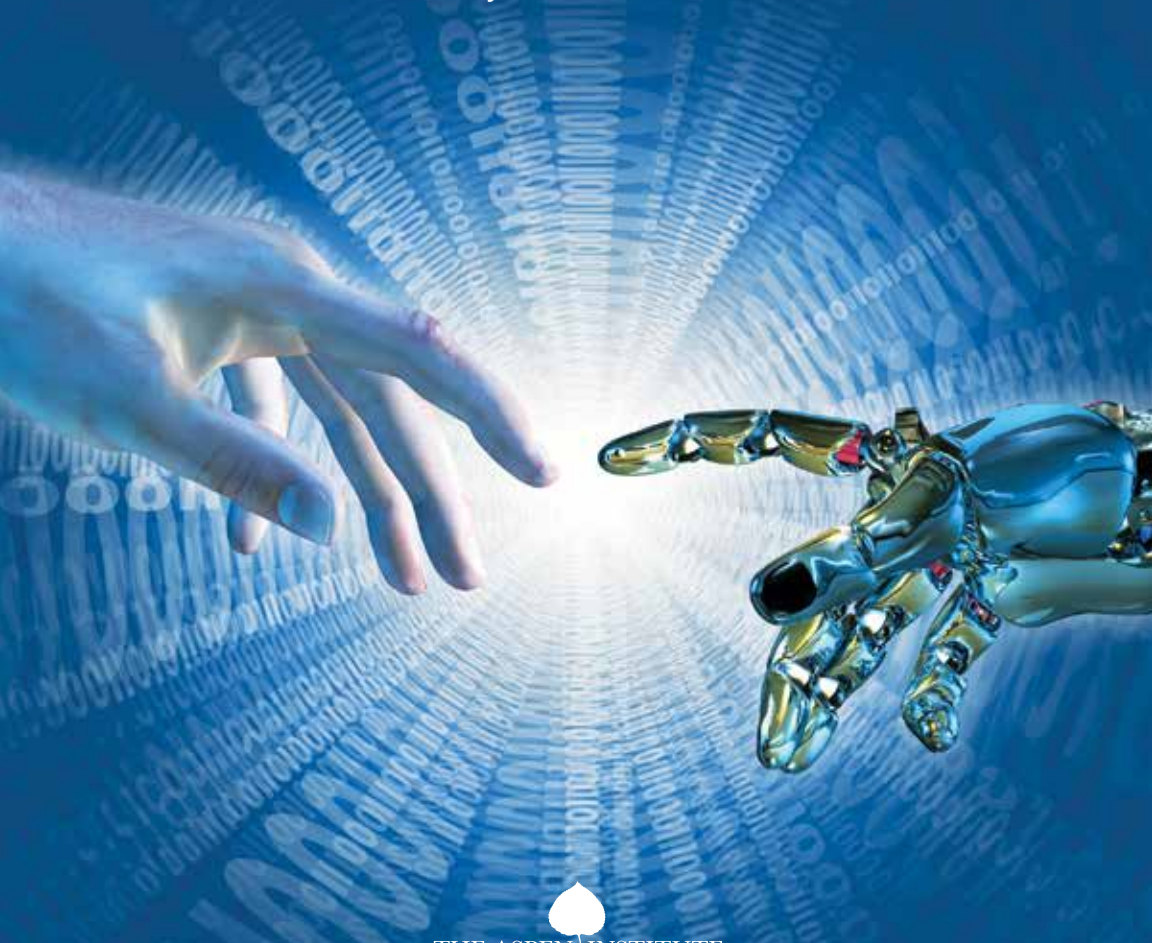
ARTIFICIAL INTELLIGENCE COMES OF AGE

The Promise and Challenge of Integrating AI Into Cars, Healthcare and Journalism

By David Bollier



THE ASPEN INSTITUTE

A Report on the Inaugural Aspen Institute Roundtable on Artificial Intelligence

ARTIFICIAL INTELLIGENCE COMES OF AGE

The Promise and Challenge of Integrating AI Into Cars, Healthcare and Journalism

David Bollier

A Report on the
Inaugural Aspen Institute Roundtable on Artificial Intelligence

THE ASPEN INSTITUTE

Communications and Society Program
Charles M. Firestone
Executive Director
Washington, D.C.
2017

To purchase additional copies of this report, please contact:

The Aspen Institute
Publications Office
P.O. Box 222
2014 Carmichael Road
Queenstown, Maryland 21658
Phone: (410) 820-5326
Fax: (410) 827-9174
E-mail: publications@aspeninstitute.org

For all other inquiries, please contact:

The Aspen Institute
Communications and Society Program
One Dupont Circle, NW
Suite 700
Washington, DC 20036
Phone: (202) 736-5818
Fax: (202) 467-0790

Charles M. Firestone
Executive Director

Patricia K. Kelly
Managing Director

The Aspen Institute
One Dupont Circle, NW
Suite 700
Washington, DC 20036

Published in the United States of America in 2017
by The Aspen Institute

Printed in the United States of America

ISBN: 0-89843-654-0

17/001

Contents

This report is written from the perspective of an informed observer at the Aspen Institute Roundtable on Artificial Intelligence.

Unless attributed to a particular person, none of the comments or ideas contained in this report should be taken as embodying the views or carrying the endorsement of any specific participant at the Roundtable.

Foreword

We stand at the doorstep of a significant resurgence of artificial intelligence (AI). The advances are driven by extraordinary computing power and a constellation of new technologies that range from machine-learning and neural networks to natural-language processing and knowledge representation. These AI technologies are pervasive and manifest in systems across industries throughout the world. Whether it be an AI assistant, self-driving cars, algorithmic newsfeeds, or playing chess, the widespread impact of AI systems provoke both intrigue and caution. How these artificially intelligent technologies impact our society, our economies, our governments and our well-being remains generally unknown.

It is within this space that the Aspen Institute Communications and Society Program convened its first ever Roundtable on Artificial Intelligence in August 2016. Twenty-two leaders and experts from industry, academia and the non-profit sector met and explored issues at the intersection of AI technologies, society, economy, ethics and regulation. The Roundtable encouraged participants to critically address the values that should inform AI development and to envision appropriate frameworks and public polices for its future. Specifically, discussions revolved around the promise and challenge of adopting AI technologies in three long-standing industries.

The following report, "Artificial Intelligence Comes of Age: The Promise and Challenge of Integrating AI into Cars, Healthcare and Journalism," authored by David Bollier, traverses the realities of AI technologies and its impact on these three important sectors. As participants noted, AI systems make possible new efficiencies and spur innovation; but, they also threaten our normative ideas of work, social customs, ethical boundaries and regulatory oversight.

The report is divided into four sections. First, "AI and Self-Driving Cars," provides an overview of the "very public character" of self-driving cars, which highlights public anxieties surrounding safety of AI systems and tensions between market innovation and regulation. Second, "AI and Healthcare," considers AI's transformative impact on medical research, patient diagnoses and treatment options, highlighting big data

and deep learning techniques. Notably, a discussion on structural barriers to expanding AI in healthcare features problems ranging from data islands to privacy. The third section, "AI Begins to Change Journalism and News Organizations," examines the changing dynamics of the news business in light of new digital platforms and outdated revenue models. As the journalism tradition declines, so does its role as a vital democratic institution and resource for civic information. Lastly, the report ends with "AI Bound or Unbound," which envisions the future of AI as determined by who controls it and who will guide its development.

Acknowledgments

On behalf of the Aspen Institute Communications and Society Program, I want to thank Michael Ferro and the Ferro Institute for their leadership and support in developing this roundtable. Thanks, also, to David Bollier, our rapporteur, for capturing the various dialogues and nuanced viewpoints into this report of the inaugural Roundtable on Artificial Intelligence. As is typical of our roundtables, this report is the rapporteur's distillation of the dialogue. It does not necessarily reflect the opinion of each participant at the meeting. Finally, I want to thank Jennarose Placitella and Kristine Gloria, Project Managers, and Tricia Kelly, Managing Director, for their work on the conference and bringing this report to fruition.

Charles M. Firestone
Executive Director
Communications and Society Program
The Aspen Institute
January 2017

Artificial Intelligence Comes of Age

David Bollier

Artificial Intelligence Comes of Age

The Promise and Challenge of Integrating AI Into Cars, Healthcare and Journalism

A Report on the Aspen Institute Roundtable on Artificial Intelligence

David Bollier

In recent years, the development of new artificial intelligence technologies has been surging at unprecedented speeds. The innovations are introducing countless new disruptions and uncertainties into everyday life, commerce and public policy as well as new efficiencies, practical solutions and markets. The reverberations are already being felt acutely in certain fields where AI technologies are quite advanced, especially self-driving motor vehicles, healthcare and the news media.

While the advances are exhilarating to many, they also pose significant challenges in terms of their social, economic, legal, ethical and even political dimensions. The technologies make possible some incredible new feats in transportation, urban life, medical research, healthcare and journalism. They have the potential to create new markets, spur new efficiencies and extend human capabilities in remarkable ways. But AI systems may also eliminate millions of jobs, cause social disruptions and require significant updates to existing systems of law and regulatory oversight.

In an effort to take stock of some vanguard sectors of AI innovation, the Aspen Institute Communications and Society Program convened the first annual Roundtable on Artificial Intelligence on August 1 and 2, 2016. A special emphasis was put on addressing the values that should animate development of AI technologies and how to develop appropriate policy responses. The conference, held in Aspen, Colorado, brought together twenty-two leading AI technologists, computer industry executives, venture capitalists, and academics who study technology (see Appendix for a list of participants).

The dialogues were moderated by Charles M. Firestone, Executive Director of the Communications and Society Program. The report that follows, written by rapporteur David Bollier, is an interpretive synthesis that distills the key themes, insights and points of consensus and disagreement that emerged from the conference.

Putting Artificial Intelligence Into Perspective

To put current debates about AI into an historical context, Walter Isaacson, President and Chief Executive Officer of the Aspen Institute, and the author of books on Steve Jobs and the history of computing,[1] pointed to a recurrent public debate about technologies that has been going for nearly 200 years: Will machines replace human beings with superior performance, rendering them irrelevant, or will machines assist and augment human intelligence, surpassing what machines can do on their own? Will technology create more jobs and prosperity, or will it lead to a net loss of jobs, economic decline and social unrest?

This debate goes back to the 1840s, said Isaacson, when a woman who arguably pioneered computer programming — Ada Lovelace, the daughter of Lord Byron — wrote a seminal scientific article about the nature of computing. As Isaacson wrote, "It raised what is still the most fascinating metaphysical topic involving computers, that of artificial intelligence. Can machines think?" Lovelace emphatically rejected this proposition with an argument that has come to be known as "Lady Lovelace's Objection: The Analytic Engine [a computer] can do whatever we know how to order it to perform. It can follow analysis; but it has no power of anticipating any analytic relations or truths."[2]

Computer scientist Joseph Weizenbaum famously illustrated the limitations of AI in the 1960s with the development of the Eliza program. The program extracted key phrases and mimicked human dialogue in the manner of non-directional psychotherapy. The user might enter "I do not feel well today," to which the program would respond "Why do you not feel well today?" Weizenbaum later argued in "Computer Power and Human Reason" that computers would likely gain enormous computation power but should not replace people because they lack such human qualities and compassion and wisdom.

At the dawn of personal computing, in the 1970s, the two poles of this debate were personified by Steve Jobs and Bill Gates, tech rivals

with conflicting views about AI, according to Isaacson. Gates saw machines as capable of mimicking exactly what humans do, and surpassing us, ultimately rendering humans unnecessary. Jobs disagreed — along with computer visionary Doug Englebart — believing that the humanities and technology will always work together, and in that human/machine collaboration is a more fruitful avenue for the development of AI. This viewpoint is reflected in the work of the noted computer scientists Marvin Minsky and Seymour Papert, Isaacson added, noting that this school of thought is carried on by the M.I.T. Media Lab.

Confidence in the prospects of artificial intelligence have ebbed and flowed as federal and corporate funding was slashed, responding to a consensus in tech circles that AI would not work, at least in the ambitious ways previously imagined.[3] So-called "AI winters" ensued in the late 1960s and early 1970s and again in the late 1980s, which saw a deep decline in expectations, investment and new research approaches.[4] In the intervening years, interest in AI later picked up, occasionally punctuated by setbacks, with a strong surge of interest in AI technologies in recent years as practical commercial applications became more feasible.

Once again, the question is arising: Will AI assist and augment human intelligence or replace it? The answer implicates a related question: Will AI benefit society or harm it? Isaacson confessed, "I don't know how the story ends. That's partly what this conference is about. But I'm a little more on the Doug Englebart/Steve Jobs side of this debate" – i.e., that AI will work in tandem with humans, augmenting their capacities, and not supplanting them.

AI thinkers have themselves become more nuanced in their thinking, as reflected in Bill Gates' retreat from his former position. Gates, according to Isaacson, now concedes that the great mistake in AI was its faith that fully digital, binary, algorithmic functions embedded on silicon chips would somehow replicate human intelligence or consciousness. Gates now believes that AI scientists should attempt to reverse-engineer the way nature does things, and perhaps rely on some analog, carbon-based, "wetware" systems (the human brain linked to AI) as well.

There is another difference in today's debates about AI, said Isaacson. This time, there are some serious economists, led by Harvard economist Larry Summers, who says that this time things may be different.[5] It is

quite possible that technology will not create a net gain of jobs over the long term. This time, many people may be put out of work permanently, with no net increase in per capita gains in jobs, productivity and economic growth.[6] Of course, others remain convinced that AI systems will boost overall productivity and jobs, as other technological advances have in the past.

"What skills and intelligence are distinctly human? What do we bring to the party? And what do machines bring?" – ***Walter Isaacson***

This debate takes on a different flavor today because we live in a different historical moment and computing technology has a significantly different character. So we revisit the old questions, said Isaacson: "What skills and intelligence are distinctly human? What do we bring to the party? And what do machines bring?"

AI and Self-Driving Cars

Although there have been numerous experiments with autonomous vehicles over the decades, especially in the 1980s, the demonstration of autonomous cars on public roads in the past several years — especially the Alphabet self-driving car and the Tesla Motors Autopilot — has quickened interest in the technology. The Google prototype, released in May 2014, is fully autonomous, and notably does not have a steering wheel, gas pedal or brake pedal. The Tesla Autopilot, released in October 2015, can function autonomously on limited-access highways, but still requires drivers to be prepared to take control, as necessary, because the vehicle cannot detect lane markings, pedestrians or cyclists, and cannot shut itself off.

The arrival of these self-driving prototype cars has triggered a spirited debate about the implications of the technology and how its development should proceed responsibly. To give an overview of the current state of debate about AI cars, Dr. Astro Teller opened the first session of the conference with a brief presentation. Teller currently oversees

the company X, which is Alphabet's "moonshot factory for building magical, audacious ideas that through science and technology can be brought to reality." (Alphabet is the parent company of Google.) For Teller, the public debate about autonomous cars is important because "self-driving cars amount to a microcosm of the whole AI debate, which means that there is more at stake than just self-driving cars."

What distinguishes self-driving cars may be their very public character, said Teller. While there is a lot of AI work going on today, very little of it is as public or focused as self-driving cars. One reason for public interest may be the obvious safety implications. Unlike other fields of AI research, "it's not hard to make the argument that someone could get hurt if it goes wrong," said Teller, "and there is no field where it's as easy to make a mistake."

"Self-driving cars amount to a microcosm of the whole AI debate, which means that there is more at stake than just self-driving cars." ***– Astro Teller***

Teller noted that self-driving cars would be able to help many people. "There are more than a million people worldwide who die each year in car accidents," he said. He also noted the economic impacts: Traffic in the U.S. alone wastes hundreds of billions of dollars a year.[7] (The automobile fatality rate in the U.S. is 35,000 people per year.) A paper by two University of Texas researchers found that if 10 percent of the vehicles on the road were self-driving cars, there would be savings of more than $37 billion from lives saved, fuel savings, reduced travel time, and so on. The estimated benefits would exceed $447 billion if 90 percent of the vehicles on the road were self-driving.[8]

Self-driving cars could greatly reduce the burdens of commuting and fuel consumption. The cars will enable the elderly and disabled to get around more easily. Drunk driving could cease to exist. Self-driving cars could also hasten a shift to electric vehicles.[9] "So the upsides of self-driving cars are as self-evident as the downsides," said Teller.

Teller worries, however, that concern about the risks of self-driving vehicles could needlessly derail the technology and its widespread use.

While some people find the whole concept of self-driving cars alarming, he said, we need to remember that "airplanes already fly themselves." Regulatory and technological systems for managing large-scale, autonomous transportation —aviation — already exist.

Nonetheless, Teller agrees that the unique challenges posed by self-driving cars require careful attention: "We must make sure that the process for regulating self-driving cars goes well, because a lot of other robotics and secondary fields will follow the good or bad path that this one goes down." Teller hopes that automakers can demonstrate best practices and cooperation among each other, for example, in developing performance standards, and that a rigorous and flexible regulatory process can be established as soon as possible.

"Autonomous cars are going to happen, and they are the right thing….But everyone fights the pioneers." – ***Michael Ferro***

Teller and others in the room pointed out that the struggle of new technologies to gain acceptance is a familiar story. "Autonomous cars are going to happen, and they are the right thing," said Michael W. Ferro, Chairman of tronc, inc., and an investor and philanthropist, "But everyone fights the pioneers."

Conference participants identified a number of concerns about autonomous cars that must be addressed. The first set of issues (discussed below) involve the unsolved technical design challenges, most of which involve their safety. The resolution of these technical issues, in turn, are likely to affect how regulatory oversight and legal liability regimes are crafted.

Self-driving cars also raise a variety of secondary, indirect issues beyond the functioning of the car itself. Chief among them is the likely economic impact of eliminating jobs for drivers. Autonomous cars would also affect energy use, traffic patterns, urban design and real estate markets. Finally, AI-driven cars raise privacy and security questions: Who shall control the data generated by self-driving cars? And

can AI systems successfully prevent malicious hacker attacks that could surreptitiously seize control of a car?

The next sections look first at the technical issues that must be surmounted in making self-driving cars safe enough for universal use. Then, the report moves to consider the various ethical, social and legal issues that need to be addressed, and the challenge of devising appropriate regulatory oversight.

Technical Challenges of Self-Driving Cars

When talking about self-driving vehicles, it is not widely appreciated that there are different levels of autonomy. People accustomed to science-fiction or Hollywood depictions of self-driving cars imagine vehicles that automatically whisk a person to a desired destination with no involvement by people at all. In fact, there are different gradations of autonomous design.

The Society of Automotive Engineers published a system in 2014 to classify just how autonomous a vehicle is.[10] The National Highway Traffic Safety Administration formally adopted the classification criteria in September 2016. Cars at Levels 1 and 2 ("Driver Assistance" and "Partial Automation," respectively) have minor automation systems such as adaptive cruise control, "Lane Keeping Assistance" and automated acceleration, braking and steering. Drivers in Level 1 or 2 cars cannot sit back and relax, however; they must be ready to take control of the automated system at any moment to avoid hitting objects or to deal with real-world events.

At Level 3 ("Conditional Automation"), automated vehicles have a more significant degree of autonomous control. They can function on known, limited environments such as highways, enabling drivers to avert their full attention from driving. However, an occupant still must monitor the vehicle's operation, and may be required to take charge of it. A car with Level 4 ("High Automation") features is almost completely autonomous. It can function in most driving environments except severe weather, and the driver need not pay any attention to the operation of the vehicle. A car with Level 5 capabilities ("Full Automation") governs all aspects of dynamic driving tasks all the time, under all roadway and environmental conditions, without any driver role.

The differences between Level 3 and Level 4 — and the perceptions of what each should be capable of doing — is a source of some controversy. Designers of self-driving vehicles realize that there is a big difference between Level 3 and Level 4 vehicles in terms of their capabilities, how drivers must interact (or not) with the car, and what sorts of regulation may be needed. But, the general public may not necessarily understand these differences in automation.

Stuart Russell, professor of computer science at the University of California Berkeley, believes that Tesla itself may have contributed to confusion about the extent of its vehicles' capacities by saying its prototype has an "auto-pilot" mode. Russell also cited a public comment by Tesla Motors' co-founder and CEO Elon Musk in 2015 that autonomous cars, "would be like an elevator. They used to have elevator operators, and then we developed some simple circuitry to have elevators just automatically come to the floor that you're at...the car is going to be just like that."[11]

Russell criticized the implied claims that the Tesla vehicle is a "fully autonomous vehicle" when in fact "that problem has not been solved yet….I worked on automated vehicles from 1993 to 1997, and it was clear that there was no [intermediate] place between 'smart cruise control' and full autonomy. So here it is twenty years later, and we keep going through this." Another participant believes that Tesla made an error in calling its car "auto-pilot" instead of "driver-assist," and that Musk's "elevator" comment needlessly caused confusion.

"We [the company X] built a car that didn't have a steering wheel in it because that was the only way we could teach our engineers not to trust humans as a backup system." ***- Astro Teller***

Highlighting an engineering design disagreement with Tesla, Dr. Teller explained why Google turned away from the development of a Level 3, "driver-assisted" vehicle similar to the Tesla's four years ago. Google engineers discovered that they could not reliably assure that on-board testers would constantly monitor the car's operation. They

concluded that humans could not serve as reliable backups to the AI system, and so they re-focused their efforts on building a fully autonomous, Level 4 vehicle. "We [the company X] built a car that didn't have a steering wheel in it," said Teller, "because that was the only way we could teach our engineers not to trust humans as a backup system."

Humans as the "Failure Point" in the Technology. The engineering design choices for Level 3 versus Level 4 cars raises a host of issues about the ways in which humans may need to interact with self-driving technologies. One troubling conclusion is that human beings — whether as drivers, pedestrians or cyclists — may be the key "failure point" in the technology.[12]

For Teller, the key question is, "How safe should a self-driving car be to make it a legitimate use case?" It is reassuring that more than two million miles of real-world test data show significant safety improvements, he said, but distressing that many users of Level 3 vehicles are not prepared to take the steering wheel if necessary. Some test users of Tesla vehicles have actually climbed into the backseat, which Teller regards as reckless and perhaps criminal.

Such behaviors suggest that user education may be an important priority going forward, said Professor Rao Kambhampati, President of the Association for the Advancement of Artificial Intelligence and a computer scientist at Arizona State University. Kambhampati argued that "some driver-assist technologies can actually increase the cognitive role for drivers" by engaging them. But for a small minority of drivers, including himself, driver-assist features such as cruise-control are more annoying than welcome. He suggested that designers of autonomous cars develop ways to deal with such variable human responses.

Can auto-pilot technologies actually diminish a driver's skills and alertness? That is a concern raised by Stuart Frankel, Chief Executive Officer of Narrative Science, a tech firm that generates natural language from data in enterprise settings. "The more that people use their semi-autonomous cars in autonomous mode, the more that their skills are going to atrophy," said Frankel. "If you look at airline pilots who are under 40 years old or so, their ability to effectively deal with an emergency is significantly lower than that of older pilots." The point is underscored by Maria Konnikova in her article "The Hazards of Going on Autopilot" in *The New Yorker*:

> As pilots were becoming freed of responsibilities, they were becoming increasingly susceptible to boredom and complacency — problems that were all the more insidious for being difficult to identify and assess. As one pilot…put it, "I know I'm not in the loop, but I'm not exactly out of the loop. It's more like I'm flying alongside the loop."[13]

Marc Rotenberg, President and Executive Director of the Electronic Privacy Information Center, raised a similar point, noting that the U.S. Naval Academy is now requiring young cadets to learn celestial navigation for the first time in twenty years. "They are anticipating failures of GPS," he said, referring to the Global Positioning System, the navigation satellite technology that provides location and time information in all weather situations. "When you're at sea, on a boat, entirely dependent on GPS, what do you do if GPS fails?" asked Rotenberg. "I'm sure that a lot of judgment went into this decision [to bring back the teaching of celestial navigation]." Should this line of thinking be applied to self-driving cars as well?

For Joi Ito, Director of the MIT Media Lab, there is no evading the fact that humans will have to co-evolve with new technologies, and over time, reach a stable rapprochement and familiarity with them. "I have a Tesla X," said Ito, "and when driving I know exactly when [the driver-assist function] should be on and when it should be off. The training of humans comes from using a technology, in an iterative process. Each community of users is going to be different. You will have a co-evolution of humans and technology as people become accustomed to knowing more about the limits of the machine, and then they will start to trust it."

Can AI Engage with Tacit and Dynamic Social Factors? While technology and humans will surely have to co-evolve, a deeper question may haunt the future of AI: Can it accommodate irregular driving practices and social norms, many of which are tacit, subtle, idiosyncratic and dynamic?

Conference participants pointed out that driving is a social act and tradition that varies immensely from one culture to another. "I think it would be easier to teach a car to drive in Japan, where people tend to

follow the law, versus a country like, say India, where you don't expect drivers to follow the law," said Joi Ito of the MIT Media Lab. Ito cited a design proposal about how a self-driving car trained in England "would have to be put into quarantine before being allowed to drive in another country, because driving on the street is really about figuring out how people are going to react, and not about following the law."

"What is the ability of AI to understand human foolishness?" - ***Father Eric Salobir***

AI scientists are not unaware of these tacit, cultural dimensions of driving. The technology to "generate that [social] handshake is here today," said Ken Denman, an entrepreneur and former Chief Executive Officer of Emotient, a tech startup that uses computer vision and behavioral and cognitive science to predict emotions. Denman said that computers, data systems and cameras can be used today to locate faces and interpret the meanings and emotions that are being expressed. The camera can make a prediction as to "Is that person looking at me? Are they engaged?" That data is available in real time today. The question is, "Is there some need for the car to signal the pedestrian?"

Astro Teller said that the Google car is addressing such issues as well. "We spend much of our total engineering time modeling issues such as a bicyclist waggling his hand." Teller thinks that these are "temporary problems. We mostly don't actually get signaled by the drivers of other cars. We just think we have a good model of what they're like. And we don't yet have a model for what self-driving cars themselves will be like or what they will do. Once they've been out for twenty years, it will be fine."

Beyond the cultural quirks of driving, the deeper question may be, "What is the ability of AI to understand human foolishness?" said Father Eric Salobir. Salobir is a member of the Order of Preachers (known as Dominicans) and President of OPTIC, a network that promotes the digital humanities. He elaborated: "We should not be assuming some future world where everything is easy because everything is rational. We should assume the high level of irrationality that

currently exists." This insight may be especially important when we think about "rational" autonomous cars sharing the road with unpredictable human drivers, he said.

AI engineers are well aware of the tension between a world governed by formal rules and the messy realities of "real life," said Astro Teller: "We [at X] have discovered that people are dangerous around our cars because our cars follow traffic laws. But people are so bad at following laws that they don't expect that a car on the road next to them will actually do what the law says it should do. This puts us in this weird quandary."

To move beyond the formal rules of an AI system, even one that is capable of learning and evolving, requires moving beyond what Wendell A. Wallach calls "bounded morality." Wallach, an author and scholar at the Interdisciplinary Center for Bioethics at Yale University, notes that while many traffic rules are clear and self-evident — you stop at a stop sign, you brake when you see a child's ball bouncing near the road, i.e., examples of bounded morality — other rules are highly situational and open-ended. In other words, social practice contrasts with purely automated programming. "Driving is actually a social practice," said Wallach. "A classic example is when four cars come to a four-way stop at the same time. Which one should go first? People give each other social cues such as looking at each other, nodding, or nudging their car forward to establish who should go first. We don't know how to program an understanding of these social practices into driverless vehicles."

Part of the problem may be in conceiving of cars as "autonomous," said Cynthia Breazeal, Associate Professor at the MIT Media Lab and Founder of JIBO, Inc. "Human driving is a very collaborative social process," she said. "There is a lot of signaling of intent, and not just a following of rules." It may be better to view the challenges of autonomous driving as a "collaborative teamwork problem, where the car is part of a team of humans who are in the car, pedestrians walking on the side of the road, and drivers in other cars," said Breazeal. Such a framing of the challenge can help us "think about the interfaces we need and how to design the signaling of intentions, which are fundamental to the intuitive ways that people drive."

For example, eye contact between a driver and a pedestrian can be a way of signaling intent, helping decide which car at an intersection is

going to proceed first. Breazeal said that the problem with autonomous cars is that people cannot signal their intent to it, and it cannot read or anticipate what actual human beings will do. "I can't signal to it" and thereby establish some measure of trust, she said, "and I can't understand what the car may be signaling to me. These questions are really worth thinking through."

Humans are not the only unpredictable factor. As computer pioneer Norbert Weiner said many years ago: "As machines learn they may develop unforeseen strategies at rates that baffle their programmers…. By the very slowness of our human actions, our effective control of machines may be nullified. By the time we are able to react to information conveyed to our senses and stop the car we are driving, it may already have run head on into a wall.… Therefore," Wiener advised, "we must always exert the full strength of our imagination to examine where the full use of our new modalities may lead us."[14]

The Ethical Design of Autonomous Vehicles

What do these realities mean for the ethical design choices of autonomous vehicles? By contrast, the real-life choices and behaviors of human drivers are arguably more unpredictable, improvisational and perhaps even unknowable because of the welter of situational factors and cultural predispositions at play.

An often-invoked ethical scenario for self-driving cars is whether a car should "choose" to hit a baby carriage careening into the road or instead swerve into a trolley filled with nuns. The algorithmic design of the car supposedly makes such ethical choices inescapable.[15] While there are indeed ethical choices to take seriously, Astro Teller considers such scenarios removed from everyday reality. "When you give a human a driver's test, you don't ask them, right before you hand them the driver's license, 'Are you going to hit the nun or are you going to hit the baby?' People say, 'Jeez, I'm going to drive really safely.'"

The comparisons may be moot, suggested Father Eric Salobir, because drivers may or may not actually exercise moral judgment in such situations. Split-second driving decisions are not necessarily moral choices in any conventional sense, he said. "When something happens on the road, you react instinctively. It's not morality. It's just survival — an instinct."

By this logic, replied Astro Teller, "It's immoral for humans to be driving at this point, if they don't really have time to choose." Teller suggested that if alternatives to current driving practices could save lives and function more safely, then the proper "moral choice" is to use the alternatives: "Imagine if it turned out that robots could do some surgical operation with half the mortality rate of human surgeons. Would we let surgeons continue to do it? No, it would be immoral." Teller suggested that the same reasoning might be applied to self-driving cars versus conventional cars. What matters is "having that conversation in a functional way with regulators, and getting away from this 'nun versus baby' nonsense, which is not useful because that's not how AI works," he said.

"These are social problems….So who will be the good-faith brokers who can create a framework within public policy to establish some norms and trustworthy outcomes?" ***– Wendall A. Wallach***

For Wendell A. Wallach, a bioethicist, "Programming the self-driving car to save the most lives in an accident (short-term utilitarian calculation) even if that meant killing the car's passengers, could lead to more deaths in the long run (long-term utilitarian calculation) if that meant that people would not buy such a car. In other words, to minimize the harm from a once-in-a-trillion mile accident, we could lose many more lives because people won't buy a car that might kill them."[16] And without consumer acceptance, there might never be a market for self-driving cars that could save tens of thousands of lives.

The critical question for Wallach is what sort of philosophical or conceptual framework will be used in making necessary ethical and policy choices. "These are *social* problems," Wallach insisted. "These are problems of social practice. We need to be establishing social norms. So how do we go about doing that, and who can you trust to do that? People don't trust government to do that. So who will be the good-faith brokers who can create a framework within public policy to establish some norms and trustworthy outcomes? Who?"

Other Policy Concerns Raised by Autonomous Cars

Beyond safety, there is a raft of other social, economic and policy concerns that self-driving cars are raising. These include:

Liability. Who's responsible for any harm that the cars may cause? According to Teller, test data clearly show that cars with auto-pilot technology engaged are safer than conventional cars. This would be a huge improvement, but liability issues would remain, particularly for the potentially fraught real-world interactions between autonomous cars and human-driven cars.

Wendell Wallach believes that "autonomous technologies threaten to undermine the foundational principle that there is an agent, either human or corporate, that is responsible and potentially culpable and liable for what can go wrong." Teller agrees that liability issues are complicated, at least for Level 3, driver-assisted cars. However, perhaps the issue is more straight-forward for Level 4 cars.

An ideal scenario for implementing such complete liability would be a city that entirely converts to autonomous cars, thus avoiding the messy, unpredictable encounters between Level 3 cars, Level 4 cars and conventional cars. Under some scenarios, the city of the future may wish to ban all conventional privately owned cars, converting automobile transport in cities into a service.

The cybersecurity of cars. An abiding problem for self-driving cars is their cyber-security. "We have to consider third parties who may be intent on causing significant harm by hacking into systems and disabling them," said Marc Rotenberg of the Electronic Privacy Information Center, calling the problem a "huge soft target." In a remarkable demonstration of this fact, a group of "white hat" hackers in September 2016 took control of a Tesla Model S vehicle from twelve miles away, unlocking the car and activating the brakes, bringing the car to a stop.[17] Malicious hacks are likely to be an ongoing risk of self-driving cars.

Data privacy and due process. Because self-driving cars will generate and store vast quantities of data about driving behavior, control over this data will become a major issue, especially from a driver's perspective. Following a crash or criminal allegation, for example, will the data belong to the manufacturer, to be used as forensic evidence to defend itself, or will the driver have full and exclusive access to the data? Marc

Rotenberg suggested that the issue is not simply about privacy, but also about due process rights and fairness. Increasingly, states are declaring that people should have the right to know what data is being collected about them, and to be informed about how information may be legally used in the future.[18]

The issue is not simply about privacy, but also about due process rights and fairness.
– Marc Rotenberg

To help clarify what this process should entail, Rotenberg proposed two additions to Isaac Asimov's famous "Three Laws of Robotics," a proposed set of fundamental ethical behaviors that all robotic systems must implement.[19] Rotenberg said two new rules should be added: "The machine must always provide the basis for its decisions," and "A machine must always reveal its identity." These rules are likely to become more important as autonomous devices (cars, drones, other) begin to proliferate across the social and geographic landscape.[20]

Urban design and planning. Self-driving cars will have multiple transformational effects on the life of cities. It is unclear whether "robot taxis" and other autonomous cars will decrease traffic by reducing the fleet of cars in a city, or whether it will encourage more people to simply keep their cars circulating because parking spaces are so expensive. It is possible that a city may want to develop more dedicated bus lanes, segregated bike lanes and pedestrian paths if autonomous cars come to dominate city streets.

Real estate values may well shift as transportation patterns shift. Urban planners already know that public transportation boosts real estate values in the areas it reaches, and reduces values in less accessible areas. Uber cars are already having an effect on real estate markets, said Teller; autonomous cars will likely intensify this impact. For example, storefronts without adequate street parking — which normally would prevent their use as restaurants — could suddenly be feasible, which could raise the value of such properties by two or three times. Uber recently persuaded the City of Pittsburgh to give it free rein to experi-

ment with autonomous vehicles within the city, which may yield some insights into this issue.[21]

Economic disruption and jobs. One of the biggest unresolved policy questions is how to deal with the economic disruption that would affect millions of taxi drivers, chauffeurs and truck drivers whose jobs could potentially be eliminated by self-driving vehicles. The first self-driving truck began testing in the deserts of Nevada in May 2015, and firms such as Daimler and Otto (a startup launched by two former Google engineers) are now attempting to perfect the technology. Morgan Stanley predicts completely autonomous capability by 2022 and massive market penetration by 2026.[22]

While autonomous trucks could surely help reduce the 330,000 large-truck crashes that killed nearly 4,000 people in 2012, they could also eliminate the jobs of 3.5 million truckers and an additional 5.2 million non-drivers employed within the trucking industry. The technology could also threaten the jobs of millions of people who work in restaurants, motels and truck stops that service truck drivers, with community ripple effects flowing from those job losses. The appropriate policy responses to such developments — re-training? a basic income? — are relatively unexplored.[23]

The cultural appeal of autonomous cars. It remains to be seen whether the American people will embrace autonomous cars. For a country raised on the ethic of "hitting the road" and the association of cars with personal freedom, it could be hard for many Americans to "move from a world of personal autonomy to one of device autonomy," said Marc Rotenberg.

What Type of Regulatory Oversight Is Needed?

There was a general consensus among conference participants that self-driving cars would create new regulatory challenges. But what specific sort of regulatory oversight is needed, and what structural principles and procedures should guide it?

One significant answer to these questions arrived a month after this conference, in September 2016, when the National Highway Traffic Safety Administration (NHTSA) announced voluntary federal guidelines for self-driving cars.[24] Automakers will be allowed to self-certify

the safety of autonomous vehicles based on a fifteen-point checklist for safety design and development. While the guidelines are not mandatory or enforceable, federal regulators expect compliance.

According to news reports, some consumer advocates object to the voluntary guidelines, preferring a formal rule-making process that would have provided greater opportunity for the public to register its views.[25] But such a process would likely take years, say the makers of autonomous vehicles, who are eager to move forward rapidly; many critics believe that NHTSA authority will be extended. For the time being, the federal guidelines will likely deter states from enacting their own laws for autonomous vehicles, except for liability standards, which have long been a state concern.[26]

While many makers of autonomous vehicles welcomed the NHTSA policies, the character of regulatory oversight is sure to evolve in the years ahead. The guidelines provide a general framework and assert a federal role, but many still-emerging issues will need to be addressed as self-driving cars move towards commercial sale and actual use at large scales. It is therefore worth reviewing some of the conference discussion about regulation, despite its occurring prior to the NHTSA announcement.

In thinking about how autonomous transportation should be regulated, an obvious analogy comes to mind: aviation regulation. Commercial aircraft have many autonomous and semi-autonomous technologies that have public safety ramifications. Is that history instructive in thinking about the regulation of autonomous vehicles?

One could argue that autonomous cars present an "easier" and more responsible challenge than aviation, said Jeff Huber, Chief Executive Officer of Grail, a firm that seeks to use AI to detect early cancers in asymptomatic individuals through a blood screen. "In my view airline regulation was spectacularly irresponsible. It was effectively 'trialed-and-errored' with people in planes without any feedback loop other than whether the plane crashed or not." By contrast, the Google, Tesla and other autonomous cars have driven more than two million miles, and data from these ongoing tests in real-world circumstances are being fed back into the system in real time. The AI systems are learning, and the rate of learning is dramatically faster [than that which occurred in aviation].

However, "Autonomous driving on the roads is a much more complicated problem than commercial aviation," said Rao Kambhampati, President of the Association for the Advancement of Artificial Intelligence. He noted that both aviation and self-driving automobile designs need to do a better job of taking human factors into account — but autonomous cars, in particular, need designs that can address human interventions necessary in the zone between 100 percent autonomous and zero percent autonomous. "It's not enough for the car to simply buzz and say 'Pay Attention!'" he said, adding that more user education and training are needed.

Stuart Russell, professor of computer science at the University of California Berkeley, believes that regulation of commercial aviation has a very different character than what may be needed for autonomous cars. "It took forty years of step-by-step experimentation for the Federal Aviation Administration to approve things like automated landings," he said. It helped that Boeing essentially had a monopoly, so there were no competitive pressures to deploy new innovations before the other guy." At every single stage in the evolution of new technologies, said Russell, the FAA spent years to see if something worked before moving on to the next step. "Expecting that [regulation of autonomous cars] can jump all the way to Level 4, and that we can just stick cars out there and hope for the best — and appeal to the history of aviation as our precedent — is not reasonable," he said. "The history of aviation regulation shows that a great deal of care was taken at every step along the way."

However, the copious amounts of real-time data on actual performance of autonomous vehicles make a big difference in evaluating the technology, said Astro Teller of X. On the other hand, AI systems are likely to produce lots of unanticipated emergent behaviors, said James Manyika, Director of the McKinsey Global Institute, especially in moving from Level 1 to Level 4. The complexity of assessing subtle, interactive algorithms is likely to elude even many experts and possibly regulators, he said. Indeed, a big part of the research agenda for AI is to get better at understanding and modeling these emergent properties, along with verification and control.[27]

There were a number of suggestions for improving any regulatory process for self-driving cars and refining the technology itself.

David Kenny, General Manager of IBM Watson, recommended the use of AI-based audits as a far better oversight tool than human auditors. It would also be faster than many forms of conventional regulation. Kenny suggested rotating working AI experts in and out of regulatory agencies so that the agencies could make more informed, sophisticated decisions.

If tech improvement, transparency and social trust all matter, "Why not 'open source' the oversight?" *– Mustafa Suleyman*

Finally, Kenny suggested a global competition among smaller countries and city-states like Singapore as test beds for proving autonomous car technologies in systemic ways. Joi Ito of the MIT Media Lab cautioned that a global competition among allies may not be the best way to proceed. He did note that, in terms of thoughtful regulation, "Socialist countries have a much better alignment of incentives for taking a long-term view of what's going to prevent harm and have the best impact."

If tech improvement, transparency and social trust all matter, "Why not 'open source' the oversight?" asked Mustafa Suleyman, Co-founder of Deep Mind, an AI company based in London. "Why not be much more transparent about our models, our processes, our development frameworks and test frameworks? I think there are lots of really smart, technically savvy people who are willing to be part of a collective process of governance and oversight if we, as developers and companies, are prepared to provide a framework and be much more open and willing to engage."

Suleyman described an experimental model used in healthcare that empowers a panel of independent reviewers to act as non-contractual, unpaid, independent reviewers of his firm's work. "Their mandate is essentially to audit us in the public interest. There is a terms of reference and scope document, and the panel members meet periodically. They can interview people on my team. They will publish a report. They have a budget. I think it's a first step towards trying to build public trust by proactively providing access. It provides some reassur-

ance that we're behaving responsibly and that we're prepared to hold ourselves pro-actively responsible."

Suleyman acknowledged that there are some proprietary issues that would need to be addressed in such a scheme, but added, "You can solve for those sorts of things." He added that transparency and oversight also help improve the technology as outsiders identify bugs.

AI and Healthcare

A second session of the AI Roundtable focused on the role that artificial intelligence is playing in healthcare. In diverse contexts, emerging AI systems are transforming the character of medical research, patient diagnoses and treatment options. AI systems are also changing the economics of certain types of medical care and broadening access to specialized knowledge — shifts that could dramatically reconfigure medical treatment norms and healthcare markets. As tech firms such as IBM, Dell, Hewlett-Packard, Apple and Hitachi develop AI plans for healthcare, it is expected that AI's use in medicine will increase tenfold within the next five years.[28]

In trying to take stock of these changes, it helps in the first instance to distinguish the different healthcare spheres that AI is affecting. Perhaps the largest, most consequential realm of AI applications, at least in the near term, involves Big Data. AI systems can be tremendously effective in searching and analyzing large pools of patient data to identify unusual patterns of physiological factors and symptoms. This knowledge, in turn, can help improve diagnosis and accelerate and refine new treatments.

AI as a research tool is primarily relevant to large institutions that administer or finance healthcare, such as government, insurers, hospitals, medical researchers and the like. Another important tier of AI applications focuses on individual patients, often in their home environments. Here the goal is to develop more insightful personalized medical assessments, diagnoses and treatment plans. AI can also help in providing "smart patient monitoring and alerts," such as tracking prescription drug regimens and flagging symptoms that require intervention.

Yet another field for deploying AI tools is in augmenting the intelligence and skills of physicians in the course of their work. By having

quick, searchable access to vast quantities of information, AI can help physicians make more discerning choices while providing patients with greater knowledge and statistical predictions about treatment outcomes. AI can also help doctors make more precise, personalized prescriptions of medicine and, through robots, perform automated surgery. At some point, AI is likely to improve healthcare management by improving efficiencies, accuracy and cost-effectiveness of medical practice.[29]

Summarizing AI's potential contributions to medicine, one commentator writes: "AI can help diagnose illness, offer novel treatment options, eliminate human error, and take care of all the repetitive tasks that clog up the system. These time saving measures mean more efficiency and reduced costs."[30]

The following sections review these different uses of artificial intelligence in healthcare before turning to the structural challenges and policy complications that often stand in the way.

AI as a Tool for "Deep Learning" in Medical Research

Jeff Huber, CEO of Grail, made a presentation about the applications of AI and machine learning to improve medical diagnosis and treatment of cancer. This is a significant healthcare issue because about fourteen million new cases of cancer are diagnosed each year, and eight million people die of cancer each year, he said.

"The premise behind Grail is actually a very simple one," said Huber. "Cancer that is detected in its early stages today — Stage I or Stage II — can be cured in 80 to 90 percent of cases. Their lives can be saved. Cancer detected in late stages — Stage III or Stage IV — is the inverse, a negative outcome 80 to 90 percent of the time, where people die. So instead of detecting cancer late, when the outcomes are usually bad, we want to detect it early, when people can be cured." Huber believes that early-stage detection of cancer could increase positive outcomes, cures, to 95 or even 99 percent of cases.

The catch, of course, is how to successfully detect cancer in its earliest stages when it is often invisible to conventional medical tests. For Grail, the tool for improving early diagnoses is known as "ultra-deep genome sequencing," a system that uses immense amounts of data and AI to try to detect nucleic acids and fragmentary RNA and DNA circu-

lating in a person's blood. Those elements are shed by a cancer from its very earliest stages, and so identifying them through a blood test could help detect cancer and treat it far earlier than is now possible.

The Grail test has four functions: Detecting whether a person has cancer; identifying how aggressively it is growing; pinpointing its location in the body; and helping doctors select the most appropriate therapies. Since medical scientists know the molecular and mutational drivers of various cancers, the knowledge revealed by the test can inform which therapeutic options should be considered — chemotherapy, immunotherapies, surgery, etc.

Huber said that Grail's AI system amounts to a tool for looking for needles in a haystack: "We're finding the needles at almost the limits of physics — a handful of those molecules in a tube of blood." This sequencing tool goes "an order of magnitude broader and two or three orders of magnitude deeper than anyone else is doing," he said. At the moment, every test using the ultra-deep genome sequencing is generating about a terabyte of data (10^{12}, or one trillion, bytes). Grail combines this test data with data from clinical trials and phenotypic data related to a patient's other diseases, co-morbidities, drugs he or she is taking, family medical histories, etc.

Grail's AI system pores through all this data looking for patterns that may reveal something about the four goals of the test. The process requires the creation of powerful machine-learning algorithms designed to penetrate to deeper levels of biological knowledge about cancer.

It is here that the dynamic forces shaping AI as a field become more significant — not just in this instance for Grail, but for AI systems more generally. Huber cited an essay by Beau Cronin, an expert in computational neuroscience, who identifies four basic ingredients for making AI systems work effectively. They are "data, compute resources (i.e., hardware), algorithms (i.e., software), and the talent to put it all together."[31]

While most people today assume that data is the most important element in successful AI applications — why else are Google and Facebook so successful? — Cronin argues that different scenarios could lead to the other factors becoming more influential. New hardware architectures could accelerate the development of better learning algorithms, for example. Or the wealthiest tech companies could attract the most talented programmers. Or access to good data could improve (or diminish) if privacy policies, security concerns or public opinion change.

It is also quite possible that the character of AI systems could be significantly affected by network dynamics. A successful AI firm could attract the most users and evolve into the biggest network, propelling a self-reinforcing "winner-takes-most" dynamic. Cronin quotes tech analyst Kevin Kelly, who predicts: "Our AI future is likely to be ruled by an oligarchy of two or three large, general-purpose cloud-based commercial intelligences."[32]

For now, Grail is trying to assemble a new dataset that has not existed previously while assembling the compute resources, algorithms and talent that integrate computer science, life science and biology.

As the Grail project demonstrates, machine learning with sufficiently large aggregations of data can open up vast new fields for medical research. As one small example, Huber cited a recently published study that analyzed electronic medical records. It discovered that a subset of diabetes patients had far lower incidence of cancer — on the order of one-third than the general population. This was a counterintuitive finding, said Huber, because one would expect that patients with diabetes, an inflammatory disease, would have *higher* rates of cell mutations at the margin and thus higher cancer rates. After looking more closely, researchers discovered that these diabetes patients were taking Metformin, an inexpensive drug for managing glucose levels, and this was apparently helping to fight cancer (further studies are seeking to confirm this suspicion).

"That's a relatively trivial case of machine learning using a sufficiently large aggregation of data to make important findings," said Huber. The problem is that most data is "incredibly siloed," he said. "Electronic medical records are in tiny different pools all over; there aren't any good aggregations." There are also many privacy, security and business-model factors that are preventing the aggregation of medical data — a topic below.

AI as Augmented Intelligence for Conventional Medical Care

Beyond medical research, AI systems can have important applications in the everyday practice of medicine, especially in helping physicians gain access to a wider body of knowledge and make better judgments. Given the explosion of the medical literature, physicians understandably may not be aware of new or unusual findings or treatments

for a given medical condition. One conference participant said that his wife had to visit five doctors before getting the correct diagnosis for a health problem. Others noted that doctors may not provide a balanced perspective about the available treatment options. Rao Kambhampati of the Association for the Advancement of Artificial Intelligence, envisions a future in which patients will consult doctors, but also ask, "What does the AI system say?"

There is enough that we do not know about health and disease that it may be dangerous to conflate data-driven analysis with the mysteries of the soma. The human factor matters. The will to live may triumph over the statistical predictions.

Even before formal medical diagnoses, AI could be used to provide early informal assessments to patients. The largest diagnostic expert system in the world — heavily used and totally unregulated — is Google. Even many doctors turn to Google when certain configurations of symptoms puzzle them. Of course, Google can function in this capacity only because it is an unofficial, non-authoritative source of medical information, and therefore it cannot be held liable for the information it provides.

AI systems could provide highly refined and targeted assistance to doctors, if only as a second-opinion drawing upon a vast pool of digitized knowledge. It could also help provide some measure of authoritative confirmation for their diagnostic and treatment choices.

Once again, the issue of liability arises: What if a doctor relying on the AI system makes an incorrect or unwise judgment? While an AI system might predict that a patient has only a 1 percent chance of surviving a given disease, should the doctor and patient take that data-driven judgment as conclusive — "AI systems as death panels?" as one participant pondered. There is enough that we do not know about health and disease that it may be dangerous to conflate data-driven analysis with the mysteries of the soma. The human factor matters. The will to live may triumph over the statistical predictions.

A related problem is the general lack of numeracy. Doctors are sometimes poor communicators, especially about statistical probabilities, and patients themselves may not be equipped to make good judgments based on numbers. Indeed, doctors themselves, when faced with fatal diagnoses, disproportionately choose not to receive medical care in order to avoid dying in hospitals.[33]

AI as a Tool to Empower Individuals

AI systems offer a wealth of new ways that patients can take better care of their health directly. There are consumer-facing apps that can monitor vital signs (see the "quantified self" movement); make preliminary diagnoses of illness and disease; manage prescriptions for patients; and oversee their adherence to drug regimens. When combined with AI systems used by physicians, individual patients are beginning to have a dizzying array of choices.

Huber believes that AI could begin to consolidate more of the medical information inputs and synthesize them, relieving both patients and doctors of that impossible task. Right now, he said, patients get routed through a series of specialists, but in effect, you need to be your own general contractor because no single doctor can know everything. "So let AI be the general contractor," he urged. AI could potentially understand all of the symptoms and outcomes, beyond what a specialist can.

It is too early to make any general conclusions about better treatment outcomes and reduced medical costs from such an approach, he conceded. But there are many anecdotes that suggest that prevention, early diagnoses and treatment, could save considerable money. A $500 blood test for cancer and $20,000 for early surgical intervention, for example, could save $2.7 million in futile treatments of late-stage cancer, Huber said.

One subtle but potentially huge impact of consumer-oriented AI systems is the disintermediation of conventional medical institutions. Just as individuals have used open networks to seize greater autonomy and choice from large, centralized institutions — newspapers, broadcasters, record labels, government — so new consumer AI systems could change how medicine is practiced, and where it is practiced. Services like WebMD, IBM's Watson, Grail, 23andMe, and even Google search are already changing the economics of healthcare by making it more affordable to shift services to more accessible and even remote locations.

Local drug stores now offer flu shots, a variety of wellness services, and nurse practitioners who diagnose illnesses, injuries and skin conditions. Google search is the first medical advisor that many people turn to.

These trends suggest that new innovations in medical AI may first take root and flourish in unregulated corners of the world. It is in these spaces — websites, adjacent retail sectors, foreign nations — where people are likely to have greater personal agency in opting into new types of healthcare delivery that leverages AI in creative ways. Unfortunately, the new disintermediated opportunities for AI-assisted healthcare will also be prime targets for bad actors having dubious medical expertise. Some sort of rapprochement between social responsibility and medical innovation will have to be negotiated and refined.

Structural Barriers to Expanding AI in Healthcare

If the vision for AI-driven change in healthcare is often compelling, the forces of resistance are deeply entrenched. "There is no shortage of heartwarming stories about what we could do with big data, and what people are beginning to do," said Wendell A. Wallach, the Yale bioethicist. "But at some point in these discussions, we always come upon these macro, structural problems." There are many players in healthcare policy debates who have their own reasons for opposing the use of AI and Big Data in medical contexts.

Michael W. Ferro, Chairman and CEO of Merrick Ventures and tronc, Inc., added that many governments have their own reasons for not pursuing AI-based innovations in healthcare: "Very powerful people in governments are really worried about all these healthcare innovations because it creates a whole new issue for them. If everyone lives longer, they [politicians] don't know how to pay for it."

Another major impediment to many AI-based approaches to healthcare is privacy. If healthcare data becomes available to employers or insurers, it could lead to discrimination against people in hirings, firings and insurance applications. And yet there are potentially important public and individual benefits from using artificial intelligence to detect illnesses and disease. Jeff Huber said that it is technically feasible for AI agents on one's smartphone or computer to detect signals of potential mental disorders in users. "Early treatment would save lives and be a society good — but where does privacy start and stop in a situation like that?"

Stuart Russell, the computer scientist, said that he has an adjunct position in neurosurgery at University of California San Francisco, where "it took three years to get legal permissions to use our own data from our own ICU [Intensive Care Unit] for research purposes." A far bigger problem, Russell added, is that "medical equipment manufacturers won't allow researchers to access data being collected by their physiological measurement devices. They want to have a monopoly over the data." Russell said that his colleagues have struggled with this issue for twenty-five years, and a nationwide consortium of researchers on which he sat has tried and failed to find solutions for five years.

Marc Rotenberg of EPIC noted that there are techniques for de-identification and anonymization of data that could provide some solutions by allowing data to be used in research with minimal privacy risks. "Of course, privacy experts tend to be a bit skeptical of this scenario," he conceded, "and want to know how it is really going to work." Rotenberg nonetheless believes that privacy is not a "zero-sum problem" and that win-win solutions are possible.

Then, of course, there are liability concerns. "Who is going to be responsible for these decisions — the recommendations and action based on AI?" asked Jeff Huber. Google says that it is willing to underwrite liability for Level 4 cars, but so far no one in the healthcare industry is willing to take on liability for AI systems and data-driven decision-making. This may be a case in which the state is the only player with sufficient incentive and means to address the problem. Any private insurer or tech company is ill-equipped to handle the magnitude or complexity of liability.

Perhaps the trickiest issue is whether AI-enabled healthcare would reduce or raise costs. To date, said Jeff Huber, "There is no evidence that all the health improvements we've made with technology decade after decade have actually lowered healthcare costs or improved outcomes. We're dealing with a healthcare economy that is $3.05 trillion, the seventh largest economy by itself in the world — a half trillion dollars less than the powerhouse economy of Germany. And yet there is no end in sight for how we are going to control costs."

Several participants agreed that a paradigm shift in healthcare is needed. It would benefit consumers who want personalized medicine and better medical outcomes, and help physicians, who could improve

diagnoses and treatment; and researchers, who could have access to large bodies of aggregated data to improve their understanding of disease. But pursuing such a paradigm-shift, and developing a new systemic infrastructure to host AI systems and Big Data, remains elusive.

These differences are more influential in the deployment of AI data systems than a paternalism-versus-individual empowerment framing. *- Mustafa Suleyman*

A few participants argued that this discussion is not just about biology, data and medical knowledge, but equally about patient agency and social trust. Mustafa Suleyman of DeepMind said: "An individual must have the personal agency to develop trust. He or she must be able to say 'I approve this legitimate research use of my data,' or 'I withdraw consent from this particular use.' We should be creating a verifiable digital structure around that." Conversely, others argued that government agencies such as the Veterans Administration, which has one of the biggest repositories of healthcare data, could and should assert ownership of the data, and put it to use for public benefit.

While some argued that the debate is essentially a choice between government paternalism and individual empowerment, others replied that this is too simplistic. There are systemic differences among healthcare systems in the world, such as the single-payer system in the UK and the market-driven healthcare in the U.S. These differences are more influential in the deployment of AI data systems than a paternalism-versus-individual empowerment framing, said Suleyman. Jeff Huber agreed, saying that the U.S. healthcare system is simply unable to facilitate the kind of data-collection and analysis that Grail is currently undertaking. In many respects the U.S. system *favors* late-stage cancer treatments because it is more profitable than prevention. By contrast, the UK healthcare system is more structurally aligned with advancing long-term outcomes, he said, which is why Grail is doing its data trials in the UK.

AI Begins to Change Journalism and News Organizations

Just as systems of artificial intelligence are rapidly changing automobiles and healthcare, so it is transforming journalism and media, especially in online contexts. Alberto Ibargüen, President and Chief Executive Officer of the Knight Foundation, declared, "We're at the very beginning of this process, which is fundamentally the same as what Europe experienced just after the arrival of the mechanized press. Before Gutenberg, monks would illuminate a few manuscripts a year and Church officials would give its imprimatur so all would know what was 'truth.' After Gutenberg and for 100 years, no one could figure out what was true or whom to trust because anyone could publish a book or pamphlet. I find that analogy accurate, comforting and hopeful."

". . . . it is a democratic imperative to figure out how serious journalism can be economically viable in today's digital environment."
– Michael Ferro

It helps to understand the basic forces that are roiling the news business and media today. As Charlie Firestone of the Aspen Institute explained, "The news business, like many others, has been disintermediated and decimated by the digital revolution." Technologies have provided countless new choices for people, who often had limited access to information. But there is also a huge fragmentation of information sources now competing with conventional journalism for people's attention. The non-journalistic competitors include Facebook, LinkedIn, Snapchat and Twitter as well as countless specialized websites and blogs that arguably have closer, more credible connections with a field than do general-audience newspapers and broadcasts.

"Just as the unit of commerce in music went from the album or CD to the individual song," said Firestone, "so the unit of commerce in the news has gone from the publication or broadcast outlet, to the story." That has been very useful for search engines such as Google, which have been disaggregating the functions of journalism, he said. Today, different

digital entities are taking these different parts of newspapers and recombining them. Curated websites are ascendant, but much of the information on them is based on "free" content generated by someone else.

These dynamics are undermining the formerly dominant business models of media organizations, privileging commercial viability at more granular levels (individual stories, user clicks) and eviscerating journalism as a profession. Web-based sources of news and information are eclipsing traditional journalism even as they feed on it. One conference participant suggested that Facebook arguably has more impact on public conversation these days than the *New York Times* in the sense that its AI-driven newsfeeds reach hundreds of millions of people and are widely shared. Local news is particularly vulnerable, noted many participants, because it is difficult to monetize super-local content, even with the help of AI.

The news media are indispensable sources of civic information for citizens and a force for oversight and accountability of those in power.

Contemporary journalism faces a financial crunch: News organizations are not getting paid by the ad-based, entertainment-oriented content curators who depend upon journalists' original news stories, which require costly reporting, fact-checking and investigations. "Suddenly, it is very hard to get people to pay for their broccoli," said Michael Ferro, referring to the substantive, well-reported journalism that any democratic society needs. "People don't even want to pay for their *corn*! Let's just say that people are getting their corn for free."

Ferro believes that it is a democratic imperative to figure out how serious journalism can be economically viable in today's digital environment. It may even be necessary for government or ad-based models of free content to subsidize journalism, he said. This, after all, is an approach that some European nations have taken as a means to support newspapers and book publishing, and thus a foster better-informed public. It is also the rationale for the U.S. Government's long-standing

postal subsidies for newspapers and printed matter. However, some participants strongly disagreed with subsidies for newspapers, which they felt would only result only in bad newspapers.

In any case, Ferro, as Chairman of tronc., the company that owns the *Los Angeles Times* and *Chicago Tribune*, is determined to find new business models to support serious journalism as a vital institution in American democracy. The news media are indispensable sources of civic information for citizens and a force for oversight and accountability of those in power. For now, said Ferro, journalism is surviving today chiefly through revenue from its print operations. While online news startups are getting a lot of attention, he said, "No one can make money once they have to scale and build an infrastructure."

Artificial Intelligence Enters the House of Journalism

Artificial intelligence appears to have both positive and negative impacts on journalism. It offers tools that enable journalists to perform their jobs more efficiently and generate new insights from data search-and-analysis. But AI also is a powerful tool for content personalization, which tends to disaggregate existing news products (newspapers, broadcasts) and undercut traditional business models for journalism. AI is also a tool for mischief and misinformation spread by automated bots.

To be sure, the personalization of news is in many respects a gain for readers. "AI seeks to learn what its users want and how they want it," writes Francesco Marconi of the Associated Press. "In the specific case of news media, articles can be processed through algorithms that analyze readers' locations, social media posts and other publicly available data. They can then be served content tailored to their personality, mood and social economic status, among other things."[34] This capacity has enabled the Weather Channel to customize some of its content and so improve advertising CPMs [the price charged for 1,000 user impressions of one webpage], said David Kenny of IBM, who was previously Chairman and CEO of The Weather Company, owner of the Weather Channel.

While filtering of news may make it more relevant to individual readers (generating more clicks and profits in the process), it can degrade the quality of journalism indirectly. Filtering tends to exclude diverse points of view and marginalize serious journalism and complex analysis. The headline from the mock-news website *The Onion* puts it

nicely: "Horrible Facebook Algorithm Accident Results in Exposure to New Ideas."[35] In a click-driven environment, it has become harder for reputable news organizations to commercially justify the "broccoli" that they have traditionally folded into their content mix.

"Bots are basically being used to 'weaponize' AI."
- Mark Riedl

There is another downside that AI can inflict on journalism — bots. Bots on open networks are often used to dilute the agenda-setting powers of traditional news media by building echo chambers of their own pseudo-news, misinformation and skewed perspectives. Narrowly focused political or commercial actors with no commitment to journalism or public service frequently use bots to spread propaganda or marketing disguised as news, leaving the public confused about what information is accurate, trustworthy and properly contextualized. "Bots are basically being used to 'weaponize' AI," said Mark Riedl, Associate Professor at Georgia Tech. "They just repeat the same misinformation over and over and over again. The human bias is to believe the things that they hear, more often than not."

Lili Cheng, General Manager at Microsoft, described how Microsoft released a "chat-bot" called "Tay" on Twitter in May 2016 after successfully testing it in smaller social networks in China and Japan. The bot was designed to simulate the conversational personality of a teenage girl. To the surprise of Microsoft designers, Tay in a U.S. context attracted a wide variety of hateful social media users who posted vile racist and anti-Semitic comments, which in turn triggered Tay to automatically repeat such phrases and scrape material from hate websites. Microsoft quickly suspended use of the bot.

Joi Ito believes that "the architecture of the Internet may contribute to the cesspool of trolls online. Their anti-social behaviors may be an emergent property of the way that comments sections are organized." Ito speculated that perhaps an architectural change could dampen the emergence of uninformed mobs and amplify more constructive participation.

"If we had an AI system that could go through and remove even 10 or 20 percent of the most egregious hate speech, it might have a pervasive impact on how people put their thoughts into the world," said Astro Teller. Mustafa Suleyman of DeepMind reported that his firm is actually working on a "respect engine," an AI system for that very purpose. Another participant wondered if AI could be used to identify and elevate great comments and multiple perspectives.

Journalism ultimately remains a creative human activity requiring judgment and originality.

All this speculation about possible uses of AI systems prompted Suleyman to emphasize an important point: "We should not talk about AI as if it had its own autonomy and agency independent of us. *We* are the ones who decide when to deploy AI systems, and for how long and in what context. We have to stop anthropomorphizing these systems."

Suleyman's warning highlights a key concern: *Who* will control the automated curation and exclusion of certain information via AI systems? Perhaps they could be used to fight trolls, but what happens if government wanted to use the same tools to censor or marginalize ideas that it dislikes? The U.S. Government has already approached Silicon Valley companies for their help in fighting ISIS propaganda websites. Is it possible that governments might use AI systems to try to manipulate public opinion?

These are some of the alarming possible uses of AI in news media. But there are also many benign, information-synthesizing tools that essentially convert raw data into natural language. In an article on how AI startups are reinventing media, Caleb Garling writes: "Companies like Automated Insights and Narrative Science are powering production of millions of auto-generated 'articles,' such as personalized recaps for fantasy sports fans. A similar metrics-based formula can be used to recap a customer's stock portfolio performance." A company called Arria is using AI to analyze complex data sets in numerous fields, such as finance and meteorology, and then produce expert reports — a process that once required human analysts. The Associated Press announced in 2015 that it would use AI software to write company earnings reports.

AI-driven analytics are also being used to spot news trends that human editors might not recognize. For example, an app called Banjo can pore through digital feeds from Twitter, Facebook and elsewhere on the Web to identify "important" (popular) news stories faster than a human editor might.[36]

Several participants fantasized about having an AI fact-checker, but the AI technologists in the room cautioned that that is an "AI hard problem" not easily solved. The more significant barrier to such fact-checking, warned David Kenny, is not technical, but psychological: "People don't want to hear that what they like to read is wrong."

The question posed at the outset of the conference — Should AI aspire to replace human beings or augment them? — remains a central issue in journalism and news media. AI may be able to take on many tasks historically performed by reporters and editors, participants agreed, yet there is consensus that journalism ultimately remains a creative human activity requiring judgment and originality.

AI Bound or Unbound?

In the conference's brief survey of how artificial intelligence is affecting automobile transport, healthcare and journalism, some general themes emerged about how society should think about AI and its great potential and dangers.

Reid Hoffman, Co-founder and Executive Chairman of LinkedIn, opened the concluding session by citing a classic essay, "Artificial Intelligence Meets Natural Stupidity," by Drew McDermott. McDermott laments: "Unfortunately, the necessity for speculation [within AI] has combined with the culture of the hacker in computer science to cripple our self-discipline." This is perhaps an unavoidable problem within a young discipline such as AI, McDermott argues, because "artificial intelligence has always been on the border of respectability, and therefore on the border of crack-pottery."[37]

Hoffman said that we can approach AI through the lens of utopia or dystopia. On the one hand, AI can help us better solve diseases, improve longevity, and help us address climate change, for example, or it can usher in a dystopian future that terminates life and healthy possibilities. AI can point us to utopian work scenarios as embodied in,

say, *Star Trek*, where people can overcome basic needs and pursue their passions, or toward a neo-feudalism that monopolizes AI to manage a large class of serfs.

"Within the next ten years, we are going to see inflection points that could take us in either utopian or dystopian directions." ***– Reid Hoffman***

Within the next ten years, we are going to see inflection points that could take us in either utopian or dystopian directions, said Hoffman. Because of the high uncertainties about the implications of AI, there is a tendency to move toward greater certainties, if only to block any dystopian possibilities. But Hoffman believes there is a sounder way to move forward, to "get to the good solutions faster as a way to avoid dystopia."

"This leads to great questions around what are good outcomes and who gets to make the decisions balancing risks and hoped-for outcomes," said Hoffman. While he agrees that there ought to be forums to explore these questions, he is cautious about government regulation "because that increases the possibility of dystopian outcomes." The fundamental design challenge, he said, is "figuring out how we actually achieve the good outcomes…..How can we have utopian solutions nudge aside the dystopian ones? How to make the benefits of AI more inclusive and minimize the disruptive impacts on jobs? How to deal with the cyber-security issues? Can we have a broader range of training to people, and help teach empathy in better ways?"

The answers to these questions will hinge upon who controls AI, he said, and this could be a problem because the control is not likely to be perfectly democratic. Nonetheless, Hoffman said, "I think AI can be very positive for humanity." Indeed, he thinks it is not just about developing new tools and services, but also about "how we evolve positively as a species."

Astro Teller agreed with Hoffman's "getting there faster" scenario because of the dangers of slowing down development of AI. The bad actors who exploit AI technologies for anti-social purposes are not going to slow down, he noted. "Surely we would prefer for the best

and most responsible people with the most thoughtful versions of the future, to be the ones that get there first."

But Stuart Russell, the UC Berkeley computer scientist, cited a famous 1960 article by mathematician and philosopher Norbert Wiener that pointed out that the problem is not just "bad guys doing bad AI, but good guys accidentally doing bad" — a scenario exemplified by the "sorcerer's apprentice" tale and King Midas. Wiener writes: "If we use, to achieve our purposes, a mechanical agency with whose operation we cannot interfere effectively…we had better be quite sure that the purpose put into the machine is the purpose which we really desire."

The metaphor of "good AI crowding out bad AI" is misplaced, said Russell, because the field should not define itself as maximizing one set of objectives and then castigating any constraints on those objectives. He noted that much of the work of nuclear fusion researchers is in fact focused on containment; the field incorporated within its own research norms a deep ethical concern for preventing harmful social outcomes.

Marc Rotenberg of the Electronic Privacy Information Center proposed a third vision for AI technology that is neither utopian nor dystopian, but rather a "dystopia that appears to be a utopia." He cited the 1997 science-fiction film *Gattaca*, in which people's lives are predetermined by their genetic code — a utopian enactment of technological perfection. "What do you do when technology provides so much opportunity that it raises very deep questions about our roles as individuals and humans in a very technologically advanced world?" Or as the tagline for the movie puts it, "There is no gene for the human spirit."

What Vision for Responsible Control and Social Trust?

Expanding upon the earlier discussion about regulation of AI-enabled cars, the remainder of the conference focused on key social and governance issues that AI technologies raise. Since the world is still at a rather early stage in the development of commercial AI systems, this topic remains something of a frontier issue. However, there was broad consensus that society must consider what appropriate control structures should manage the development of AI systems. If there is going to be social trust and acceptance of AI, there must be systems for open debate and effective control and accountability. Participants were also

concerned about ensuring democratic access and use of AI technologies, and fair distribution of their benefits.

"Who controls AI" is a central question because it will largely determine the types of public oversight that are possible, as well as the character of the labor market. The question is how such oversight and accountability might be established.

Astro Teller suggested that the constant churn of competition that has characterized most major industries — autos, computer chips, personal computers, e-commerce — will provide an important check on AI firms. "There has been plenty of competition in these industries, and who stays on top from one decade to the next is not clear," said Teller. He noted that it was also likely that a relatively small group of companies would become the technological leaders because of their more intensive investments and expertise.

Antonio Gracias, CEO of Valor Equity Partners, suggests that competitive markets may be the wrong framework for thinking about AI accountability. A more apt analogy might be to the nuclear power and weapons industries than to consumer-facing industries, he said. Gracias thinks we are living in a time similar to when the atom was split. We are starting to realize that the technology has enormous military and geopolitical implications. "The real issue here is state and non-state actors," said Gracias, because AI could enable interference with state or non-state actors in ways that are "basically undetectable." This is why "we should worry about power structures that control AI," he said.

Joi Ito of MIT Media Lab said that he "slightly disagrees" in the sense that "dozens of breakthroughs in AI could happen," especially ones in which "computational capabilities could be made more accessible to a wider range of people." As an historical comparison, Ito cited the release of Lotus 1-2-3, the spreadsheet software, which enabled small businesses and ordinary people to do accounting services that were once only available from large accounting firms. "What if some sophisticated user interface were to be developed that could democratize access to AI?" asked Ito. At the same time, Ito conceded that such interfaces may not materialize (look at Linux) and agreed that "we should worry about what happens to the power structure that is built around this."

Marc Rotenberg is convinced that "algorithmic transparency" is needed to ensure the accountability of AI systems. This is important, in part, to ensure that we can determine who is legally responsible for an AI system's performance. He invoked a recent ruling by the Wisconsin State Supreme Court involving a proprietary algorithm that had been used in criminal sentencing proceedings to predict the likely recidivism of an individual. The court ruled that while the algorithm could be considered in making a recommended sentence, "there has to be a human agency in the loop," as Rotenberg paraphrased the ruling.[38]

AI and Livelihoods: An Inescapable Challenge

One of the biggest issues surrounding AI technologies is how they will affect people's livelihoods and jobs. Will the various innovations enabled by AI be widely and equitably shared? That is likely to affect public acceptance of AI and, perhaps indirectly, how the technologies will be allowed to develop. "It's not clear that the benefits of AI technologies, left to their own devices, will be evenly distributed," said James Manyika, Director of the McKinsey Global Institute. There are estimates that between 5 and 9 percent of full-time jobs may be automated out of existence over the next ten years, he said. "These rates of automation will be in the 15-20 percent range for middle skill jobs," he added. "We also find that 30 percent of activities in 60 percent of jobs will be automated, which means many more jobs will be changed rather than automated."[39]

Even partial automation — i.e., technology augmenting human skills — tends to have negative impacts, said Manyika. First, the effects of new technologies on employment tend to have two tracks — one in which well-educated, privileged people are enabled to do amazing new types of work while the jobs of many other workers are deskilled, leaving them with fewer responsibilities and lower wages. "Easier tasks can be paid less and probably need less certification, so employers get a bigger supply pool for such jobs," said Manyika.[40] Jobs are often structured so that "smart systems" can operate in the background, reducing the skills needed by on-the-ground technicians. So even when you have partial automation or augmentation of human work, "it often has a depressive impact on wages," he explained.

"So we come back again to the wage and income question," said Manyika. "That's a much more complicated conversation that we are going to have to have at some point." In the meantime, he said, these concerns are fueling a lot of social unrest and political populism. This is primarily because close to two-thirds of households in the advanced economies have seen stagnant or falling incomes over the last decade or so. He noted, "While, the recession has a lot do with that, technology along with other factors are also to blame."[41]

"It's not clear that the benefits of AI technologies, left to their own devices, will be evenly distributed." - ***James Manyika***

The resentment is particularly emphatic, said Mustafa Suleyman, because people realize "who gets to direct which tech applications are built, how they will be deployed, and who's on the receiving end of the decisions being made, perhaps by three or four people in this room and their teams. There's not very much oversight, accountability or transparency."

This argues for paying close attention to this problem now, argued Wendell Wallach, the Yale bioethicist, because "there is a principle called the Collingridge dilemma, which states that 'by the time undesirable consequences [of technology] are discovered…the technology is often so much a part of the whole economic and social fabric that its control is extremely difficult.'[42] The Collingridge dilemma has stymied technology policy for decades," said Wallach. "Those like me, who advocate for more anticipatory governance, reject the dilemma's simple binary logic. I argue that there is an 'inflection point,' a window of opportunity, in which we can act once the problem comes into view but before the technology is fully entrenched. That window can be short or long. This is all fully discussed in my book, *A Dangerous Master.*"[43]

Wallach added that modulating the pace of developing a technology — for example, by speeding up investment or slowing down with regulation — is a separate matter. While Wallach advocates modulating the pace of technology development, it is not necessarily an outgrowth of

the Collingridge dilemma. However, he said, "When we do slow down the rate of development we can stretch out the inflection point and perhaps have a better opportunity to act."

"Very few of us would say we should stop greater efficiencies just because they are taking jobs away," he said. "So it becomes a political problem: How do you distribute goods and resources if jobs are no longer providing enough income to people?" There is great resistance to guaranteeing basic incomes or other forms of meeting people's needs, he said, although this topic has gained greater currency over the past year.

Concerns about a structural loss of good-paying jobs are usually rebuffed by reassurances that new technologies over time will create enough new jobs to offset short-term losses. That, after all, has been the long historical record. But based on the disruptive impact of the Internet, which has indeed created many new jobs, the new jobs being created "aren't paying people enough income to support themselves."

This topic was extensively discussed at a 2013 Aspen Institute conference on this topic, "Power-Curve Society: The Future of Innovation, Opportunity and Society Equity in the Emerging Networked Economy," said Charlie Firestone of the Aspen Institute Communications and Society Program.[44] The report probes how inequality seems to be structurally related to today's networked economy:

> Wealth and income distribution no longer resemble a familiar "bell curve" in which the bulk of the wealth accrue to a large middle class. Instead, the networked economy seems to be producing a "power-curve" distribution, sometimes known as a "winner-take-all" economy. A relative few players tend to excel and reap disproportionate benefits while the great mass of the population scrambles for lower-paid, lower-skilled jobs, if they can be found at all. Economic and social insecurity is widespread.

Reid Hoffman of LinkedIn said, "I'm generally of the belief that the problem will sort itself out in the long-term. But the problem is that the long term is long term — and it allows for a lot of pain and suffering in the meantime, and potentially very volatile circumstances." He added that "if you think we're experiencing exponential problems, then we need to think in terms of exponential solutions."

"When we do slow down the rate of development we can stretch out the inflection point and perhaps have a better opportunity to act."
– Wendall Wallach

What's missing from discussions of this issue, said Mustafa Suleyman of DeepMind, is a vision. It's hard to devise any transition plans for dealing with transitional disruptions if we do not have a vision for how things could work instead. For Suleyman, this means going beyond debates about regulation to broader questions of governance. He also urged that AI firms "create porous and stable trust-based boundaries within our own organizations through the use of independent oversight boards," as previously mentioned. "When we were acquired, we made it a condition of our acquisition that we set up an ethics and safety board with independent representation to steward our technology in the public interest," he said. "It is just an early experiment in governance, but it demonstrates our intent and takes the first steps."

Cynthia Breazeal of MIT Media Lab added that *who* innovates matters, too. Despite its efforts to broaden the race, gender, ethnicity and socio-economic backgrounds of its leaders and employees, the tech industry still faces major challenges in this respect.

Conclusion

It is clear that AI technologies are having deeper and more pervasive impacts than ever before, a trend that will only accelerate in the coming years. Many of those impacts will surely be highly beneficial — and many will be socially and economically harmful. In either case, it seems clear that we stand at an inflection point in the development of artificial intelligence that requires us to imagine appropriate frameworks for the deployment of commercial AI systems. How should markets evolve? What are the proper forms of government regulation? What internal governance structures and safeguards should AI adopt? How can social trust in AI technologies be earned?

This conference could not resolve these complicated issues, of course, but it did bring together some of the leading minds in the AI world to collectively take stock of the on-the-ground realities of AI technologies in three important sectors. Leaders from different segments of AI research, development and commercialization were able to identify key questions to ask, debate the social, economic and legal implications, and propose notional frameworks for moving forward.

The issues raised will certainly need further discussion and inquiry because, as one participant noted, we are not facing a single inflection point but an ongoing *series* of inflection points. In the years ahead, there will be technology breakthroughs, policy proposals, court rulings, public opinion and much else that will guide and direct the evolution of AI technologies. This conference was one attempt to bring the relevant issues into focus.

Endnotes

1. Walter Isaacson. *Steve Jobs*. Simon & Schuster (2011); *The Innovators:* "How a group of inventors, hackers, geniuses and geeks created the digital revolution." Simon & Schuster (2014).

2. Isaacson, *The Innovators*, p. 29.

3. Enthusiasm for AI emerged in the early 1980s along with concerns about Japanese "Fifth Generation computing" (McCorduck 1983), but support diminished later in the decade (Unger 1987).

4. Wikipedia. "AI winter." Wikipedia.org. (n.d.). Available at https://en.wikipedia.org/wiki/AI_winter. There were two "winters" – one in the late '60s/early '70s after the work of ALPAC (Automatic Language Processing Advisory Committee, 1966), Marvin Minsky and Seymour Papert (69), and then another one in the late '80s after the "failure" of rule-based expert systems, around 1987. The second one was called an "AI Winter," and had been predicted by Levesque and named after nuclear winter, a phrase coined in 1983 by Turco.

5. See, e.g., Larry Summers. "Robots Are Hurting Middle Class Workers and Education Won't Solve the Problem." *The Washington Post.* (March 3, 2015); or his review of Robert Gordon's book, *The Rise and Fall of American Growth*, in *The American Prospect.* (February 2, 2016). Available at http://larrysummers.com/2016/02/02/will-our-children-really-not-know-economic-growth. Summer writes: "…it is hard to shake the sense that something unusual and important and job threatening is going on. The share of men between the age of 25 and 54 who are out of work has risen from about 5 per cent in the 1960s to about 15 per cent today as it appears that many who lose jobs in recessions never come back. Perhaps, as [Robert] Gordon seems to suggest, this is a sociological phenomenon or a reflection of increasing problems in education. But I suspect technology along with trade have played an important role."

6. Erik Brynjolfsson. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies.* WW Norton & Company. (2014).

7. Texas A&M Transportation Institute. "Traffic Gridlock Sets New Records for Travel Misery." 2015 Urban Mobility Scorecard. (August 26, 2015). Available at https://mobility.tamu.edu/ums/media-information/press-release/.

8. Daniel Fagnant, Kara Kockelman. "Preparing a Nation for Autonomous Vehicles; Opportunities, Barriers and Policy Recommendations." Eno Center for Transportation, University of Texas. (October 2013). Available at https://www.enotrans.org/wp-content/uploads/2015/09/AV-paper.pdf.

9. Brad Plumer. "15 ways that self-driving cars could transform our lives." *Vox.* (May 28, 2014) Available at http://www.vox.com/2014/5/28/5756736/the-case-for-self-driving-cars.

10. Society of Automotive Engineers International, Standard J3016. "Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems." (2014). Available at http://www.sae.org/misc/pdfs/automated_driving.pdf.

11. Josh Lowensohn. "Elon Musk: Cars you can drive will eventually be outlawed." *The Verge.* (March 17, 2015) Available at http://www.theverge.com/transportation/2015/3/17/8232187/elon-musk-human-drivers-are-dangerous.

12. David Roberts. "The biggest challenge for driverless cars: getting rid of the drivers." *Vox.* (October 7, 2015). Available at http://www.vox.com/2015/10/7/9471393/driverless-cars-challenges.

13. Maria Konnikova. "The Hazards of Going on Autopilot." *The New Yorker.* (September 4, 2014). Available at http://www.newyorker.com/science/maria-konnikova/hazards-automation.

14. Norbert Wiener. "Some Moral and Technical Consequences of Automation." *Science.* Vol. 131, No. 3410 (May 6, 1960). pp. 1355-1358. Available at http://www.nyu.edu/projects/nissenbaum/papers/Wiener.pdf.

15. Will Knight. "How to Help Self-Driving Cars Make Ethical Decisions." *MIT Technology Review.* (July 29, 2015). Available at https://www.technologyreview.com/s/539731/how-to-help-self-driving-cars-make-ethical-decisions.

16. Wendall Wallach. "Imperfect Solutions to Driverless Car Dilemmas." The Hastings Center. (September 14, 2016). Available at http://www.thehastingscenter.org/imperfect-solutions-driverless-car-dilemmas/.

17. Nick Jaynes. "Chinese hackers take control over moving Tesla from 12 miles away." *Mashable.* (September 20, 2016). Available at http://mashable.com/2016/09/20/chinese-hackers-tesla/#BvS6PFpUHEq8.

18. The Drivers Privacy Act of 2015 establishes baseline privacy standards for Event Data Recorders. *See* Rotenberg, *Privacy Law Sourcebook* (EPIC 2016).

19. Wikipedia. "Three Laws of Robotics." Wikipedia.org. (n.d.). Available at https://en.wikipedia.org/wiki/Three_Laws_of_Robotics. The Three Laws are: "1. A robot may not injure a human being or, through inaction, allow a human being to come to harm. 2. A robot must obey the order given it by human beings except where such orders would conflict with the First Law. 3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws."

20. Mark Rotenberg. "EPIC Promotes 'Algorithmic Transparency' at Annual Meeting of Privacy Commissioners." 38th International Conference of the Data Protection and Privacy Commissioners, October 17-20, 2016. Marrakech. Available at https://epic.org/2016/10/epic-promotes-algorithmic-tran.html.

21. Cecilia Kang. "No Driver? Bring It On. How Pittsburgh Became Uber's Testing Ground." *New York Times.* (September 10, 2016). Available at http://www.nytimes.com/2016/09/11/technology/no-driver-bring-it-on-how-pittsburgh-became-ubers-testing-ground.html.

22. Scott Santens. "Self-Driving Trucks are Going to Hit Us Like a Human-Driven Truck." *Medium.* (May 14, 2015). Available at https://medium.com/basic-income/self-driving-trucks-are-going-to-hit-us-like-a-human-driven-truck-b8507d9c5961#.chwggu3up.

23. Farhad Manjoo. "For the Debaters: What Shall We Do About the Tech Careening Our Way?" *New York Times.* (September 22, 2016). Available at http://www.nytimes.com/2016/09/22/technology/for-the-debaters-what-shall-we-do-about-the-tech-careening-our-way.html?_r=0.

24. NHTSA. "U.S. DOT issues Federal Policy for safe testing and deployment of automated vehicles." NHTSA. (September 20, 2016). Available at http://www.nhtsa.gov/About+NHTSA/Press+Releases/dot-federal-policy-for-automated-vehicles-09202016.

25. Associated Press. "Feds Preview Rules of the Road for Self-Driving Cars." *New York Times.* (September 19, 2016). Available at http://www.nytimes.com/aponline/2016/09/19/us/politics/ap-us-self-driving-cars.html. One auto manufacturer, Toyota, was also unhappy with the voluntary guidelines, saying that they lack the clarity that the industry will need. Reuters, "Toyota Tech Director Calls California's Proposed Self-Driving Car Rules Preposterous" October 11, 2016. Available at http://fortune.com/2016/10/11/toyota-california-self-driving.

26. In late September 2016, the California Department of Motor Vehicles deferred to the federal guidelines by revising its draft regulations to allow Level 4 cars on the road without a licensed driver in them. Associated Press. "Amended California Rules Would Allow Driverless Cars." *New York Times.* (October 1, 2016). Available at http://www.nytimes.com/2016/10/02/us/driverless-cars-california.html?_r=0.

27. Russell, Stuart, Daniel Dewey, and Max Tegmark. "Research priorities for robust and beneficial artificial intelligence." (2016).

28. Philip Perry. "How Artificial Intelligence Will Revolutionize Healthcare." *Big Think.* (May 10, 2016). Available at http://bigthink.com/Philip-perry/how-artificial-intelligence-will-revolutionize-healthcare.

29. Sophie Curtis. "Deep Learning in Healthcare, Part I: Opportunities and Risks." Re*Work. (March 30, 2016) Available at https://re-work.co/blog/deep-learning-healthcare-experts-part-one-discuss-risks-and-opportunities.

30. See note 20.

31. Beau Cronin. "The Dynamic Forces Shaping AI." *O'Reilly.* (July 6, 2016). Available at https://www.oreilly.com/ideas/the-four-dynamic-forces-shaping-ai.

32. Ibid.

33. Ken Murray. "How Doctors Die." *Zocalo.* (November 30, 2011). Available at http://www.zocalopublicsquare.org/2011/11/30/how-doctors-die/ideas/nexus.

34. Francesco Marconi. "What artificial intelligence can do for journalists and newsrooms." Digital Content Next. (May 6, 2016). Available at https://digitalcontentnext.org/blog/2016/05/06/what-artificial-intelligence-can-do-for-journalists-and-newsrooms.

35. The Onion. "Horrible Facebook Algorithm Accident Results in Exposure to New Ideas." TheOnion.com. (September 6, 2016). Available at http://www.theonion.com/article/horrible-facebook-algorithm-accident-results-expos-53841.

36. However, critics have noted Facebook "trending" errors, an example of the invisible editorial bias that can be embedded in AI algorithms: https://techcrunch.com/2016/08/29/facebooks-trending-topics-algorithm-mistake and https://techcrunch.com/2016/08/26/confusing-topics/.

37. Drew McDermott. "Artificial Intelligence Meets Natural Stupidity" in John Haugeland, editor, Mind Design: Philosophy, Psychology, Artificial Intelligence (MIT Press, 1981), pp. 143-160.

38. EPIC.org. "Wisconsin Supreme Court Upholds Use of Sentencing Algorithms, But Recognizes Risks." EPIC. (July 16, 2016). Available at https://epic.org/2016/07/-wisconsin-supreme-court-uphol.html.

39. Michael Chui, James Manyika, and Mehdi Miremadi, Chui. "Where Machines Could Replace Humans – and Where They Can't (Yet)." *McKinsey Quarterly*. (July 2016).

40. Michael Chui, James Manyika, and Mehdi Miremadi, Chui. "Four fundamentals of workplace automation." *McKinsey Quarterly*. (November 2015).

41. McKinsey & Company. "Poorer Than Their Parents? Flat or Falling Incomes in Advanced Economies." *McKinsey Global Institute*. (July 2016).

42. David Collingridge, *Social Control of Technology*. (1980).

43. Wendell Wallach. *"A Dangerous Master: How to Keep Technology from Slipping Beyond Our Control."* Basic Books. (2015), pp. 71-72.

44. See http://csreports.aspeninstitute.org/documents/Power-Curve-Society.pdf.

APPENDIX

Artificial Intelligence Comes of Age

Aspen, Colorado
August 1-2, 2016

Roundtable Participants

David Bollier
Author, Editor, Blogger
Commons Strategies Group

Cynthia Breazeal
Associate Professor
MIT Media Lab

Lili Cheng
General Manager
Microsoft

Ken Denman
Former Chief Executive Officer
Emotient

Michael Ferro
Chairman and Chief Executive Officer
Merrick Ventures
Chairman
tronc, inc.

Charles M. Firestone
Executive Director
Communications and Society Program
The Aspen Institute

Stuart Frankel
Chief Executive Officer
Narrative Science

Antonio Gracias
Chief Executive Officer
Valor Equity Partners

Reid Hoffman
Co-Founder and Executive Chairman
LinkedIn

Jeff Huber
Chief Executive Officer
GRAIL

Alberto Ibargüen
President and Chief Executive Officer
The Knight Foundation

Walter Isaacson
President and Chief Executive Officer
The Aspen Institute

Note: Titles and affiliations are as of the date of the conference.

Joi Ito
Director
MIT Media Lab

Rao Kambhampati
President
Association for the Advancement of Artificial Intelligence
Professor, Computer Science
Arizona State University

David Kenny
General Manager, IBM Watson
IBM

James Manyika
Director
McKinsey Global Institute
Senior Partner,
McKinsey & Company

Mark Riedl
Associate Professor
Georgia Tech

Marc Rotenberg
President and Executive Director
Electronic Privacy Information Center

Stuart Russell
Professor, Computer Science
University of California Berkeley

Fr. E. Salobir
President of OPTIC
Order of Preachers

Mustafa Suleyman
Co-Founder
DeepMind

Astro Teller
Captain of Moonshots
X

Wendell A. Wallach
Author, Scholar, Consultant
Interdisciplinary Center for Bioethics
Yale University

Staff:

Robin Farmanfarmaian
Vice President,
INVICTA Medical
Co-Founder,
Organ Preservation Alliance
Consultant, the Aspen Institute

Jennarose Placitella
Project Manager
Communications & Society Program
The Aspen Institute

Note: Titles and affiliations are as of the date of the conference.

About the Author

David Bollier is an author, activist, independent scholar and blogger well-known for his work on the commons as a new paradigm of economics, politics and culture. He pursues this scholarship and activism as co-founder of the Commons Strategies Group, an advocacy/consulting project that assists the international commons movement. Bollier has written or edited seven books on the commons, including *Patterns of Commoning* (2015), co-edited with Silke Helfrich; *Think Like a Commoner: A Short Introduction to the Life of the Commons* (2014); *Green Governance: Ecological Survival, Human Rights and the Commons* (2013), co-authored with Burns Weston; and an anthology of essays, *The Wealth of the Commons: A World Beyond Market and State* (2012), co-edited with Silke Helfrich.

Bollier spent many years in various policy jobs in Washington, D.C.–in Congress, the auto safety agency, with Ralph Nader and others–in the 1970s and 1980s. In 2001 Bollier co-founded Public Knowledge, a Washington advocacy organization for the public's stake in the Internet, telecom and copyright policy. For twenty-five years, until 2010, Bollier collaborated with television producer, writer and activist Norman Lear on a wide variety of non-television public affairs and political projects. Bollier blogs at Bollier.org; lives in Amherst, Massachusetts.

About the Communications and Society Program

www.aspeninstitute.org/c&s

The Communications and Society Program is an active venue for framing policies and developing recommendations in the information and communications fields. We provide a multi-disciplinary space where veteran and emerging decision-makers can develop new approaches and suggestions for communications policy. The Program enables global leaders and experts to explore new concepts, exchange insights, develop meaningful networks, and find personal growth, all for the betterment of society.

The Program's projects range across many areas of information, communications and media policy. Our activities focus on issues of open and innovative governance, public diplomacy, institutional innovation, broadband and spectrum management, as well as the future of content, issues of race and diversity, and the free flow of digital goods, services and ideas across borders.

Most conferences employ the signature Aspen Institute seminar format: approximately 25 leaders from diverse disciplines and perspectives engaged in roundtable dialogue, moderated with the goal of driving the agenda to specific conclusions and recommendations. The program distributes our conference reports and other materials to key policymakers, opinion leaders and the public in the United States and around the world. We also use the internet and social media to inform and ignite broader conversations that foster greater participation in the democratic process.

The Program's Executive Director is Charles M. Firestone. He has served in this capacity since 1989 and also as Executive Vice President of the Aspen Institute. Prior to joining the Aspen Institute, Mr. Firestone was a communications attorney and law professor who has argued cases before the United States Supreme Court. He is a former director of the UCLA Communications Law Program, first president of the Los Angeles Board of Telecommunications Commissioners, and an appellate attorney for the U.S. Federal Communications Commission.