, workforce development, the number of scientific articles, patents, trade, and manufacturing, and found the Chinese have advanced rapidly in every area. No longer is it a country that copies to succeed; rather, it has a skilled workforce with great technological expertise.

A McKinsey Global Institute study found "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."[74] Although its authors found China currently lags the United States and the United Kingdom in AI deployment, the sheer size of China's AI market gives that country valuable opportunities for pilot testing and future development. Along with the United States, it has great capacity for AI deployment and is moving ahead quickly on many fronts.

Human Control

The ultimate AI issue in an era of megachange is how to maintain human control in the face of digital technologies that are advancing rapidly in their intentionality, intelligence, and adaptability. Beyond all the issues of privacy, security, fairness, and geopolitics lies the fundamental concern that the world is at a major inflection point where machines are ascending in their ability to make decisions and humans are losing control. As AI advances, the fear is algorithms will make so many decisions that individuals will be supplanted by hidden or indecipherable code.

Writer Yuval Noah Harari puts it most bluntly through an equation he calls $B \times C \times D = HH$. By that mathematical formulation, he warns "biological knowledge multiplied by computing power, multiplied by data equals the ability to hack humans." His fear is that "the ability to hack humans: to create an algorithm that understands me better than

I understand myself, and can therefore manipulate me, enhance me, or replace me. And this is something that our philosophical baggage and all our belief in, you know, human agency and free will, and the customer is always right, and the voter knows best, it just falls apart once you have this kind of ability."[75]

Given these kinds of concerns, it is important to delve into concrete cases of AI deployment to see whether it is possible to retain human autonomy in the contemporary period. We will elaborate in later chapters on how we assess the opportunities and risks of health information technologies, AI for school assignments, autonomous vehicles, e-commerce, and autonomous weapons systems. The governance and ethical problems vary from sector to sector, so it is important to understand the particularities of each application.

But one crucial point to remember across all of these areas is that there are particular ways people can retain control and are likely to do so in the future. This includes policy decisions, regulatory actions, legal liability, corporate self-policing, and public opinion that demands reasonable safeguards. If we utilize policy, legal, regulatory, self-regulatory, and consumer tools, there is little reason to fear the technologies that are moving full steam ahead. After all, when HAL turned into a murderer in *2001: A Space Odyssey*, Dave deactivated him. Even with powerful supercomputers, humans have many ways to control technology. Our task is to figure out how to gain digital benefits while minimizing AI's detrimental, discriminatory, or dangerous features.