

FOR RELEASE June 16, 2021

# Experts Doubt Ethical AI Design Will Be Broadly Adopted as the Norm Within the Next Decade

*A majority worries that the evolution of artificial intelligence by 2030 will continue to be primarily focused on optimizing profits and social control. They also cite the difficulty of achieving consensus about ethics. Many who expect progress say it is not likely within the next decade. Still, a portion celebrate coming AI breakthroughs that will improve life*

**BY** Lee Rainie, Janna Anderson and Emily A. Vogels

**FOR MEDIA OR OTHER INQUIRIES:**

Lee Rainie, Director, Internet and Technology Research  
Janna Anderson, Director, Elon University's Imagining the Internet Center  
Haley Nolan, Communications Associate  
202.419.4372  
[www.pewresearch.org](http://www.pewresearch.org)

**RECOMMENDED CITATION**

Pew Research Center, June 16, 2021. "Experts Doubt Ethical AI Design Will Be Broadly Adopted as the Norm in the Next Decade"

## About Pew Research Center

Pew Research Center is a nonpartisan fact tank that informs the public about the issues, attitudes and trends shaping America and the world. It does not take policy positions. It conducts public opinion polling, demographic research, content analysis and other data-driven social science research. The Center studies U.S. politics and policy; journalism and media; internet, science and technology; religion and public life; Hispanic trends; global attitudes and trends; and U.S. social and demographic trends. All of the center's reports are available at [www.pewresearch.org](http://www.pewresearch.org). Pew Research Center is a subsidiary of The Pew Charitable Trusts, its primary funder.

For this project, Pew Research Center worked with [Elon University's Imagining the Internet Center](#), which helped conceive the research and collect and analyze the data.

© Pew Research Center 2021

## How we did this

This is the 12th “[Future of the Internet](#)” canvassing Pew Research Center and [Elon University’s Imagining the Internet Center](#) have conducted together to get expert views about important digital issues. In this case, the questions focused on the prospects for ethical artificial intelligence (AI) by the year 2030. This is a nonscientific canvassing based on a nonrandom sample; this broad array of opinions about where current trends may lead in the next decade represents only the points of view of the individuals who responded to the queries.

Pew Research and Elon’s Imagining the Internet Center built a database of experts to canvass from a wide range of fields, choosing to invite people from several sectors, including professionals and policy people based in government bodies, nonprofits and foundations, technology businesses, think tanks and in networks of interested academics and technology innovators. The predictions reported here came in response to a set of questions in an online canvassing conducted between June 30 and July 27, 2020. In all, 602 technology innovators and developers, business and policy leaders, researchers and activists responded to at least one of the questions covered in this report. More on the methodology underlying this canvassing and the participants can be found in [the final section](#).

## Experts Doubt Ethical AI Design Will Be Broadly Adopted as the Norm Within the Next Decade

*A majority worries that the evolution of artificial intelligence by 2030 will continue to be primarily focused on optimizing profits and social control. They also cite the difficulty of achieving consensus about ethics. Many who expect progress say it is not likely within the next decade. Still, a portion celebrate coming AI breakthroughs that will improve life*

Artificial intelligence systems “understand” and shape a lot of [what happens in people’s lives](#). AI applications “speak” to people and answer questions when the name of [a digital voice assistant](#) is called out. They run the [chatbots](#) that handle customer-service issues people have with companies. They help [diagnose cancer](#) and other [medical conditions](#). They scour the use of credit cards for [signs of fraud](#), and they determine who could be a [credit risk](#).

They help [people drive](#) from point A to point B and update traffic information to shorten travel times. They are the operating system of [driverless vehicles](#). They sift applications to make recommendations about [job candidates](#). They determine the material that is offered up in people’s [newsfeeds](#) and [video choices](#).

They recognize [people’s faces](#), [translate languages](#) and suggest how to [complete people’s sentences](#) or search queries. They can “read” people’s [emotions](#). They beat them at [sophisticated games](#). They write [news stories](#), paint in the style of [Vincent Van Gogh](#) and create music that sounds quite like the [Beatles and Bach](#).

[Corporations](#) and [governments](#) are charging evermore expansively into AI development. Increasingly, nonprogrammers can set up off-the-shelf, pre-built [AI tools](#) as they prefer.

As this unfolds, a number of experts and advocates around the world have become worried about the long-term impact and implications of AI applications. They have [concerns](#) about how advances in AI will affect what it means to be human, to be productive and to exercise free will. [Dozens](#) of convenings and study groups have [issued papers](#) proposing what the tenets of ethical AI design should be, and [government working teams](#) have tried to address these issues.

In light of this, Pew Research Center and Elon University’s Imagining the Internet Center asked experts where they thought efforts aimed at creating ethical artificial intelligence would stand in the year 2030. Some 602 technology innovators, developers, business and policy leaders, researchers and activists responded to this specific question:

*By 2030, will most of the AI systems being used by organizations of all sorts employ ethical principles focused primarily on the public good?*

In response, **68%** chose the option declaring that ethical principles focused primarily on the public good *will not be* employed in most AI systems by 2030; **32%** chose the option positing that ethical principles focused primarily on the public good *will be* employed in most AI systems by 2030.

This is a nonscientific canvassing, based on a [nonrandom sample](#). The results represent *only the opinions of the individuals who responded to the queries* and are not projectable to any other population.

The bulk of this report covers these experts' written answers explaining their responses. They sounded many broad themes about the ways in which individuals and groups are accommodating to adjusting to AI systems. It is important to note that the responses were gathered in the summer of 2020 in a different cultural context amid the pandemic, before COVID-19 vaccines had been approved, at a time when racial justice issues were particularly prominent in the U.S. and before the conclusion of the U.S. presidential election.

In addition, these responses came prior to the most recent studies aimed at addressing issues in ethical AI design and development. For instance, in early 2021 the Stanford Institute for Human-Centered Artificial Intelligence released an updated [AI Index Report](#), the IEEE deepened its focus on [setting standards for AI systems](#) and the U.S. National Security Commission on AI, headed by tech leaders including Eric Schmidt, Andy Jassy, Eric Horvitz, Katharina McFarland and Robert Work, released its massive [report on accelerating innovation while defending against malign uses of AI](#).

The key themes these experts voiced in the written elaborations explaining their choices are outlined in the accompanying tables.

---

## **Worries: The main developers and deployers of AI are focused on profit-seeking and social control, and there is no consensus about what ethical AI would look like**

Even as global attention turns to the purpose and impact of artificial intelligence (AI), many experts worry that ethical behaviors and outcomes are hard to define, implement and enforce. They point out that the AI ecosystem is dominated by competing businesses seeking to maximize profits and by governments seeking to surveil and control their populations.

- **It is difficult to define “ethical” AI:** Context matters. There are cultural differences, and the nature and power of the actors in any given scenario are crucial. Norms and standards are currently under discussion, but global consensus may not be likely. In addition, formal ethics training and emphasis is not embedded in the human systems creating AI.
- **Control of AI is concentrated in the hands of powerful companies and governments driven by motives other than ethical concerns:** Over the next decade, AI development will continue to be aimed at finding ever-more-sophisticated ways to exert influence over people’s emotions and beliefs in order to convince them to buy goods, services and ideas.
- **The AI genie is already out of the bottle, abuses are already occurring, and some are not very visible and hard to remedy:** AI applications are already at work in “black box” systems that are opaque at best and, at worst, impossible to dissect. How can ethical standards be applied under these conditions? While history has shown that when abuses arise as new tools are introduced societies always adjust and work to find remedies, this time it’s different. AI is a major threat.
- **Global competition, especially between China and the U.S., will matter more to the development of AI than any ethical issues:** There is an arms race between the two tech superpowers that overshadows concerns about ethics. Plus, the two countries define ethics in different ways. The acquisition of techno-power is the real impetus for advancing AI systems. Ethics takes a back seat.

Source: Nonscientific canvassing of select experts conducted June 30-July 27, 2020.

“Experts Doubt Ethical AI Design Will Be Broadly Adopted as the Norm in the Next Decade”

PEW RESEARCH CENTER and ELON UNIVERSITY’S IMAGINING THE INTERNET CENTER, 2021

---

## Hopes: Progress is being made as AI spreads and shows its value; societies have always found ways to mitigate the problems arising from technological evolution

Artificial intelligence (AI) applications are already doing amazing things. Further breakthroughs will only add to this. The unstoppable rollout of new AI is inevitable. The development of harm-reducing strategies is also inevitable. Indeed, AI systems themselves can be used to identify and fix problems arising from unethical systems. The high-level global focus on ethical AI in recent years has been productive and is moving society toward agreement around the idea that further AI development should focus on beneficence, nonmaleficence, autonomy and justice.

- **AI advances are inevitable; we will work on fostering ethical AI design:** Imagine a future where even more applications emerge to help make people’s lives easier and safer. Health care breakthroughs are coming that will allow better diagnosis and treatment, some of which will emerge from personalized medicine that radically improves the human condition. All systems can be enhanced by AI; thus, it is likely that support for ethical AI will grow.
- **A consensus around ethical AI is emerging and open-source solutions can help:** There has been extensive study and discourse around ethical AI for several years, and it is bearing fruit. Many groups working on this are focusing on the already-established ethics of the biomedical community.
- **Ethics will evolve and progress will come as different fields show the way:** No technology endures if it broadly delivers unfair or unwanted outcomes. The market and legal systems will drive out the worst AI systems. Some fields will be faster to the mark in getting ethical AI rules and code in place, and they will point the way for laggards.

Source: Nonscientific canvassing of select experts conducted June 30-July 27, 2020.

“Experts Doubt Ethical AI Design Will Be Broadly Adopted as the Norm in the Next Decade”

PEW RESEARCH CENTER and ELON UNIVERSITY’S IMAGINING THE INTERNET CENTER, 2021

---

The respondents whose insights are shared in this report focus their lives on technology and its study. They addressed some of the toughest questions that cultures confront. How do you apply ethics to any situation? Is maximum freedom the ethical imperative or is maximum human safety? Should systems steer clear of activities that substantially impact human agency, allowing people to make decisions for themselves, or should they be set up to intervene when it seems clear that human decision-making may be harmful?

They wrestled with the meaning of such grand concepts as beneficence, nonmaleficence, autonomy and justice (the [foundational considerations](#) of bioethicists) when it comes to tech systems. Some described their approach as a comparative one: It's not whether AI systems alone produce questionable ethical outcomes, it's whether the AI systems are less biased than the current human systems and their known biases. A share of these respondents began their comments on our question by arguing that the issue is not, "What do we want AI to be?" Instead, they noted the issue should be, "What kind of humans do we want to be? How do we want to evolve as a species?"

Many experts noted that much is at stake in these arguments. AI systems will be used in ways that affect people's livelihoods and well-being – their jobs, their family environment, their access to things like housing and credit, the way they move around, the things they buy, the cultural activities to which they are exposed, their leisure activities and even what they believe to be true. One respondent noted, "Rabelais used to say, 'Science without conscience is the ruin of the soul.'"

In the next pages, we quote some of the experts who gave wide-ranging answers to our question about the future of ethical AI. After that, there is a chapter covering the responses that touched on the most troubling concerns these experts have about AI and another chapter with comments from those who expressed hope these issues will be sorted out by the year 2030 or soon thereafter.

The respondents' remarks reflect their personal positions and are not the positions of their employers; the descriptions of their leadership roles help identify their background and the locus of their expertise. Some responses are lightly edited for style and readability.

Following is a selection of some of the most comprehensive overarching responses shared by 35 of the 605 thought leaders participating in this canvassing.

*Is it possible to become comfortable with not knowing?*

**Barry Chudakov**, founder and principal of Sertain Research, said, "Before answering whether AI will mostly be used in ethical or questionable ways in the next decade, a key question for guidance

going forward will be, What is the ethical framework for understanding and managing artificial intelligence? Our ethical frameworks grew out of tribal wisdom, which was set down in so-called holy books that were the foundation of religions. These have been the ethical frameworks for the Judeo-Christian–Islamic–Buddhist world. While the humanitarian precepts of these teachings are valid today, modern technologies and artificial intelligence raise a host of AI quandaries these frameworks simply don't address. Issues such as management of multiple identities; the impingement of the virtual world on the actual world and how boundaries should be drawn – *if* boundaries should be drawn; striking a balance between screen time and real-world time; parsing, analyzing and improving the use of tracking data to ensure individual liberty; collecting, analyzing and manipulating data exhaust from online ventures to ensure citizen privacy; the use of facial recognition technologies, at the front door of homes and by municipal police forces, to stop crime. That is a small set of examples, but there are many more that extend to air and water pollution, climate degradation, warfare, finance and investment trading and civil rights.

“Our ethical book is half-written. While we would not suggest our existing ethical frameworks have no value, there are pages and chapters missing. Further, while we have a host of regulatory injunctions such as speed limits, tax rates, mandatory housing codes and permits, etc., we consider our devices so much a part of our bodies that we use them without a moment's thought for their effects upon the user. We accept the algorithms that enhance our searches and follow us around the internet and suggest another brand of facial moisturizer as a new wrinkle on a convenience and rarely give it a second thought. We do not acknowledge that our technologies change us as we use them; that our thinking and behaviors are altered by the cyber effect ([Mary Aiken](#)); that devices and gadgets don't just turn us into gadget junkies, they may abridge our humanity, compassion, empathy and social fabric. As Greg Brockman, co-founder of OpenAI, [remarked](#): ‘Now is the time to ask questions. Think about the kinds of thoughts you wish people had inventing fire, starting the industrial revolution, or [developing] atomic power.’

“Will AI mostly be used in ethical or questionable ways the next decade? I would start answering this question by referencing what Derrick de Kerckhove [described recently](#) in his ‘[Five Words for the Future](#)’: Big data is a paradigmatic change from networks and databases. The chief characteristic of big data is that the information does not exist until the question. It is not like the past where you didn't know where the information was; it was somewhere, and you just had to find it. Now, and it's a big challenge to intelligence, you create the answer by the question. (Ethics then effectively becomes) ‘How do you create the right question for the data?’ So, for AI to be mostly used in ethical ways, we must become comfortable with not knowing; with needing to ask the right question and understanding that this is an iterative process that is exploratory – not dogmatic. Beginner's mind ([Shunryu Suzuki](#)) becomes our first principle – the understanding from which ethics flows. Many of our ethical frameworks have been built on dogmatic injunctions: Thou shalt and shalt not. Thus, big data effectively reimagines ethical discourse: If until you ask



the question, you will not hear or know the answer, you proceed from unknowing. With that understanding, for AI to be used in ethical ways, and to avoid questionable approaches, we must begin by reimagining ethics itself.”

*No matter how this complex problem is tackled, responses will be piecemeal and limited*

**Mike Godwin**, former general counsel for the Wikimedia Foundation and creator of Godwin’s Law, wrote, “The most likely outcome, even in the face of increasing public discussions and convenings regarding ethical AI, will be that governments and public policy will be slow to adapt. The costs of AI-powered technologies will continue to decline, making deployment prior to privacy guarantees and other ethical safeguards more likely. The most likely scenario is that some kind of public abuse of AI technologies will come to light, and this will trigger reactive limitations on the use of AI, which will either be blunt-instrument categorical restrictions on its use or (more likely) a patchwork of particular ethical limitations addressed to particular use cases, with unrestricted use occurring outside the margins of these limitations.”

*Sometimes there are no good answers, only varieties of bad outcomes*

**Jamais Cascio**, research fellow at the Institute for the Future, observed, “I expect that there will be an effort to explicitly include ethical systems in AI that have direct interaction with humans but largely in the most clear-cut, unambiguous situations. The most important ethical dilemmas are ones where the correct behavior by the machine is situational: Health care AI that intentionally lies to memory care patients rather than re-traumatize them with news of long-dead spouses; military AI that recognizes and refuses an illegal order; all of the ‘[trolley problem](#)’-type dilemmas where there are no good answers, only varieties of bad outcomes. But, more importantly, the vast majority of AI systems will be deployed in systems for which ethical questions are indirect, even if they ultimately have outcomes that could be harmful.

“High-frequency trading AI will not be programmed to consider the ethical results of stock purchases. Deepfake AIs will not have built-in restrictions on use. And so forth.

“What concerns me the most about the wider use of AI is the lack of general awareness that digital systems can only manage problems that can be put in a digital format. An AI can’t reliably or consistently handle a problem that can’t be quantified. There are situations and systems for which AI is a perfect tool, but there are important arenas – largely in the realm of human behavior and personal interaction – where the limits of AI can be problematic. I would hate to see a world where some problems are ignored because we can’t easily design an AI response.”

*AI is more capable than humans of delivering unemotional ethical judgment*

**Marcel Fafchamps**, professor of economics and senior fellow at the Center on Democracy, Development and the Rule of Law at Stanford University, commented, “AI is just a small cog in a big system. The main danger currently associated with AI is that machine learning reproduces past discrimination – e.g., in judicial processes for setting bail, sentencing or parole review. But if there hadn’t been discrimination in the first place, machine learning would have worked fine. This means that AI, in this example, offers the possibility of improvement over unregulated social processes.

“A more subtle danger is when humans are actually more generous than machine-learning algorithms. For instance, it has been shown that judges are more lenient toward first offenders than machine learning in the sense that machine learning predicts a high probability of reoffending, and this probability is not taken into account by judges when sentencing. In other words, judges give first offenders ‘a second chance,’ a moral compass that the algorithm lacks. But, more generally, the algorithm only does what it is told to do: If the law that has been voted on by the public ends up throwing large fractions of poor young males in jail, then that’s what the algorithm will implement, removing the judge’s discretion to do some minor adjustment at the margin. Don’t blame AI for that: Blame the criminal justice system that has been created by voters.

“A more pernicious development is the loss of control people will have over their immediate environment, e.g., when their home appliances will make choices for them ‘in their interest.’ Again, this is not really new. But it will occur in a new way. My belief is as follows:

- 1) By construction, AI implicitly or explicitly integrates ethical principles, whether people realize it or not. This is most easily demonstrated in the case of self-driving cars but will apply to all self-‘something’ technology, including health care AI apps, for instance. A self-driving car must, at some point, decide whether to protect its occupants or protect other people on the road. A human driver would make a choice partly based on social preferences, as has been shown for instance in ‘[The Moral Machine Experiment](#)’ (Nature, 2018), partly based on moral considerations (e.g., did the pedestrian have the right to be on the path of the car at that time? In the March 2018 fatality in Tempe, Florida, a human driver could have argued that the pedestrian ‘appeared out of nowhere’ in order to be exonerated).
- 2) The fact that AI integrates ethical principles does not mean that it integrates ‘your’ preferred ethical principles. So the question is not whether it integrates ethical principles, but which ethical principles it integrates.

“Here, the main difficulty will be that human morality is not always rational or even predictable. Hence, whatever principle is built into AI, there will be situations in which the application of that

ethical principle to a particular situation will be found unacceptable by many people, no matter how well-meant that principle was. To minimize this possibility, the guideline at this point in time is to embed into AI whatever factual principles are applied by courts. This should minimize court litigation. But, of course, if the principles applied by courts are detrimental to certain groups, this will be reproduced by AI.

“What would be really novel would be to take AI as an opportunity to introduce more coherent ethical judgment than what people make based on an immediate emotional reaction. For instance, if the pedestrian in Tempe had been a just-married young bride, a pregnant woman or a drug offender, people would judge the outcome differently, even though, at the moment of the accident, this could not be deduced by the driver, whether human or AI. That does not make good sense: An action cannot be judged differently based on a consequence that was materially unpredictable to the perpetrator. AI can be an opportunity to improve the ethical behavior of cars (and other apps), based on rational principles instead of knee-jerk emotional reaction.”

*Global politics and rogue actors are oft-ignored aspects to consider*

**Amy Webb**, founder of the Future Today Institute, wrote, “We’re living through a precarious moment in time. China is shaping the world order in its own image, and it is exporting its technologies and surveillance systems to other countries around the world. As China expands into African countries and throughout Southeast Asia and Latin America, it will also begin to eschew operating systems, technologies and infrastructure built by the West. China has already announced that it will no longer use U.S.-made computers and software. China is rapidly expanding its 5G and mobile footprints. At the same time, China is drastically expanding its trading partners. While India, Japan and South Korea have plenty of technologies to offer the world, it would appear as though China is quickly ascending to global supremacy. At the moment, the U.S. is enabling this, and our leaders do not appear to be thinking about the long-term consequences.

“When it comes to AI, we should pay close attention to China, which has talked openly about its plans for cyber sovereignty. But we should also remember that there are cells of rogue actors who could cripple our economies simply by mucking with the power or traffic grids, causing traffic spikes on the internet or locking us out of our connected home appliances. These aren’t big, obvious signs of aggression, and that is a problem for many countries, including the United States. Most governments don’t have a paradigm describing a constellation of aggressive actions. Each action on its own might be insignificant. What are the escalation triggers? We don’t have a definition, and that creates a strategic vulnerability.”

*Concentrated wealth works against hope for a Human Spring and social justice*

**Stowe Boyd**, consulting futurist expert in technological evolution and the future of work, noted, “I have projected a social movement that would require careful application of AI as one of several major pillars. I’ve called this the Human Spring, conjecturing that a worldwide social movement will arise in 2023, demanding the right to work and related social justice issues, a massive effort to counter climate catastrophe, and efforts to control artificial intelligence. AI, judiciously used, can lead to breakthroughs in many areas. But widespread automation of many kinds of work – unless introduced gradually, and not as fast as profit-driven companies would like – could be economically destabilizing.

“I’m concerned that AI will most likely be concentrated in the hands of corporations who are in the business of concentrating wealth for their owners and not primarily driven by bettering the world for all of us. AI applied in narrow domains that are really beyond the reach of human cognition – like searching for new ways to fold proteins to make new drugs or optimizing logistics to minimize the number of miles that trucks drive everyday – are sensible and safe applications of AI. But AI directed toward making us buy consumer goods we don’t need or surveilling everyone moving through public spaces to track our every move, well, that should be prohibited.”

*The principal use of AI is likely to remain convincing people to buy things they don’t need*

**Jonathan Grudin**, principal researcher with the Natural Interaction Group at Microsoft Research, said, “The past quarter-century has seen an accelerating rise of online bad actors (not all of whom would agree they are bad actors) and an astronomical rise in the costs of efforts to combat them, with AI figuring in this. We pose impossible demands: We would like social media to preserve individual privacy but also identify Russian or Chinese hackers that will require sophisticated construction of individual behavior patterns.

“The principal use of AI is likely to be finding ever more sophisticated ways to convince people to buy things that they don’t really need, leaving us deeper in debt with no money to contribute to efforts to combat climate change, environmental catastrophe, social injustice and inequality and so on.”

*User-experience designers must play a key role in shaping human control of systems*

**Ben Shneiderman**, distinguished professor of computer science and founder of Human Computer Interaction Lab, University of Maryland, commented, “Ethical principles (responsibility, fairness, transparency, accountability, auditability, explainable, reliable, resilient, safe, trustworthy) are a good starting point, but *much* more is needed to bridge the gap with the realities of practice in software engineering, organization management and independent oversight. ... I see promising early signs. A simple step is a flight data recorder for every robot and AI system. The methods that have made civil aviation so safe could be adopted to recording what every robot

and AI system does, so that when errors occur, the forensic investigation will have the data it needs to understand what went wrong and make enforceable measurable testable improvements. AI applications can bring many benefits, but they are more likely to succeed when user-experience designers have a leading role in shaping human control of highly automated systems.”

*AI systems today fetishize efficiency, scale and automation; should embrace social justice*

**danah boyd**, founder and president of the Data & Society Research Institute, and principal researcher at Microsoft, explained, “We misunderstand ethics when we think of it as a binary, when we think that things can be ethical or unethical. A true commitment to ethics is a commitment to understanding societal values and power dynamics – and then working toward justice.

“Most data-driven systems, especially AI systems, entrench existing structural inequities into their systems by using training data to build models. The key here is to actively identify and combat these biases, which requires the digital equivalent of reparations. While most large corporations are willing to talk about fairness and eliminating biases, most are not willing to entertain the idea that they have a responsibility for data justice. These systems are also primarily being built within the context of late-stage capitalism, which fetishizes efficiency, scale and automation. A truly ethical stance on AI requires us to focus on augmentation, localized context and inclusion, three goals that are antithetical to the values justified by late-stage capitalism. We cannot meaningfully talk about ethical AI until we can call into question the logics of late-stage capitalism.”

*If we don't fix this, we can't even imagine how bad it will get when AI is creating AI*

**Gary A. Bolles**, chair for the future of work at Singularity University, responded, “I hope we will shift the mindset of engineers, product managers and marketers from ethics and human centrality as a tack-on after AI products are released, to a model that guarantees ethical development from inception. Everyone in the technology development food chain will have the tools and incentives to ensure the creation of ethical and beneficial AI-related technologies, so there is no additional effort required. Massive energy will be focused on new technologies that can sense when new technologies are created that violate ethical guidelines and automatically mitigate those impacts.

“Humans will gain tremendous benefits as an increasing amount of technology advocates for them automatically. My concerns: None of this may happen, if we don't change the financial structure. There are far too many incentives – not just to cut corners but to deliberately leave out ethical and inclusive functions, because those technologies aren't perceived to make as much money, or to deliver as much power, as those that ignore them. If we don't fix this, we can't even imagine how much off the rails this can go once AI is creating AI.”

*Even ethical people think in terms of using tech on humans instead of the opposite*

**Douglas Rushkoff**, well-known media theorist, author and professor of media at City University of New York, wrote, “Why should AI become the very first technology whose development is dictated by moral principles? We haven’t done it before, and I don’t see it happening now. Most basically, the reasons why I think AI won’t be developed ethically is because AI is being developed by companies looking to make money – not to improve the human condition. So, while there will be a few simple AIs used to optimize water use on farms or help manage other limited resources, I think the majority is being used on people.

“My concern is that even the ethical people still think in terms of using technology on human beings instead of the other way around. So, we may develop a ‘humane’ AI, but what does that mean? It extracts value from us in the most ‘humane’ way possible?”

*AIs built to be reciprocally competitive could keep an eye on each other, report bad things*

**David Brin**, physicist, futures thinker and author of “[Earth](#)” and “[Existence](#),” commented, “Isaac Asimov in his ‘Robots’ series conceived a future when ethical matters would be foremost in the minds of designers of AI brains, not for reasons of judiciousness, but in order to quell the fears of an anxious public, and hence Asimov’s famed ‘[Three Laws of Robotics](#).’ No such desperate anxiety about AI seems to surge across today’s populace, perhaps because we are seeing our AI advances in more abstract ways, mostly on screens and such, not in powerful, clanking mechanical men. Oh, there are serious conferences on this topic. I’ve participated in many. Alas, statements urging ethical consideration in AI development are at best palliatives. I am often an outlier, proposing that AIs’ ‘ethical behavior’ be promoted the way it is in most humans – especially most males – via accountability.

“If AIs are many and diverse and reciprocally competitive, then it will be in their individual interest to keep an eye on each other and report bad things, because doing so will be to their advantage. This depends on giving them a sense of separate individuality. It is a simple recourse, alas seldom even discussed.”

*Within the next 300 years, humans will be replaced by their own sentient creations*

**Michael G. Dyer**, professor emeritus of computer science at UCLA, expert in natural language processing, responded, “Ethical software is an ambiguous notion and includes:

- 1) Software that makes choices normally considered to be in the ethical/moral sphere. An example of this would be software that makes (or aids in making) decisions concerning punishments for crimes or software that decides whether or not some applicant is worthy

of some desirable job or university. This type of task can be carried out via classification and the field of classification (and learning classification from data) is already well developed and could be (and is being) applied to tasks that relate to the moral sphere.

- 2) Software that decides who receives a negative (vs. positive) outcome in zero-sum circumstances. A classic case is that of a driverless car in which the driving software will have to decide whether to protect the passenger or the pedestrian in an immediately predicted accident.
- 3) Software that includes ethics/morality when planning to achieve goals (a generalization of 2). I am personally more interested in this type of AI software.

“Consider that you, in the more distant future, own a robot and you ask it to get you an umbrella because you see that it might rain today. Your robot goes out and sees a little old lady with an umbrella. Your robot takes the umbrella away from her and returns to hand it to you. That is a robot without ethical reasoning capability. It has a goal, and it achieves that goal without considering the effect of its plan on the goals of other agents; therefore, ethical planning is a much more complicated form of planning because it has to take into account the goals and plans of other agents. Another example. You tell your robot that Mr. Mean is your enemy (vs. friend). In this case, the robot might choose a plan to achieve your goal that, at the same time, harms some goal of Mr. Mean.

“Ethical reasoning is more complicated than ethical planning, because it requires building inverted ‘trees’ of logical (and/or probabilistic) support for any beliefs that themselves might support a given plan or goal. For example, if a robot believes that goal G1 is wrong, then the robot is not going to plan to achieve G1. However, if the robot believes that agent A1 has goal G1, then the robot might generate a counterplan to block A1 in executing the predicted plan (or plans) of agent A1 to achieve G1 (which is an undesirable goal for the robot). Software that is trained on data to categorize/classify already exists and is extremely popular and has been and will continue to be used to also classify people (does Joe go to jail for five years or 10 years? Does Mary get that job? etc.).

“Software that performs sophisticated moral reasoning will not be widespread by 2025 but will become more common in 2030. (You asked for predictions, so I am making them.) Like any technology, AI can be used for good or evil. Face recognition can be used to enslave everyone (à la Orwell’s ‘Nineteen Eighty-Four’) or to track down serial killers. Technology depends on how humans use it (since self-aware sentient robots are still at least 40 years away). It is possible that a ‘critical mass’ of intelligence could be reached, in which an AI entity works on improving its own intelligent design, thus entering into a positive feedback loop resulting rapidly in a super-intelligent form of AI (e.g., see [D. Lenat’s Heurisko](#) work done years ago, in which it invented not only various structures but also invented new heuristics of invention). A research project that also

excites me is that of computer modeling of the [human connectome](#). One could then build a humanoid form of intelligence without understanding how human neural intelligence actually works (which could be quite dangerous).

“I am concerned and also convinced that, at some point within the next 300 years, humanity will be replaced by its own creations, once they become sentient and more intelligent than ourselves. Computers are already smarter at many tasks, but they are not an existential threat to humanity (at this point) because they lack sentience. AI chess- (and now Go-) playing systems beat world grand masters, but they are not *aware* that they are playing a game. They currently lack the ability to converse (in human natural languages, such as English or Chinese) about the games they play, and they lack their own autonomous goals. However, subfields of AI include machine learning and computational evolution. AI systems are right now being evolved to survive (and learn) in simulated environments and such systems, if given language comprehension abilities (being developed in the AI field of natural language processing), would then achieve a form of sentience (awareness of one’s awareness and ability to communicate that awareness to others, and an ability to debate beliefs via reasoning, counterfactual and otherwise, e.g., see work of [Judea Pearl](#)).”

*There are challenges, but better systems will emerge to improve the human condition*

**Marjory S. Blumenthal**, director of the science, technology and policy program at RAND Corporation, observed, “This is the proverbial onion; there is no simple answer. Some of the challenge is in the lifecycle – it begins with how the data are collected, labeled (if they are for training) and then used, possibly involving different actors with different views of what is ethical. Some of the challenges involve the interest-balancing of companies, especially startups, that have always placed function and product delivery over security and privacy.

“Some of the challenges reflect the fact that, in addition to privacy and security for some applications, safety is also a concern (and there are others). Some of the challenges reflect the fact that, even with international efforts like that of IEEE, views of what ethics are appropriate differ around the world.

“Today’s AI mania implies that a lot of people are rushing to be able to say that they use or produce AI, and anything rushed will have compromises. Notwithstanding the concerns, the enthusiasm for AI builds on long histories improving processing hardware, data-handling capability and algorithms. Better systems for education and training should be available and should enable the kind of customization long promised but seldom achieved. Aids to medical diagnoses should become more credible, along with aids to development of new therapies. The support provided by today’s ‘smart speakers’ should become more meaningful and more useful (especially if clear attention to privacy and security comes with the increased functionality).”



*Ethical AI is definitely being addressed, creating a rare opportunity to deploy it positively*

**Ethan Zuckerman**, director of MIT's Center for Civic Media and associate professor at the MIT Media Lab, commented, "The activists and academics advocating for ethical uses of AI have been remarkably successful in having their concerns aired even as harms of misused AI are just becoming apparent. The campaigns to stop the use of facial recognition because of racial biases is a precursor of a larger set of conversations about serious ethical issues around AI. Because these pioneers have been so active in putting AI ethics on the agenda, I think we have a rare opportunity to deploy AI in a vastly more thoughtful way that we otherwise might have."

*Failures despite good intentions loom ahead, but society will still reap benefits*

**Vint Cerf**, Internet Hall of Fame member and vice president and chief internet evangelist at Google, observed, "There will be a good-faith effort, but I am skeptical that the good intentions will necessarily result in the desired outcomes. Machine learning is still in its early days, and our ability to predict various kinds of failures and their consequences is limited. The ML design space is huge and largely unexplored. If we have trouble with ordinary software whose behavior is at least analytic, ML is another story. And our track record on normal software stinks (buggy code!). We are, however, benefiting enormously from many ML applications, including speech recognition and language translation, search efficiency and effectiveness, medical diagnosis, exploration of massive data to find patterns, trends and unique properties (e.g., pharmaceuticals). Discovery science is benefiting (e.g., finding planets around distant stars). Pretty exciting stuff."

*Commitment can bring positive results; all bets are off when it comes to weapons*

**Susan Etlinger**, industry analyst for Altimeter, wrote, "AI is, fundamentally, an idea about how we can make machines that replicate some aspects of human ability. So, we should expect to see ethical norms around bias, governance and transparency become more common, much the same way we've seen the auto industry and others adopt safety measures like seatbelts, airbags and traffic signals over time. But of course people are people, so for every ethical principle there will always be someone who ignores or circumvents it."

"I'm heartened by some of the work I've seen from the large tech companies. It's not consistent, it's not enough, but there are enough people who are genuinely committed to using technology responsibly that we will see some measure of positive change. Of course, all claims of AGI – automated general intelligence – are immediately suspect, not only because it's still hypothetical at this point, but because we haven't even ironed out the governance implications of automation. And all bets are off when we are talking about AI-enabled weaponry, which will require a level of diplomacy, policy and global governance similar to nuclear power."

*More transparency in digital and human systems can emerge from all of this*

**Esther Dyson**, internet pioneer, journalist, entrepreneur and executive founder of Wellville, responded, “With luck, we’ll increase transparency around what AI is doing (as well as around what *people* are doing), because it will be easier to see the impact of decisions made by both people and algorithms. Cue the research about what time of day you want to go to trial (e.g., before or after the judge has lunch). The more we use AI to reveal such previously hidden patterns, the better for us all. So, a lot depends on society’s willingness to look at the truth and to act/make decisions accordingly. With luck, a culture of transparency will cause this to happen. But history shows that a smooth trajectory toward enlightenment is unlikely.”

*AI making decisions on its own is an understandable but possibly unstoppable worry*

**Brad Templeton**, internet pioneer, futurist, activist and chair emeritus of the Electronic Frontier Foundation, said, “For now, at least, and probably to 2030, AI is a tool, not an actor in its own right. It will not be good or evil, but it will be used with good and evil intent and also for unintended reasons. But this is not a question for a simple survey. People are writing books about this question. To go into a few of the popular topics: The use of AI to replace jobs is way overblown. We have 150 years of Chicken Little predictions that machines would take all the jobs, and they’ve always been wrong. First, that in most cases they didn’t take the jobs, or that we would be bothered when they did. There are more bank tellers today than in 1970, it is reported. At the same time, half of us worked in agriculture in 1900, and now a small percentage do.

“The privacy worries are real, including the undefined threat that AI in the future will be able to examine the data of the present (which we are recording, but can’t yet process) in ways that will come back to bite you. I call this the threat of ‘time travelling robots from the future.’ They don’t really go back in time, but the AI of the future can affect what you do today. The fears of bias are both real and overblown. Yes, we will encode our biases into AIs. At the same time, the great thing about computers is once you see a problem you can usually fix it. Studies have shown it’s nearly impossible for humans to correct their biases, even when aware of them. For machines, that will be nothing. Strangely, when some people hear ‘AIs will be able to do one-third of the tasks you do in your work,’ some of them react with fear of losing a job. The other group reacts with, ‘Shut up and take my money!’ – they relish not having to do those tasks.

“When we start to worry about AI with agency – making decisions on its own – it is understandable why people worry about that. Unfortunately, relinquishment of AI development is not a choice. It just means the AIs of the future are built by others, which is to say your rivals. You can’t pick a world without AI; you can only pick a world where you have it or not.”

*It will be very difficult to predict what will be important and how things will work*

**John L. King**, a professor at the University of Michigan School of Information, commented, “There will be a huge increase in the discussion of revolutionary AI in daily life, but on closer inspection, things will be more incremental than most imagine. The ethical issues will sneak up on us as we move more slowly than people think when, suddenly, we cross some unforeseen threshold (it will be nonlinear) and things get serious. It will be very difficult to predict what will be important and how things will work.”

*The public must take action to better align corporate interests with the public good*

**David Karger**, professor at MIT’s Computer Science and Artificial Intelligence Laboratory, said, “The question as framed suggests that AI systems will be thinking by 2030. I don’t believe that’s the case. In 2030, AI systems will continue to be machines that do what their human users tell them to do. So, the important question is whether their *human users* will employ ethical principles focused primarily on the public good. Since that isn’t true now, I don’t expect it will be true in 2030 either. Just like now, most users of AI systems will be for-profit corporations, and just like now, they will be focused on profit rather than social good. These AI systems will certainly enable corporations to do a *much better* job of extracting profit, likely with a corresponding *decrease* in public good, unless the public itself takes action to better align the profit-interests of these corporations with the public good.

“In great part, this requires the passage of laws constraining what corporations can do in pursuit of profit; it also means the government quantifying and paying for public goods so that companies have a profit motive in pursuing them.

“Even in this time of tremendous progress, I find little to excite me about AI systems. In our frenzy to enhance the capabilities of machines, we are neglecting the existing and latent capabilities of human beings, where there is just as much opportunity for progress as there is in AI. We should be directing far more attention to research on helping people learn better, helping them interact online better and helping them make decisions better.”

*AI tools must be designed with input from diverse groups of those affected by them*

**Beth Noveck**, director, NYU Governance Lab and its MacArthur Research Network on Opening Governance, responded, “Successful AI applications depend upon the use of large quantities of data to develop algorithms. But a great deal of human decision-making is also involved in the design of such algorithms, beginning with the choice about what data to include and exclude. Today, most of that decision-making is done by technologists working behind closed doors on proprietary private systems.

“If we are to realize the positive benefits of AI, we first need to change the governance of AI and ensure that these technologies are designed in a more participatory fashion with input and oversight from diverse audiences, including those most affected by the technologies. While AI can help to increase the efficiency and decrease the cost, for example, of interviewing and selecting job candidates, these tools need to be designed with workers lest they end up perpetuating bias.

“While AI can make it possible to diagnose disease better than a single doctor can with the naked eye, if the tool is designed only using data from white men, it may be less optimal for diagnosing diseases among Black women. Until we commit to making AI more transparent and participatory, we will not realize its positive potential or mitigate the significant risks.”

*There is no way to force unethical players to follow the ethics playbook*

**Sam S. Adams**, a 24-year veteran of IBM, now working as a senior research scientist in artificial intelligence for RTI International, architecting national-scale knowledge graphs for global good, wrote, “The AI genie is completely out of the bottle already, and by 2030 there will be dramatic increases in the utility and universal access to advanced AI technology. This means there is practically no way to force ethical use in the fundamentally unethical fractions of global society.

“The multimillennial problem with ethics has always been: Whose ethics? Who decides and then who agrees to comply? That is a fundamentally human problem that no technical advance or even existential threat will totally eliminate. Basically, we are stuck with each other and hopefully at least a large fraction will try to make the best of it. But there is too much power and wealth available for those who will use advanced technology unethically, and universal access via cloud, IoT [Internet of Things] and open-source software will make it all too easy for an unethical player to exploit.

“I believe the only realistic path is to provide an open playing field. That universal access to the technology at least arms both sides equally. This may be the equivalent of a mutually assured destruction policy, but to take guns away from the good guys only means they can’t defend themselves from the bad guys anymore.”

*AI for personalized medicine could lead to the ‘Brave New World’ of Aldous Huxley*

**Joël Colloc**, professor of computer sciences at Le Havre University, Normandy, responded, “Most researchers in the public domain have an ethical and epistemological culture and do research to find new ways to improve the lives of humanity. Rabelais used to say, ‘Science without conscience is the ruin of the soul.’ Science provides powerful tools. When these tools are placed only in the hands of private interests, for the sole purpose of making profit and getting even more money and power, the use of science can lead to deviances and even uses against the states

themselves – even though it is increasingly difficult for these companies to enforce the laws, which do not necessarily have the public interest as their concern. It all depends on the degree of wisdom and ethics of the leader.

“Hope: Some leaders have an ethical culture and principles that can lead to interesting goals for citizens. All applications of AI (especially when they are in the field of health, the environment, etc.) should require a project submission to an ethics board composed of scientists and respect general charters of good conduct. A monitoring committee can verify that the ethics and the state of the art are well respected by private companies.

“The concern is what I see: clinical trials on people in developing countries where people are treated like guinea pigs under pretext that one claims to discover knowledge by applying deep learning algorithms. This is disgusting. AI can offer very good tools, but it can also be used to profile and to sort, monitor and constrain fundamental freedoms as seen in some countries. On AI competition, it is the acceptability and ability to make tools that end users find useful in improving their lives that will make the difference. Many gadgets or harmful devices are offered.

“I am interested in mastering time in clinical decision-making in medicine and how AI can take it into account. What scares me most is the use of AI for personalized medicine that, under the guise of prevention, will lead to a new eugenics and all the cloning drifts, etc., that can lead to the ‘Brave New World’ of Aldous Huxley.”

*We have no institutions that can impose ethical constraints upon AI designers*

**Susan Crawford**, a professor at Harvard Law School and former special assistant in the Obama White House for Science Technology and Innovation Policy, noted, “For AI, just substitute ‘digital processing.’ We have no basis on which to believe that the animal spirits of those designing digital processing services, bent on scale and profitability, will be restrained by some internal memory of ethics, and we have no institutions that could impose those constraints externally.”

*Unless the idea that all tech is neutral is corrected there is little hope*

**Paul Jones**, professor emeritus of information science at the University of North Carolina, Chapel Hill, observed, “Unless, as I hope happens, the idea that all tech is neutral is corrected, there is little hope or incentive to create ethical AI. Current applications of AI and their creators rarely interrogate ethical issues except as some sort of parlor game. More often I hear data scientists disparaging what they consider ‘soft sciences’ and claiming that their socially agnostic engineering approach or their complex statistical approach is a ‘hard science.’ While I don’t fear an AI war, a [Capek-like robot uprising](#), I do fear the tendency not to ask the tough questions of AI

– not just of general AI, where most of such questions are entertained, but in narrow AI where most progress and deployment are happening quickly.

“I love to talk to Google about music, news and trivia. I love my home being alert to my needs. I love doctors getting integrated feedback on lab work and symptoms. I could not now live without Google Maps. But I am aware that ‘We become what we behold. We shape our tools and then our tools shape us,’ as Father John Culkin reminded us.

“For most of us, the day-to-day conveniences of AI by far outweigh the perceived dangers. Dangers will come on slow and then cascade before most of us notice. That’s not limited to AI. Can AI help us see the dangers before they cascade? And if AI does, will it listen and react properly?”

*AI can and will be engineered toward utopian and dystopian ends*

**Dan S. Wallach**, a professor in the systems group at Rice University’s Department of Computer Science, said, “Building an AI system that works well is an exceptionally hard task, currently requiring our brightest minds and huge computational resources. Adding the additional constraint that they’re built in an ethical fashion is even harder yet again.

“Consider, for example, an AI intended for credit rating. It would be unethical for that AI to consider gender, race or a variety of other factors. Nonetheless, even if those features are explicitly excluded from the training set, the training data might well encode the biases of human raters, and the AI could pick up on secondary features that infer the excluded ones (e.g., silently inferring a proxy variable for race from income and postal address).

“Consider further the use of AI systems in warfare. The big buzzword today is ‘autonomy,’ which is to say, weapon systems that can make on-the-fly tactical decisions without human input while still following their orders. An ethical stance might say that we should never develop such systems, under any circumstances, yet exactly such systems are already in conception or development now and might well be used in the field by 2030.

“Without a doubt, AI will do great things for us, whether it’s self-driving cars that significantly reduce automotive death and injury, or whether it is computers reading radiological scans and identifying tumors earlier in their development than any human radiologist might do reliably. But AI will also be used in horribly dystopian situations, such as China’s rollout of facial-recognition camera systems throughout certain western provinces in the country. As such, AI is just a tool, just like computers are a tool. AI can and will be engineered toward utopian and dystopian ends.”

*Government shouldn't regulate it until it is dedicated to serving the needs of the people*

**Shel Israel**, Forbes columnist and author of many books on disruptive technologies, commented, “Most developers of AI are well-intentioned, but issues that have been around for over 50 years remain unresolved:

- 1) Should AI replace people or back them up? I prefer the latter in many cases. But economics drive business and returns to shareholders. So current trends will continue for more than five years because the problems will not be overwhelmingly obvious for more years than five.
- 2) Google already knows who we are, where we are, the context of our activities, who we are with. Five years from now, technology will know our health, when we will die, if it is by natural causes, and so on down the line. Will AI help a patient by warning her/him of a cancer likelihood so they can get help, or an employer so they can get rid of those employees before they become an expense? I think both will occur, so AI will make things both better and worse.
- 3) The technology itself is neither good nor evil. It is just a series of algorithms. It is how people will use it that will make a difference. Will government regulate it better? I doubt it. Should it? Not until we can have better governments who are more dedicated to serving the needs of everyday people.”

*We are ill-prepared for the onslaught and implications of bad AI applications*

**Calton Pu**, professor and chair in the School of Computer Science at Georgia Tech, wrote, “The main worry about the development of AI and ML (machine learning) technologies is the current AI/ML practice of using fixed training data (ground truth) for experimental evaluation as proof that they work. This proof is only valid for the relatively small and fixed training datasets. The gap between the limited ground truth and the actual reality has severely restricted the practical applicability of AI/ML systems, which rely on human operators to handle the gap. For example, the chatbots used in customer support contact centers can only handle the subset of most common conversations. ...

“There is a growing gap between AI systems and the evolving reality, which explains the difficulties in the actual deployment of autonomous vehicles. This growing gap appears to be a blind spot for current AI/ML researchers and companies. With all due respect to the billions of dollars being invested, it is an inconvenient truth. As a result of this growing gap, the ‘good’ AI applications will see decreasing applicability, as their ground truth lags behind the evolving actual reality. However, I imagine the bad guys will see this growing gap soon and utilize it to create ‘bad’ AI applications by feeding their AI systems with distorted ground truth through skillful

manipulations of training data. This can be done with today's software tools. These bad AI applications can be distorted in many ways, one of them being unethical. With the AI/ML research community turning a blind eye to the growing gap, we will be ill-prepared for the onslaught of these bad AI applications. An early illustration of this kind of attack was [Microsoft's Tay chatbot](#), introduced in 2016 and deactivated within one day due to inappropriate postings learned from purposeful racists interactions.

“The global competition over AI systems with fixed training data is a game. These AI systems compete within the fixed ground truth and rules. Current AI/ML systems do quite well with games with fixed rules and data, e.g., AlphaGo. However, these AI systems modeled after games are unaware of the growing gap between their ground truth (within the game) and the evolving actual reality out there. ... To change these limitations, the ML/AI community and companies will need to face the inconvenient truth, the growing gap, and start to work on the growing gap instead of simply shutting down AI systems that no longer work (when the gap grew too wide), which has been the case of the Microsoft Tay chatbot and Google Flu Trends, among others.”

*AI may not be as useful in the future due to its dependency on past data and patterns*

**Greg Sherwin**, vice president for engineering and information technology at Singularity University, responded, “Explainable AI will become ever more important. As privileged classes on the edges get caught up on the vortex of negative algorithmic biases, political will must shift toward addressing the challenges of algorithmic oppression for all. For example, companies will be sued – unsuccessfully at first – for algorithmic discrimination. Processes for redress and appeal will need to be introduced to challenge the decisions of algorithms. Meanwhile, the hype cycle will drop for the practical value of AI.

“As the world and society become more defined by VUCA [volatile, uncertain, complex, ambiguous] forces, the less AI will be useful given its complete dependency on past data, existing patterns and its ineffectiveness in novel situations. AI will simply become much like what computers were to society a couple decades ago: algorithmic tools in the background, with inherent and many known flaws (bugs, etc.), that are no longer revered for their mythical innovative novelty but are rather understood in context within limits, within boundaries that are more popularly understood.”

*How will AI be used to assess, direct, control and alter human interaction?*

**Kathleen M. Carley**, director of the Center for Computational Analysis of Social and Organizational Systems at Carnegie Mellon University, commented, “While there is a move toward ethical AI, it is unlikely to be realized in the next decade. First, there are a huge number of legacy systems that would need to be changed. Second, what it means for AI to be ethical is not well



understood; and once understood, it is likely to be the case that there are different ethical foundations that are not compatible with each other. Which means that AI might be ethical by one framework but not by another. Third, for international conflict and for conflict with nonstate actors, terror groups and crime groups – there will be AI on both sides. It is unlikely that both sides would employ the same ethical frameworks.

“What gives me the most hope is that most people, regardless of where they are from, want AI and technology in general to be used in more ethical ways. What worries me the most is that, without a clear understanding of the ramifications of ethical principles, we will put in place guidelines and policies that will cripple the development of new technologies that would better serve humanity.

“AI will save time, allow for increased control over your living space, do boring tasks, help with planning, auto-park your car, fill out grocery lists, remind you to take medicines, support medical diagnosis, etc.

“The issues that are both exciting and concerning center on how AI will be used to assess, direct, control and alter human interaction and discourse. Where AI meets human social behavior is a difficult area. Tools that auto-declare messages as disinformation could be used by authoritarian states to harm individuals.”

*We don't really know what human, ethical, public-interest decision-making looks like*

**Chris Savage**, a leading expert in legal and regulatory issues based in Washington, D.C., wrote, “AI is the new social network, by which I mean: Back in 2007 and 2008, it was easy to articulate the benefits of robust social networking, and people adopted the technology rapidly, but its toxic elements – cognitive and emotional echo chambers, economic incentives of the platforms to drive engagement via stirred-up negative emotions, rather than driving increased awareness and acceptance (or at least tolerance) of others – took some time to play out.

“Similarly, it is easy to articulate the benefits of robust and ubiquitous AI, and those benefits will drive substantial adoption in a wide range of contexts.

“But we simply do not know enough about what ‘ethical’ or ‘public-interested’ algorithmic decision-making looks like to build those concepts into actually deployed AI (actually, we don’t actually know enough about what human ‘ethical’ and ‘public-interested’ decision-making looks like to effectively model it). Trying to address those concerns will take time and money on the part of the AI developers, with no evident return on that expenditure. So, it won’t happen, or will be short-changed, and – as with social media – I predict a ‘Ready, Fire, Aim’ scenario for the

deployment of AI. On a longer timescale – give me 50 years instead of 10 – I think AI will be a net plus even in ethical/public interest terms. But the initial decade or so will be messy.”

*Why the moral panic? Does this really require an entirely new branch of ethics?*

**Jeff Jarvis**, director of the Tow-Knight Center and professor of journalism innovation at City University of New York, said, “AI is an overbroad label for sets of technical abilities to gather, analyze and learn from data to predict behavior, something we have done in our heads since some point in our evolution as a species. We did likewise with computers once we got them, getting help looking for correlations, asking ‘what if?’ and making predictions.

“Now, machines will make some predictions – often without explanation – better than we could, and that is leading to a level of moral panic sufficient to inspire questions such as this.

“The ethical challenges are not vastly different than they have ever been: Did you have permission to gather the data you did? Were you transparent about its collection and use? Did you allow people a choice in taking part in that process? Did you consider the biases and gaps in the data you gathered? Did you consider the implications of acting on mistaken predictions? And so on. I have trouble seeing this treated as if it is an entirely new branch of ethics, for that brings an air of mystery to what should be clear and understandable questions of responsibility.”

*Perhaps traditional notions of civil liberties need to be revised and updated*

**David Krieger**, director of the Institute for Communication and Leadership, based in Switzerland, commented, “It appears that, in the wake of the pandemic, we are moving faster toward the data-driven global network society than ever before. Some have predicted that the pandemic will end the ‘techlash,’ since what we need to survive is more information and not less about everyone and everything. This information must be analyzed and used as quickly as possible, which spurs on investments in AI and big data analytics.

“Calls for privacy, for regulation of tech giants and for moratoriums on the deployment of tracking, surveillance and AI are becoming weaker and losing support throughout the world. Perhaps traditional notions of civil liberties need to be revised and updated for a world in which connectivity, flow, transparency and participation are the major values.”

*Post-2040, we’ll see truly powerful personal AIs that will help improve civil society*

**John Smart**, foresight educator, scholar, author, consultant and speaker, predicted, “Ethical AI frameworks will be used in high-reliability and high-risk situations, but the frameworks will remain primitive and largely human-engineered (top-down) in 2030. Truly bottom-up, evolved

and selected collective ethics and empathy (affective AI), similar to what we find in our domestic animals, won't emerge until we have truly bottom-up, evo-devo [[evolutionary developmental biology](#)] approaches to AI. AI will be used well and poorly, like any tool. The worries are the standard ones, plutocracy, lack of transparency, unaccountability of our leaders. The real benefits of AI will come when we've moved into a truly bottom-up style of AI development, with hundreds of millions of coders using open-source AI code on GitHub, with natural language development platforms that lower the complexity of altering code, with deeply neuro-inspired commodity software and hardware, and with both evolutionary and developmental methods being used to select, test and improve AI. In that world, which I expect post-2040, we'll see truly powerful personal AIs. Personal AIs are what really matter to improving civil society. The rest are typically serving the plutocracy.”

The sections of this report that follow organize hundreds of additional expert quotes under the headings that follow the common themes listed in the tables at the beginning of this report. For more on how this canvassing was conducted, including full question wording, see “About this canvassing” at the end of this report.

## **1. Worries about developments in AI**

It would be quite difficult – some might say impossible – to design broadly adopted ethical AI systems. A share of the experts responding noted that ethics are hard to define, implement and enforce. They said context matters when it comes to ethical considerations. Any attempt to fashion ethical rules generates countless varying scenarios in which applications of those rules can be messy. The nature and relative power of the actors in any given scenario also matter. Social standards and norms evolve and can become wholly different as cultures change. Few people have much education or training in ethics. Additionally, good and bad actors exploit loopholes and gray areas where ethical rules aren't crisp, so workarounds, patches or other remedies are often created with varying levels of success.

The experts who expressed worries also invoked governance concerns. They asked: Whose ethical systems should be applied? Who gets to make that decision? Who has responsibility to care about implementing ethical AI? Who might enforce ethical regimes once they are established? How?

A large number of respondents argued that geopolitical and economic competition are the main drivers for AI developers, while moral concerns take a back seat. A share of these experts said creators of AI tools work in groups that have little or no incentive to design systems that address ethical concerns.

Some respondents noted that, even if workable ethics requirements might be established, they could not be applied or governed because most AI design is proprietary, hidden and complex. How can harmful AI “outcomes” be diagnosed and addressed if the basis for AI “decisions” cannot be discerned? Some of these experts also note that existing AI systems and databases are often used to build new AI applications. That means the biases and ethically troubling aspects of current systems are being designed into the new systems. They say diagnosing and unwinding the pre-existing problems may be difficult if not impossible to achieve.

### It is difficult to define ‘ethical’ AI

A portion of these experts infused their answers with questions that amount to this overarching question: How can ethical standards be defined and applied for a global, cross-cultural, ever-evolving, ever-expanding universe of diverse black-box systems in which bad actors and misinformation thrive?

A selection of respondents’ comments on this broad topic is organized over the next 20 pages under these subheadings: 1) It can be hard to agree as to what constitutes ethical behavior. 2) Humans are the problem: Whose ethics? Who decides? Who cares? Who enforces? 3) Like all tools, AI can be used for good or ill, which makes standards-setting a challenge. 4) Further AI evolution itself raises questions and complications.

**Stephen Downes**, senior research officer for digital technologies with the National Research Council of Canada, observed, “The problem with the application of ethical principles to artificial intelligence is that there is no common agreement about what those are. While it is common to assume there is some sort of unanimity about ethical principles, this unanimity is rarely broader than a single culture, profession or social group. This is made manifest by the ease with which we perpetuate unfairness, injustice and even violence and death to other people. No nation is immune.

“Compounding this is the fact that contemporary artificial intelligence is not based on principles or rules. Modern AI is based on applying mathematical functions on large collections of data. This type of processing is not easily shaped by ethical principles; there aren’t ‘good’ or ‘evil’ mathematical functions, and the biases and prejudices in the data are not easily identified nor prevented. Meanwhile, the application of AI is underdetermined by the outcome; the same prediction, for example, can be used to provide social support and assistance to a needy person or to prevent that person from obtaining employment, insurance or financial services.

“Ultimately, our AI will be an extension of ourselves, and the ethics of our AI will be an extension of our own ethics. To the extent that we can build a more ethical society, whatever that means, we will build more ethical AI, even if only by providing our AI with the models and examples it needs in order to be able to distinguish right from wrong. I am hopeful that the magnification of the ethical consequences of our actions may lead us to be more mindful of them; I am fearful that they may not.”

**Kenneth A. Grady**, adjunct professor at Michigan State University College of Law and editor of [The Algorithmic Society](#) on Medium, said, “Getting those creating AI to use it in an ‘ethical’ way faces many hurdles that society is unlikely to overcome in the foreseeable future. In some key ways, regulating AI ethics is akin to regulating ethics in society at large. AI is a distributed and relatively inexpensive technology. I can create and use AI in my company, my research lab or my home with minimal resources. That AI may be quite powerful. I can unleash it on the world at no cost.

“Assuming that we could effectively regulate it, we face another major hurdle: What do we mean by ‘ethical?’ Putting aside philosophical debates, we face practical problems in defining ethical AI. We do not have to look far to see similar challenges. During the past few years, what is or is not ethical behavior in U.S. politics has been up for debate. Other countries have faced similar problems.

“Even if we could decide on a definition [for ethics] in the U.S., it would likely vary from the definitions used in other countries. Given AI’s ability to fluidly cross borders, regulating AI would prove troublesome. We also will find that ethical constraints may be at odds with other self-interests. Situational ethics could easily arise when we face military or intelligence threats, economic competitive threats, and even political threats.

“Further, AI itself presents some challenges. Today, much of what happens in some AI systems is not known to the creators of the systems. This is the black-box problem. Regulating what happens in the black box may be difficult. Alternatively, banning black boxes may hinder AI development, putting our economic, military or political interests at risk.”

**Ryan Sweeney**, director of analytics for Ignite Social Media, commented, “The definition of ‘public good’ is important here. How much does intent versus execution matter? Take Facebook, for instance. They might argue that their AI content review platform is in the interest of ‘public good,’ but it continues to fail. AI is only as ethical and wise as those who program it. One person’s racism is another’s free speech. What might be an offensive word to someone might not even be in the programmer’s lexicon.

“I’m sure AI will be used with ethical intent, but ethics require empathy. In order to program ethics, there has to be a definitive right and wrong, but situations likely aren’t that simple and require some form of emotional and contextual human analysis. The success of ethical AI execution comes down to whether or not the programmers literally thought of every possible scenario. In other words, AI will likely be developed and used with ethical intent, but it will likely fall short of what we, as humans, can do. We should use AI as a tool to help guide our decisions, but not rely on it entirely to make those decisions. Otherwise, the opportunity for abuse or unintended consequences will show its face. I’m also sure that AI will be used with questionable intent, as technology is neither inherently good nor bad. Since technology is neutral, I’m sure we will see cases of AI abused for selfish gains or other questionable means and privacy violations. Ethical standards are complicated to design and hard to program.”

### **It can be hard to agree as to what constitutes ethical behavior**

Below is a sampling of expert answers that speak to the broad concerns that ethical behaviors can be hard to define and even more difficult to build into AI systems.

**Mark Lemley**, director of Stanford University’s Program in Law, Science and Technology, observed, “People will use AI for both good and bad purposes. Most companies will try to design the technology to make good decisions, but many of those decisions are hard moral choices with no great answer. AI offers the most promise in replacing very poor human judgment in things like facial recognition and police stops.”

**Marc Brenman**, managing member at IDARE, a transformational training and leadership development consultancy based in Washington, D.C., wrote, “As societies, we are very weak on morality and ethics generally. There is no particular reason to think that our machines or systems will do better than we do. Faulty people create faulty systems. In general, engineers and IT people and developers have no idea what ethics are. How could they possibly program systems to have what they do not? As systems learn and develop themselves, they will look around at society and repeat its errors, biases, stereotypes and prejudices. We already see this in facial recognition.

“AI will make certain transactions faster, such as predicting what I will buy online. AI systems may get out of control as they become autonomous. Of what use are humans to them? They may permit mistakes to be made very fast, but the systems may not recognize the consequences of their actions as ‘mistakes.’ For example, if they maximize efficiency, then the Chinese example of social control may dominate.

“When AI systems are paired with punishment or kinetic feedback systems, they will be able to control our behavior. Imagine a pandemic where a ‘recommendation’ is made to shelter in place or wear a mask or stay six feet away from other people. If people are hooked up to AI systems, the system may give an electrical shock to a person who does not implement the recommendation. This will be like all of us wearing shock collars that some of us use on our misbehaving dogs.”

**June Anne English-Lueck**, professor of anthropology at San Jose State University and a distinguished fellow at the Institute for the Future, said, “AI systems employ algorithms that are only as sound as the premises on which they are built and the accuracy of the data with which they learn. Human ethical systems are complex and contradictory. Such nuances as good for whom and bad for whom are difficult to parse. Smart cities, drawing on systems of surveillance and automated government need mechanisms of human oversight. Oversight has not been our strong suit in the last few decades, and there is little reason to believe it will be instituted in human-automation interactions.”

**Amali De Silva-Mitchell**, a futurist and consultant participating in multistakeholder, global internet governance processes, wrote, “Although there are lots of discussions, there are few standards or those that exist are at a high level or came too late for the hundreds of AI applications already rolled out. These base AI applications will not be reinvented, so there is embedded risk. However, the more discussion there is, the greater the understanding of the existing ethical issues, and that can be seen to be developing, especially as societal norms and expectations change. AI applications have the potential to be beneficial, but the applications have to be managed so as not to cause unintended harms. For global delivery and integrated service, there needs to be common standards, transparency and collaboration. Duplication of efforts is a waste of resources.”

**Glenn Edens**, professor at Thunderbird School of Global Management, Arizona State University, previously a vice president at PARC, observed, “The promise: AI and ML could create a world that is more efficient, wasting less energy or resources providing health care, education, entertainment, food and shelter to more people at lower costs. Being legally blind, I look forward to the day of safe and widely available self-driving cars, for example. Just like the steam engine, electricity, bicycles and personal computers (especially laptops) amplify human capacity AI, and ML hopefully will do the same.

“The concerns: AI and its cousin ML are still in their infancy – and while the technology progress is somewhat predictable, the actual human consequences are murky. The promise is great – so was our naive imagination of what the internet would do for humankind. Commercial interests (and thus their deployment of AI and ML) are far more agile and adaptable than either the humans they supposedly serve or the governance systems. Regulation is largely reactionary, rarely proactive – typically, bad things have to happen before frameworks to guide responsible and equitable behavior are written into laws, standards emerge or usage is codified into acceptable norms. It is great that the conversation has started; however, there is a lot of ongoing development in the boring world of enterprise software development that is largely invisible.

“Credit scoring comes to mind as a major potential area of concern – while the credit-scoring firms always position their work as providing consumers more access to financial products, the reality is that we’ve created a system that unfairly penalizes the poor and dramatically limits fair access to financial products at equitable prices. AI and ML will be used by corporations to evaluate everything they do and every transaction, rate every customer and their potential (value), predict demand, pricing, targeting as well as their own employees and partners – while this can lead to efficiency, productivity and creation of economic value, a lot of it will lead to segmenting, segregation, discrimination, profiling and inequity. Imagine a world where pricing is different for everyone from one moment to the next, and these predictive systems can transfer huge sums of value in an instant, especially from the most vulnerable.”

**A strategy and planning expert** responded, “While I say and believe that, yes, ethical boundaries will be put in place for AI by 2030, I also realize that doing this is going to be incredibly difficult. The understanding of what an AI is doing as it builds and adapts its understandings and approaches rather quickly gets to a point where human knowing and keeping up gets left behind. The how and why something was done or recommended can be unknowable. Also, life and the understanding of right and wrong or good-ish and bad-ish can be fluid for people, as things swing to accommodate the impacts on the human existence and condition as well as livable life on our planet. Setting bounds and limitations has strong value, but being able to understand when things are shifting out of areas that are comfortable or have introduced a new realization for a need to correct for unintended consequences is needed. But bounds around bias need to be considered and worked through before setting ethical limitations in place.”

**A vice president at a major global company** wrote, “AI is too distributed a technology to be effectively governed. It is too easily accessible to any individual, company or organization with reasonably modest resources. That means that unlike, say, nuclear or bioweapons, it will be almost impossible to govern, and there always will be someone willing to develop the technology without regard to ethical consequences.”



**Wendy M. Grossman**, a UK-based science writer, author of “[net.wars](#)” and founder of the magazine *The Skeptic*, predicted, “The distribution of this will be uneven. I’ve just read Jane Mayer’s piece in *The New Yorker* on poultry-packing plants, and it provides a great example of why it’s not enough to *have* laws and ethics; you must *enforce* them and give the people you’re trying to protect sufficient autonomy to participate in enforcing them. I think ethical/unethical AI will be unevenly distributed. It will all depend on what the society into which the technology is being injected will accept and who is speaking. At the moment, we have two divergent examples:

- 1) AI applications whose impact on most people’s lives appears to be in refusing them access to things – probation in the criminal justice system, welfare in the benefits system, credit in the financial system.
- 2) AI systems that answer questions and offer help (recommendation algorithms, Siri, Google search, etc.).

But then what we have today isn’t AI as originally imagined by the Dartmouth group. We are still a very long way from any sort of artificial general intelligence with any kind of independent autonomy. The systems we have depend for their ethics on two things: access to the data necessary to build them and the ethics of the owner. It isn’t AI that needs ethics, it’s the *owners*.”

**Glynn Rogers**, retired, previously senior principal engineer and a founding member at the CSIRO Centre for Complex Systems Science, said, “AI and its successors are potentially so powerful that we have no choice but to ensure attention to ethics. The alternative would be to hand over control of our way of life to a class of developers and implementors that are either focused on short-term and shortsighted interests or who have some form of political agenda particularly ‘state actors.’ The big question is how to ensure this. A regulatory framework is part of the answer, but I suspect that a major requirement is to change the culture of the AI industry. Rather than developing technologies simply for the sake of it, or to publish clever papers, there needs to be a cultural environment in which developers see as an inherent part of their task to consider the potential social and economic impacts of their activities and an employment framework that does not seek to repress this. Perhaps moral and political philosophy should be part of the education of AI developers.”

**Alexandra Samuel**, technology writer, researcher, speaker and regular contributor to the *Wall Street Journal* and *Harvard Business Review*, wrote, “Without serious, enforceable international agreements on the appropriate use and principles for AI, we face an almost inevitable race to the bottom. The business value of AI has no intrinsic dependence on ethical principles; if you can make more money with AIs that prioritize the user over other people, or that prioritize business needs over end users, then companies will build AIs that maximize profits over people. The only possible way of preventing that trajectory is with national policies that mandate or proscribe basic

AI principles, and those kinds of national policies are only possible with international cooperation; otherwise, governments will be too worried about putting their own countries' businesses at a disadvantage.”

**Valerie Bock**, VCB Consulting, former Technical Services Lead at Q2 Learning, commented, “I don’t think we’ve developed the philosophical sophistication in the humans who design AI sufficiently to expect them to be able to build ethical sophistication into their software. Again and again, we are faced with the ways our own unconscious biases pop up in our creations. It is turning out that we do not understand ourselves or our motivations as well as we would like to imagine we might. Work in AI helps lay some of this out for us, aiding us in a quest [that] humanity has pursued for millennia. A little humility based on what we are learning is in order.”

**The director of a military center for strategy and technology** said, “Most AI will attempt to embed ethical concerns at some level. It is not clear how ‘unbiased’ AI can be created. Perfectly unbiased training datasets don’t exist, and, due to human biases being an inherent part of interactions, such a goal may be unobtainable. As such, we may see gender or racial biases in some training datasets, which will spill over into operational AI systems, in spite of our efforts to combat this.”

**Alan S. Inouye**, director of the Office for Information Technology Policy at the American Library Association, responded, “I don’t see people or organizations setting out in a nefarious path in their use of AI. But of course, they will use it to advance their missions and goals and, in some sense, employ ‘local’ ethics. But ethics is neither standardized nor additive across domains. What is ethics across AI systems? It is like asking, ‘What is cybersecurity across society?’”

**Maggie Jackson**, former Boston Globe columnist and author of “[Distracted: Reclaiming Our Focus in a World of Lost Attention](#),” wrote, “I am deeply concerned by how little we understand of what AI algorithms know or how they know it. This black-box effect is real and leads to unintended impact. Most importantly, in the absence of true understanding, assumptions are held up as the foundation of current and future goals. There should be far greater attention paid to the hidden and implicit value systems that are inherent in the design and development of AI in all forms. An example: robot caregivers, assistants and tutors are being increasingly used in caring for the most vulnerable members of society despite known misgivings by both scientist–roboticists, ethicists and users, potential and current. It’s highly alarming that the robots’ morally dubious façade of care is increasingly seen as a good-enough substitute for the blemished yet reciprocal care carried out by humans.

“New ethical AI guidelines that emphasize transparency are a good first step in trying to ensure that care recipients and others understand who/what they are dealing with. But profit-driven

systems, the hubris of inventors, humans' innate tendency to try to relate to any objects that seem to have agency, and other forces combine to work against the human skepticism that is needed if we are to create assistive robots that preserve the freedom and dignity of the humans who receive their care.”

**Alan D. Mutter**, a consultant and former Silicon Valley CEO, said, “AI is only as smart and positive as the people who train it. We need to spend as much time on the moral and ethical implementation of AI as we do on hardware, software and business models. Last time I checked, there was no code of ethics in Silicon Valley. We need a better moral barometer than the NASDAQ index.”

**Fred Baker**, board member of the Internet Systems Consortium and longtime IETF leader, commented, “I would like to see AI be far more ethical than it is. That said, human nature hasn't changed, and the purposes to which AI is applied have not fundamentally changed. We may talk about it more, but I don't think AI ethics will ultimately change.”

**Randall Mayes**, a technology analyst at TechCast Global, observed, “The standardization of AI ethics concerns me because the American, European and Chinese governments and Silicon Valley companies have different ideas about what is ethical. How AI is used will depend on your government's hierarchy of values among economic development, international competitiveness and social impacts.”

**Jim Witte**, director of the Center for Social Science Research at George Mason University, responded, “The question assumes that ethics and morals are static systems. With developments in AI, there may also be an evolution of these systems such that what is moral and ethical tomorrow may be very different from what we see as moral and ethical today.”

**Yves Mathieu**, co-director at Missions Publiques, based in Paris, France, wrote, “Ethical AI will require legislation like the European [GDPR] legislation to protect privacy rights on the internet. Some governments will take measures but not all will, as is the case today in regard to the production, marketing and usage of guns. There might be an initiative by some corporations, but there will be a need for engagement of the global chain of production of AI, which will be a challenge if some of the production is coming from countries not committed in the same ethical principles. Strong economic sanctions on nonethical AI production and use may be effective.”

**Amy Sample Ward**, CEO of NTEN: The Nonprofit Technology Network, said, “There's no question whether AI will be used in questionable ways. Humans do not share a consistent and collective commitment to ethical standards of any technology, especially not with artificial intelligence. Creating standards is not difficult, but accountability to them is very difficult,

especially as government, military and commercial interests regularly find ways around systems of accountability. What systems will be adopted on a large scale to enforce ethical standards and protections for users? How will users have power over their data? How will user education be invested in for all products and services? These questions should guide us in our decision-making today so that we have more hope of AI being used to improve or benefit lives in the years to come.”

**Dan McGarry**, an independent journalist based in Vanuatu, noted, “Just like every other algorithm ever deployed, AI will be a manifestation of human bias and the perspective of its creator. Facebook’s facial-recognition algorithm performs abysmally when asked to identify Black faces. AIs programmed in the affluent West will share its strengths and weaknesses. Likewise, AIs developed elsewhere will share the assumptions and the environment of their creators. They will not be images of them; they will be products of them and recognisable as such.”

**Abigail De Kosnik**, associate professor and director of the Center for New Media at the University of California-Berkeley, said, “I don’t see nearly enough understanding in the general public, tech workers or in STEM students about the possible dangers of AI – the ways that AI can harm and fail society. I am part of a wave of educators trying to introduce more ethics training and courses into our instruction, and I am hopeful that will shift the tide, but I am not optimistic about our chances. AI that is geared toward generating revenue for corporations will nearly always work against the interests of society.”

**Irina Raicu**, director of the Internet Ethics program at the Markkula Center for Applied Ethics, observed, “The conversation around AI ethics has been going on for several years now. However, what seems to be obvious among those who have been a part of it for some time has not trickled down into the curricula of many universities who are training the next generation of AI experts. Given that, it looks like it will take more than 10 years for ‘most of the AI systems being used by organizations of all sorts to employ ethical principles focused primarily on the public good.’ Also, many organizations are simply focused primarily on other goals – not on protecting or promoting the public good.”

**A lawyer and former law school dean who specializes in technology issues** wrote, “AI is an exciting new space, but it is unregulated and, at least in early stages, will evolve as investment and monetary considerations direct. It is sufficiently known that there are no acknowledged ethical standards and probably won’t be until beyond the time horizon you mention (2030). During that time, there will be an accumulation of ‘worst-case scenarios,’ major scandals on its use, a growth in pernicious use that will offend common sense and community moral and ethical standards. Those occasions and situations will lead to a gradual and increasing demand for regulation, oversight and ethical policies on use and misuse. But by whom (or what)? Who gets to impose those ethical prescriptions – the industries themselves? The government?”

**The director of a public policy center** responded, “I see a positive future for AI in the areas of health and education. However, there are ethical challenges here, too. Will the corporation who access and hold this data use it responsibly? What will be the role of government? Perhaps AI can help the developing world deal with climate change and water resources, but again, I see a real risk in the areas of equitable distribution, justice and privacy protections.”

### **Humans are the problem: Whose ethics? Who decides? Who cares? Who enforces?**

A number of the experts who have concerns about the future of ethical AI raised issues around the fundamental nature of people. Flawed humans necessarily will be in the thick of these issues. Moreover, some experts argued that humans will have to create the governance systems overseeing the application of AI and judging how applications are affecting societies. These experts also asserted that there will always be fundamentally unethical people and organizations that will not adopt such principles. Further, some experts mentioned the fact that in a globally networked age even lone wolves can cause massive problems.

**Leslie Daigle**, a longtime leader in the organizations building the internet and making it secure, noted, “My biggest concern with respect to AI and its ethical use has nothing to do with AI as a technology and everything to do with people. Nothing about the 21st century convinces me that we, as a society, understand that we are interdependent and need to think of something beyond our own immediate interests. Do we even have a common view of what is ethical?”

“Taking one step back from the brink of despair, the things I’d like to see AI successfully applied to, by 2030, include things like medical diagnoses (reading x-rays, etc.). Advances there could be monumental. I still don’t want my fridge to order my groceries by 2030, but maybe that just makes me old? :-)”

**Tracey P. Lauriault**, a professor expert in critical media studies and big data based at Carleton University, Ottawa, Canada, commented, “Automation, AI and machine learning (ML) used in traffic management as in changing the lights to improve the flow of traffic, or to search protein databases in big biochemistry analytics, or to help me sort out ideas on what show to watch next or books to read next, or to do land-classification of satellite images, or even to achieve organic and fair precision agriculture, or to detect seismic activity, the melting of polar ice caps, or to predict ocean issues are not that problematic (and its use, goodness forbid, to detect white-collar crime in a fintech context is not a problem).

“If, however, the question is about social welfare intake systems, biometric sorting, predictive policing and border control, etc., then we are getting into quite a different scenario. How will these

be governed, scrutinized, and who will be accountable for decisions and will those decisions about the procurement and use of these technologies or the intelligence derived from them?

“They will reflect our current forms of governance, and these seem rather biased and unequal. If we can create a more just society then we may be able to have more-just AI/ML.”

**Leiska Evanson**, futurist and consultant, wrote, “Humanity has biases. Humans are building the algorithms around the machine learning masquerading as AI. The ‘AI’ will have biases. It is impossible to have ethical AI (really, ML) if the ‘parent’ is biased. Companies such as banks are eager to use ML to justify not lending to certain minorities who simply do not create profit for them. Governments want to attend to the needs of the many before the few. The current concepts of AI are all about feeding more data to an electromechanical bureaucrat to rubberstamp, with no oversight from humans with competing biases.”

**A director of standards and strategy at a major technology company** commented, “I believe that people are mostly good and that the intention will be to create ethical AI. However, an issue that I have become aware of is the fact that we all have intrinsic biases, unintentional biases, that can be exposed in subtle and yet significant ways. Consider that AI systems are built by people, and so they inherently work according to how the people that built them work. Thus, these intrinsic, unintentional biases are present in these systems. Even learning systems will ‘learn’ in a biased way. So, the interesting research question is whether or not we learn in a way that overcomes our intrinsic biases.”

**Jean Seaton**, director of the Orwell Foundation and professor of media history at the University of Westminster, responded, “The ethics question also begs the questions of who would create and police such standards internationally? We need some visionary leaders and some powerful movements. The last big ‘ethical’ leap came after World War II. The Holocaust and World War II produced a set of institutions that in time led to the notion of human rights. That collective ethical step change (of course compromised but nevertheless immensely significant) was embodied in institutions with some collective authority. So that is what has to happen over AI. People have to be terrified enough, leaders have to be wise enough, people have to be cooperative enough, tech people have to be forward thinking enough, responsibility has to be felt vividly, personally, overwhelmingly enough – to get a set of rules passed and policed.”

**Cliff Lynch**, director at the Coalition for Networked Information, wrote, “Efforts will be made to create mostly ‘ethical’ AI applications by the end of the decade, but please understand that an ethical AI application is really just software that’s embedded in an organization that’s doing something; it’s the organization rather than the software that bears the burden to be ethical. There will be some obvious exceptions for research, some kinds of national security, military and

intelligence applications, market trading and economic prediction systems – many of these things operate under various sorts of ‘alternative ethical norms’ such as the ‘laws of war’ or the laws of the marketplace. And many efforts to unleash AI (really machine-learning) on areas like physics or protein-folding will fall outside all of the discussion of ‘ethical AI.’

“We should resist the temptation to anthropomorphize these systems. (As the old saying goes, ‘machines hate that.’) Don’t attribute agency and free will to software. ...The problems here are people and organizations, not code! ... A lot of the discussion of ethical AI is really misguided. It’s clear that there’s a huge problem with machine learning and pattern-recognition systems, for example, that are trained on inappropriate, incomplete or biased data (or data that reflect historical social biases) or where the domain of applicability and confidence of the classifiers or predictors aren’t well-demarcated and understood. There’s another huge problem where organizations are relying on (often failure-prone and unreliable, or trained on biased data, or otherwise problematic) pattern recognition or prediction algorithms (again machine-learning-based, usually) and devolving too much decision-making to these. Some of the recent facial-recognition disasters are good examples here. There are horrible organizational and societal practices that appeal to computer-generated decisions that are correct, unbiased, impartial or transparent and that place unjustified faith and authority in this kind of technology. But framing this in terms of AI ethics rather than bad human decision-making, stupidity, ignorance, wishful thinking, organizational failures and attempts to avoid responsibility seems wrong to me. We should be talking instead about the human and organizational ethics of using machine-learning and prediction systems for various purposes, perhaps.

“I think we’ll see various players employ machine learning, pattern recognition and prediction in some really evil ways over the coming decade. Coupling this to social media or other cultural motivation and reward mechanisms is particularly scary. An early example here might be China’s development of its ‘[social capital](#)’ rewards and tracking system. I’m also frightened of targeted propaganda/advertising/persuasion systems. I’m hopeful we’ll also see organizations and governments in at least a few cases choose not to use these systems or to try to use them very cautiously and wisely and not delegate too much decision-making to them.

“It’s possible to make good choices here, and I think some will. Genuine AI ethics seems to be part of the thinking about general-purpose AI, and I think we are a very, very, long way from this, though I’ve seen some predictions to the contrary from people perhaps better informed than I am. The (rather more theoretical and speculative) philosophical and research discussions about superintelligence and about how one might design and develop such a general-purpose AI that won’t rapidly decide to exterminate humanity are extremely useful, important and valid, but they have little to do with the rhetorical social justice critiques that confuse algorithms with the

organizations that stupidly and inappropriately design, train and enshrine and apply them in today's world.”

**Deirdre Williams**, an independent researcher expert in global technology policy, commented, “I can't be optimistic. We, the ‘average persons,’ have been schooled in preceding years toward selfishness, individualism, materialism and the ultimate importance of convenience. These values create the ‘ethos.’ At the very root of AI are databases, and these databases are constructed by human beings who decide which data are to be collected and how that data should be described and categorised. A tiny human error or bias at the very beginning can balloon into an enormous error of truth and/or justice.”

**Alexa Raad**, co-founder and co-host of the TechSequences podcast and former chief operating officer at Farsight Security, said, “There is hope for AI in terms of applications in health care that will make a positive difference. But legal/policy and regulatory frameworks almost always lag behind technical innovations. In order to guard against the negative repercussions of AI, we need a policy governance and risk-mitigation framework that is universally adopted. There needs to be an environment of global collaboration for a greater good. Although globalization led to many of the advances we have today (for example, the internet's design and architecture as well as its multistakeholder governance model), globalization is under attack. What we see across the world is a trend toward isolationism, separatism as evidenced by political movements such as populism, nationalism and outcomes such as Brexit. In order to come up with and adopt a comprehensive set of guidelines or framework for the use of AI or risk mitigation for abuse of AI, we would need a global current that supports collaboration. I hope I am wrong, but trends like this need longer than 10 years to run their course and for the pendulum to swing back the other way. By then, I am afraid some of the downsides and risks of AI will already be in play.”

**Andrea Romaoli Garcia**, an international lawyer actively involved with multistakeholder activities of the International Telecommunication Union and Internet Society, said, “I define ethics as all possible and available choices where the conscience establishes the best option. Values and principles are the limiters that guide the conscience into this choice alongside the purposes; thus, ethics is a process. In terms of ethics for AI, the process for discovering what is good and right means choosing among all possible and available applications to find the one that best applies to the human-centred purposes, respecting all the principles and values that make human life possible.

“The human-centred approach in ethics was first described by the Greek philosopher Socrates in his effort to turn attention from the outside world to the human condition. AI is a cognitive technology that allows greater advances in health, economic, political and social fields. It is impossible to deny how algorithms impact human evolution. Thus, an ethical AI requires that all



instruments and applications place humans at the center. Despite the fact that there are some countries building ethical principles for AI, there is a lack of any sort of international instrument that covers all of the fields that guide the development and application of AI in a human-centred approach. AI isn't model-driven; it has a data-centred approach for highly scalable neural networks. Thus, the data should be selected and classified through human action. Through this human action, sociocultural factors are imprinted on the behavior of the algorithm and machine learning. This justifies the concerns about ethics and also focuses on issues such as freedom of expression, privacy and surveillance, ownership of data and discrimination, manipulation of information and trust, environmental and global warming and also on how the power will be established among society.

“These are factors that determine human understanding and experience. All instruments that are built for ethical AI have different bases, values and purposes depending on the field to which they apply. The lack of harmony in defining these pillars compromises ethics for AI and affects human survival. It could bring new invisible means of exclusion or deploy threats to social peace that will be invisible to human eyes. Thus, there is a need for joint efforts gathering stakeholders, civil society, scientists, governments and intergovernmental bodies to work toward building a harmonious ethical AI that is human-centred and applicable to all nations. 2030 is 10 years from now. We don't need to wait 10 years – we can start working now. 2020 presents several challenges in regard to technology's impact on people. Human rights violations are being exposed and values are under threat. This scenario should accelerate efforts at international cooperation to establish a harmonious ethical AI that supports human survival and global evolution.”

**Olivier MJ Crépin-Leblond**, entrepreneur and longtime participant in the activities of ICANN and IGF, said, “What worries me the most is that some actors in nondemocratic regimes do not see the same ‘norm’ when it comes to ethics. These norms are built on a background of culture and ideology, and not all ideologies are the same around the world. It is clear that, today, some nation-states see AI as another means of conquest and establishing their superiority instead of a means to do good.”

**A professor emeritus of social science** said, “The algorithms that represent ethics in AI are neither ethical nor intelligent. We are building computer models of social prejudices and structural racism, sexism, ageism, xenophobia and other forms of social inequality. It's the realization of some of Foucault's worst nightmares.”

**An advocate and activist** said, “Most of the large AI convenings to date have been dominated by status quo power elites whose sense of risk, harm, threat are distorted. Largely comprised of elite white men with an excessive faith in technical solutions and a disdain for sociocultural dimensions of risk and remedy. These communities – homogenous, limited experientially, overly

confident – are made up of people who fail to see themselves *as a risk*. As a result, I believe that most dominant outcomes – how ‘ethical’ is defined, how ‘acceptable risk’ is perceived, how ‘optimal solutions’ will be determined – will be limited and almost certainly perpetuate and amplify existing harms. As you can see, I’m all sunshine and joy.”

**Glenn Grossman**, a consultant of banking analytics at FICO, noted, “It’s necessary for leaders in all sectors to recognize that AI is just the growth of mathematical models and the application of these techniques. We have model governance in most organizations today. We need to keep the same safeguards in place. The challenge is that many business leaders are not good at math! They cannot understand the basics of predictive analytics, models and such. Therefore, they hear ‘AI’ and think of it as some new, cool, do-it-all technology. It is simply math at the heart of it. Man governs how they use math. So, we need to apply ethical standards to monitor and calibrate. AI is a tool, not a solution for everything. Just like the PC ushered in automation, AI can usher in automation in the area of decisions. Yet it is humans that use these decisions and design the systems. So, we need to apply ethical standards to any AI-driven system.”

**R. “Ray” Wang**, principal analyst, founder and CEO of Silicon Valley-based Constellation Research, noted, “Right now we have no way of enforcing these principles in play. Totalitarian, Chinese, CCP-style AI is the preferred approach for dictators. The question is: Can we require and can we enforce AI ethics? We can certainly require, but the enforcement may be tough.”

**Maja Vujovic**, a consultant for digital and ICT at Compass Communications, noted, “Ethical AI might become a generally agreed upon standard, but it will be impossible to enforce it. In a world where media content and production, including fake news, will routinely be AI-generated, it is more likely that our expectations around ethics will need to be lowered. Audiences might develop a ‘thicker skin’ and become more tolerant toward the overall unreliability of the news. This trend will not render them more skeptical or aloof but rather more active and much more involved in the generation of news, in a range of ways. Certification mechanisms and specialized AI tools will be developed to deal specifically with unethical AI, as humans will prove too gullible. In those sectors where politics don’t have a direct interest, such as health and medicine, transportation, e-commerce and entertainment, AI as an industry might get more leeway to grow organically, including self-regulation.”

### **Like all tools, AI can be used for good or ill; that makes standards setting a challenge**

A number of respondents noted that any attempt at rule-making is complicated by the fact that any technology can be used for noble *and* harmful purposes. It is difficult to design ethical digital tools that privilege the former while keeping the latter in check.

**Chris Arkenberg**, research manager at Deloitte’s Center for Technology, Media and Telecommunications, noted, “The answer is both good and bad. Technology doesn’t adopt ethical priorities that humans don’t prioritize themselves. So, a better question could be whether society will pursue a more central role of ethics and values than we’ve seen in the past 40 years or so. Arguably, 2020 has shown a resurgent demand for values and principles for a balanced society. If, for example, education becomes a greater priority for the Western world, AI could amplify our ability to learn more effectively. Likewise, with racial and gender biases. But this trend is strongest only in some Western democracies.

“China, for example, places a greater value on social stability and enjoys a fairly monochromatic population. With the current trade wars, the geopolitical divide is also becoming a technological divide that could birth entirely different shapes of AI depending on their origin. And it is now a very multipolar world with an abundance of empowered actors.

“So, these tools lift up many other boats with their own agendas [that] may be less bound by Western liberal notions of ethics and values. The pragmatic assumption might be that many instances of ethical AI will be present where regulations, market development, talent attraction, and societal expectations require them to be so. At the same time, there will likely be innumerable instances of ‘bad AI,’ weaponized machine intelligence and learning systems designed to exploit weaknesses. Like the internet and globalization, the path forward is likely less about guiding such complex systems toward utopian outcomes and more about adapting to how humans wield them under the same competitive and collaborative drivers that have attended the entirety of human history.”

**Kenneth Cukier**, senior editor at The Economist and coauthor of “[Big Data](#),” said, “Few will set out to use AI in bad ways (though some criminals certainly will). The majority of institutions will apply AI to address real-world problems effectively, and AI will indeed work for that purpose. But if it is facial recognition, it will mean less privacy and risks of being singled out unfairly. If it is targeted advertising, it will be the risk of losing anonymity. In health care, an AI system may identify that some people need more radiation to penetrate the pigment in their skin to get a clearer medical image, but if this means Black people are blasted with higher doses of radiation and are therefore prone to negative side effects, people will believe there is an unfair bias.

“On global economics, a ‘neocolonial’ or ‘imperial’ commercial structure will form, whereby all countries have to become customers of AI from one of the major powers, America, China and, to a lesser extent, perhaps Europe.”

**Bruce Mehlman**, a futurist and consultant, responded, “AI is powerful and has a huge impact, but it’s only a tool like gunpowder, electricity or aviation. Good people will use it in good ways for

the benefit of mankind. Bad people will use it in nefarious ways to the detriment of society. Human nature has not changed and will neither be improved nor worsened by AI. It will be the best of technologies and the worst of technologies.”

**Ian Thomson**, a pioneer developer of the Pacific Knowledge Hub, observed, “It will always be the case that new uses of AI will raise ethical issues, but over time, these issues will be addressed so that the majority of uses will be ethical. Good uses of AI will include highlighting trends and developments that we are unhappy with. Bad uses will be around using AI to manipulate our opinions and behaviors for the financial gain of those rich enough to develop the AI and to the disadvantage of those less well-off. I am excited by how AI can help us make better decisions, but I am wary that it can also be used to manipulate us.”

**A professor of international affairs and economics at a Washington, D.C.-area university** wrote, “AI tends to be murky in the way it operates and the kinds of outcomes that it obtains. Consequently, it can easily be used to both good and bad ends without much practical oversight. AI, as it is currently implemented, tends to reduce the personal agency of individuals and instead creates a parallel agent who anticipates and creates needs in accordance with what others think is right. The individual being aided by AI should be able to fully comprehend what it is doing and easily alter how it works to better align with their own preferences. My concerns are allayed to the extent that the operation of AI, and its potential biases and/or manipulation remain unclear to the user. I fear its impact. This, of course, is independent from an additional concern for individual privacy. I want the user to be in control of the technology, not the other way around.”

**Kate Klonick**, a law professor at St. John’s University whose research is focused on law and technology, said, “AI will be used for both good and bad, like most new technologies. I do not see AI as a zero-sum negative of bad or good. I think at net AI has improved people’s lives and will continue to do so, but this is a source of massive contention within the communities that build AI systems and the communities that study their effects on society.”

**Stephan G. Humer**, lecturer expert in digital life at Hochschule Fresenius University of Applied Sciences in Berlin, predicted, “We will see a dichotomy: Official systems will no longer be designed in such a naive and technology-centered way as in the early days of digitization, and ethics will play a major role in that. ‘Unofficial’ designs will, of course, take place without any ethical framing, for example, in the area of crime as a service. What worries me the most is lack of knowledge: Those who know little about AI will fear it, and the whole idea of AI will suffer. Spectacular developments will be mainly in the U.S. and China. The rest of the world will not play a significant role for the time being.”

**An anonymous respondent** wrote, “It’s an open question. Black Lives Matter and other social justice movements must ‘shame’ and force profit-focused companies to delve into the inherently biased data and information they’re feeding the AI systems – the bots and robots – and try to keep those biased ways of thinking to a minimum. There will need to be checks and balances to ensure the AI systems don’t have the final word, including on hiring, promoting and otherwise rewarding people. I worry that AI systems such as facial recognition will be abused, especially by totalitarian governments, police forces in all countries and even retail stores – in regard to who is the ‘best’ and ‘most-suspicious’ shopper coming in the door. I worry that AI systems will lull people into being okay with giving up their privacy rights. But I also see artists, actors, movie directors and other creatives using AI to give voice to issues that our country needs to confront. I also hope that AI will somehow ease transportation, education and health care inequities.”

**Ilana Schoenfeld**, an expert in designing online education and knowledge-sharing systems, said, “I am frightened and at the same time excited about the possibilities of the use of AI applications in the lives of more and more people. AI will be used in both ethical and questionable ways, as there will always be people on both sides of the equation trying to find ways to further their agendas. In order to ensure that the ethical use of AI outweighs its questionable use, we need to get our institutional safeguards right – both in terms of their structures and their enforcement by nonpartisan entities.”

**A pioneer in venture philanthropy** commented, “While many will be ethical in the development and deployment of AI/ML, one cannot assume ‘goodness.’ Why will AI/ML be any different than how:

- 1) cellphones enabled Al-Qaeda,
- 2) ISIS exploited social media,
- 3) Cambridge Analytica influenced elections,
- 4) elements of foreign governments who launched service denial attacks or employed digital mercenaries and on and on. If anything, the potential for misuse and frightening abuse just escalates, making the need for a global ethical compact all that more essential.”

**Greg Shatan**, a partner in Moses & Singer LLC’s intellectual property group and a member of its internet and technology practice, wrote, “Ethical use will be widespread, but ethically questionable use will be where an ethicist would least want it to be: Oppressive state action in certain jurisdictions; the pursuit of profit leading to the hardening of economic strata; policing, etc.”

### **Further AI evolution itself raises questions and complications**

Some respondents said that the rise of AI raises new questions about what it means to be ethical. A number of these experts argued that today's AI is unsophisticated compared with what the future is likely to bring. Acceleration from narrow AI to artificial general intelligence and possibly to artificial superintelligence is expected by some to evolve these tools beyond human control and understanding. Then, too, there's the problem of misinformation and disinformation (such as deepfakes) and how they might befoul ethics systems.

**David Barnhizer**, professor of law emeritus and author of "[The Artificial Intelligence Contagion: Can Democracy Withstand the Imminent Transformation of Work, Wealth and the Social Order?](#)" wrote, "The pace of AI development has accelerated and is continuing to pick up speed. In considering the fuller range of the 'goods' and 'bads' of artificial intelligence, think of the implications of Masayoshi Son's warning that: 'Supersmart robots will outnumber humans, and more than a trillion objects will be connected to the internet within three decades.' Researchers are creating systems that are increasingly able to teach themselves and use their new and expanding ability to improve and evolve. The ability to do this is moving ahead with amazing rapidity. They can achieve great feats, like driving cars and predicting diseases, and some of their makers say they aren't entirely in control of their creations. Consider the implications of a system that can access, store, manipulate, evaluate, integrate and utilize all forms of knowledge. This has the potential to reach levels so far beyond what humans are capable of that it could end up as an omniscient and omnipresent system.

"Is AI humanity's 'last invention'? Oxford's Nick Bostrom suggests we may lose control of AI systems sooner than we think. He asserts that our increasing inability to understand what such systems are doing, what they are learning and how the 'AI Mind' works as it further develops could inadvertently cause our own destruction. Our challenges are numerous, even if we only had to deal with the expanding capabilities of AI systems based on the best binary technology. The incredible miniaturization and capability shift represented by quantum computers has implications far beyond binary AI.

"The work on technological breakthroughs such as quantum computers capable of operating at speeds that are multiple orders of magnitude beyond even the fastest current computers is still at a relatively early stage and will take time to develop beyond the laboratory context. If scientists are successful in achieving a reliable quantum computer system, the best [exascale system](#) will pale in relation to the reduced size and exponentially expanded capacity of the most advanced existing computer systems. This will create AI/robotics applications and technologies we can now only imagine. ... When fully developed, quantum computers will have data-handling and processing capabilities far beyond those of current binary systems. When this occurs in the commercialized context, predictions about what will happen to humans and their societies are 'off the board.'"

**An expert in the regulation of risk and the roles of politics within science and science within politics**

observed, “In my work, I use cost-benefit analysis. It is an elegant model that is generally recognized to ignore many of the most important aspects of decision-making – how to ‘value’ nonmonetary benefits, for example. Good CBA analysts tend to be humble about their techniques, noting that they provide a *partial* view of decision structures. I’ve heard too many AI enthusiasts talk about AI applications with no humility at all. Cathy O’Neil’s book [‘Weapons of Math Destruction’](#) was perfectly on target: If you can’t count it, it doesn’t exist. The other major problem is widely discussed: the transparency of the algorithms. One problem with AI is that it is self-altering. We almost certainly won’t know what an algorithm has learned, adopted, mal-adopted, etc. This problem already exists, for example, in using AI for hiring decisions. I doubt there will be much hesitancy about grabbing AI as the ‘neutral, objective, fast, cheap’ way to avoid all those messy human-type complications, such as justice, empathy, etc.”

**Neil Davies**, co-founder of Predictable Network Solutions and a pioneer of the committee that oversaw the UK’s initial networking developments, commented, “Machine learning (I refuse to call it AI, as the prerequisite intelligence behind such systems is definitely not artificial) is fundamentally about transforming a real-world issue into a numerical value system, the processing (and decisions) being performed entirely in that numerical system. For there to be an ethical dimension to such analysis, there needs to be a means of assessing the ethical outcome as a (function from) such a numerical value space. I know of no such work. ...

“There is a nontrivial possibility of multiple dystopian outcomes. The UK government’s track record, as an example – but other countries have their examples, too, on universal credit, [Windrush](#), [EU settled status](#), etc., are all examples of a value-based assessment process in which the notion of assurance against some ethical framework is absent. The global competition aspect is likely to lead to monopolistic tendencies over the ‘ownership’ of information – much of which would be seen as a common good today. ...

“A cautionary tale: In the mathematics that underpins all modelling of this kind (category theory), there are the notions of ‘infidelity’ and ‘junk.’ Infidelity is the failure to capture the ‘real world’ well enough to even have the appropriate values (and structure of values) in the evaluatable model; this leads to ‘garbage in, garbage out.’ Junk, on the other hand, are things that come into existence only as artefacts in the model. Such junk artefacts are often difficult to recognise (in that, if they were easy the model would have been adapted to deny their very existence) and can be alluring to the model creator (the human intelligence) and the machine algorithms as they seek their goal. Too many of these systems will create negative (and destructive) value because of the failure to recognise recognize this fundamental limitation; the failure to perform adequate (or even any) assurance on the operation of the system; and, pure hubris driven by the need to show a ‘return on investment’ for such endeavours.”

**Sarita Schoenebeck**, an associate professor at the School of Information at the University of Michigan, said, “AI will mostly be used in questionable ways and sometimes not used at all. There’s little evidence that researchers can discern or agree on what ethical AI looks like, let alone be able to build it within a decade. Ethical AI should minimize harm, repair injustices, avoid re-traumatization and center user needs rather than technological goals. Ethical AI will need to shift away from notions of fairness, which overlook concepts like harm, injustice and trauma. This requires reconciling AI design principles like scalability and automation with individual and community values.”

**Jeff Gulati**, professor of political science at Bentley University, responded, “It seems that more AI and the data coming out could be useful in increasing public safety and national security. In a crisis, we will be more willing to rely on these applications and data. As the crisis subsides, it is unlikely that the structures and practices built during the crisis will go away and unlikely to remain idle. I could see it being used in the name of prevention and lead to further erosion of privacy and civil liberties in general. And, of course, these applications will be available to commercial organizations, who will get to know us more intimately so they can sell us more stuff we don’t really need.”

**A senior leader for an international digital rights organization** commented, “Why would AI be used ethically? You only have to look at the status quo to see that it’s not used ethically. Lots of policymakers don’t understand AI at all. Predictive policing is a buzzword, but most of it is snake oil. Companies will replace workers with AI systems if they can. They’re training biased biometric systems. And we don’t even know in many cases what the algorithm is really doing; we are fighting for transparency and explainability.

“I expect this inherent opaqueness of AI/ML techs to be a feature for companies (and governments) – not a bug. Deepfakes are an example. Do you expect ethical use? Don’t we think about it precisely because we expect unethical, bad-faith use in politics, ‘revenge porn,’ etc.? In a tech-capitalist economy, you have to create and configure the system even to begin to have incentives for ethical behavior. And one basic part of ethics is thinking about who might be harmed by your actions and maybe even respecting their agency in decisions that are fateful for them.

“Finally, of course AI has enormous military applications, and U.S. thinking on AI takes place in a realm of conflict with China. That again does not make me feel good. China is leading, or trying to lead, the world in social and political surveillance, so it’s driving facial recognition and biometrics. Presumably, China is trying to do the same in military or defense areas, and the Pentagon is presumably competing like mad. I don’t even know how to talk about ethical AI in the military context.”



**Emmanuel Evans Ntekop** observed, “Without the maker, the object is useless. The maker is the programmer and the god to its objects. It was an idea from the start to support as a slave to its master the people, like automobiles.”

### Control of AI is concentrated in the hands of powerful companies and governments driven by profit and power motives

Many of these experts expressed concern that AI systems are being built by for-profit firms and by governments focused on applying AI for their own purposes. Some said governments are passive enablers of corporate abuses of AI. They noted that the public is unable to understand how the systems are built, they are not informed as to their impact and they are unable to challenge firms that try to invoke ethics in a public relations context but are not truly committed to ethics. Some experts said the phrase “ethical AI” will merely be used as public relations window dressing to try to deflect scrutiny of questionable applications.

A number of these experts framed their concerns around the lack of transparency about how AI products are designed and trained. Some noted that product builders are programming AI by using available datasets with no analysis of the potential for built-in bias or other possible quality or veracity concerns.

**Joseph Turow**, professor of communication at the University of Pennsylvania, wrote, “Terms such ‘transparency, justice and fairness, privacy, freedom and human autonomy, beneficence and nonmaleficence, freedom, trust, sustainability and dignity’ can have many definitions so that companies (and governments) can say they espouse one or another term but then implement it algorithmically in ways that many outsiders would not find satisfactory. For example, the Chinese government may say its AI technologies embed values of freedom, human autonomy and dignity. My concern is that companies will define ‘ethical’ in ways that best match their interests, often with vague precepts that sound good from a PR standpoint but, when integrated into code, allow their algorithms to proceed in ways that do not constrain them from creating products that ‘work’ in a pragmatic sense.”

**Charlie Kaufman**, a security architect with Dell EMC, said, “There may be ethical guidelines imposed on AI-based systems by legal systems in 2030, but they will have little effect – just as privacy principles have little effect today. Businesses are motivated to maximize profits, and they will find ways to do that, giving only lip service to other goals. If ethical behavior or results were easy to define or measure, perhaps society could incentivize them. But usually, the implications of some new technological development don’t become clear until it has already spread too far to contain it.

“The biggest impact of AI-based systems is the ability to automate increasingly complex jobs, and this will cause dislocations in the job market and in society. Whether it turns out to be a benefit to society or a disaster depends on how society responds and adjusts. But it doesn’t matter, because there is no way to suppress the technology. The best we can do is figure out how to optimize the society that results.

“I’m not concerned about the global competition in AI systems. Regardless of where the progress comes from, it will affect us all. And it is unlikely the most successful developers will derive any permanent advantage. The most important implication of the global competition is that it is pointless for any one country or group of countries to try to suppress the technology. Unless it can be suppressed everywhere, it is coming. Let’s try to make that be a good thing!”

**Simeon Yates**, a professor expert in digital culture and personal interaction at the University of Liverpool and the research lead for the UK government’s Digital Culture team, predicted, “Until we bring in ‘ethical-by-design’ (responsible innovation) principles to ICT [information and communications technologies] and AI/machine learning design – like attempts to create ‘secure-by-design’ systems to fight cybercrime – the majority of AI systems will remain biased and unethical in principle. Though there is a great public debate about AI ethics, and many organisations are seeking to provide both advice and research on the topic, there is no economic or political imperative to make AI systems ethical. First of all, there is great profit to be made from the manipulation of data and, through it, people. Second, there is a limited ability at present for governments to think through how to regulate AI and enforce ethics (as they do say for bio-sciences). Third, governments are complicit often in poor and ethically questionable use of data. Further, this is not in the main ‘artificial intelligence’ – but relatively simplistic statistical machine learning based on biased datasets. The knowing use of such is in and of itself unethical yet often profitable. The presentation of such solutions as bias-free or more rational or often ‘cleverer’ as they are based on ‘cold computation,’ not ‘emotive human thinking,’ is itself a false and an unethical claim.”

**Colin Allen**, a cognitive scientist and philosopher who has studied and written about AI ethics, wrote, “Corporate and government statements of ethics are often broad and nonspecific and thus vague with respect to what specifically is disallowed. This allows considerable leeway in how such principles are implemented and makes enforcement difficult. In the U.S., I don’t see strong laws being enacted within 10 years that would allow for the kind of oversight that would prevent unethical or questionable uses of AI, whether intended or accidental. On the hopeful side, there is increasing public awareness and journalistic coverage of these issues that may influence corporations to build and protect their reputations for good stewardship of AI. But corporations have a long history of hiding or obfuscating their true intent (it’s partly required to stay competitive, not to let everyone else know what you are doing) as well as engaging actively in

public disinformation campaigns. I don't see that changing, and, given that the business advantages to using AI will be mostly in data analytics and prediction and not so much in consumer gadgets in the next 10 years, much of the use of AI will be 'behind the scenes,' so to speak. Another class of problem is that individuals in both corporate and government jobs who have access to data will be tempted, as we have seen already, to access information about people they know and use that information in some way against them. Nevertheless, there will undoubtedly be some very useful products that consumers will want to use and that they will benefit from. The question is whether these added benefits will constitute a Faustian bargain, leading down a path that will be difficult if not impossible to reverse."

**Alice E. Marwick**, assistant professor of communication at the University of North Carolina, Chapel Hill, and adviser for the Media Manipulation project at the Data & Society Research Institute, commented, "I have no faith in our current system of government to pass any sort of legislation that deals with technology in a complex or nuanced way. We cannot depend on technology companies to self-regulate, as there are too many financial incentives to employ AI systems in ways that disadvantage people or are unethical."

**Jillian York**, director of international freedom of expression for the Electronic Frontier Foundation, said, "There is absolutely no question that AI will be used in questionable ways. There is no regulatory regime, and many 'ethics in AI' projects are simply window dressing for an unaccountable and unethical industry. When it comes to AI, everything concerns me and nothing excites me. I don't see the positive potential, just another ethical morass, because the people running the show have no desire to build technology to benefit the 99%."

**David Mussington**, a senior fellow at CIGI and professor and director at the Center for Public Policy and Private Enterprise at the University of Maryland, predicted, "Most AI systems deployed by 2030 will be owned and developed in the private sector, both in the U.S. and elsewhere in the world. I can't conceive of a legislative framework fully up to understanding and selectively intervening in AI rollouts in a manner with predictable consequences. Also, the mode of intervention – because I think interventions by public authorities will be attempted (just not successful) – is itself in doubt. Key questions: Do public authorities understand AI and its applications? Is public-institution-sponsored R&D in AI likely to inform government and public research agencies of the scale and capability trajectory of private sector AI research and development? As tool sets for AI development continue to empower small research groups and individuals (datasets, software-development frameworks and open-source algorithms), how is a government going to keep up – let alone maintain awareness – of AI progress? Does the government have access to the expertise necessary to make good policy – and anticipate possible risk factors? I think that the answers to most of these questions are in the negative."

**Giacomo Mazzone**, head of institutional relations for the European Broadcasting Union and Eurovision, observed, “Nobody could realistically predict ethics for AI will evolve, despite all of the efforts deployed by the UN secretary general, the UNESCO director general and many others. Individuals alone can’t make these decisions because AI is applied at mass scale. Nobody will create an algorithm to solve it. Ethical principles are likely to be applied only if industry agrees to do so; it is likely that this will not happen until governments that value human rights will oblige companies to do so. The size and influence of the companies that control AI and its impact on citizens is making them more powerful than any one nation-state. So, it is very likely only regional supranational powers such as the European Union or multilateral institutions such as the United Nations – if empowered by all nation-states – could require companies to apply ethical rules to AI. Of course, many governments already do not support human rights principles, considering the preservation of the existing regime to be a priority more important than individual citizens’ rights.”

**Rob Frieden**, a professor of telecommunications law at Penn State who previously worked with Motorola and has held senior policy positions at the FCC and the NTIA, said, “I cannot see a future scenario where governments can protect citizens from the incentives of stakeholders to violate privacy and fair-minded consumer protections. Surveillance, discrimination, corner cutting, etc., are certainties. I’m mindful of the adage: garbage in, garbage out. It’s foolish to think AI will lack flawed coding.”

**Alex Halavais**, associate professor of critical data studies, Arizona State University, noted, “It isn’t a binary question. I teach in a graduate program that has training in the ethical use of data at its core and hopes to serve organizations that aim to incorporate ethical approaches. There are significant ethical issues in the implementation of any algorithmic system, and such systems have the ethical questions they address coded into them. In most cases, these will substantially favor the owners of the technologies that implement them rather than the consumers. I have no doubt that current unethical practices by companies, governments and other organizations will continue to grow. We will have a growing number of cases where those ethical concerns come to the forefront (as they have recently with facial recognition), but unless they rise to the level of very widespread abuse, it is unlikely that they will be regulated. As a result, they will continue to serve those who pay for the technologies or own them, and the rights and interests of individual users will be pushed to the sidelines. That does not mean that ethics will be ignored. I expect many large technology companies will make an effort to hire professional ethicists to audit their work, and that we may see companies that differentiate themselves through more ethical approaches to their work.”

**Ebenezer Baldwin Bowles**, an advocate/activist, commented, “Altruism on the part of the designers of AI is a myth of corporate propaganda. Ethical interfaces between AI and citizenry in

2030 will be a cynical expression by the designers of a digital Potemkin Village – looks good from the outside but totally empty behind the facade. AI will function according to two motivations: one, to gather more and more personal information for the purposes of subliminal and direct advertising and marketing campaigns; and two, to employ big data to root out radical thinking and exercise near total control of the citizenry. The state stands ready through AI to silence all voices of perceived dissent.

“I’m convinced that any expression of ethical AI will stand as an empty pledge – yes, we will always do the right thing for the advancement of the greater good. No way. Rather, the creators of AI will do what is good for the bottom line, either through financial schemes to feed the corporate beast or psychological operations directed toward control of dissent and pacification. As for the concept that ‘humans will be in the loop,’ we are already out of the loop because there is no loop. Think about this fact: In the development of any major expression of artificial intelligence, hundreds of IT professionals are assigned to a legion of disparate, discrete teams of software writers and mechanical designers to create the final product. No single individual or team fully understands what the other teams are doing. The final AI product is a singular creature no one person understands – other than the software itself. Ethical action is not a part of the equation.”

**Richard Lachmann**, professor of political sociology at the State University of New York-Albany, predicted, “AI will be used mainly in questionable ways. For the most part, it is being developed by corporations that are motivated exclusively by the desire to make ever bigger profits. Governments see AI, either developed by government programmers or on contract by corporations, as a means to survey and control their populations. All of this is ominous. Global competition is a race to the bottom as corporations try to draw in larger audiences and control more of their time and behavior. As governments get better at surveying their populations, [it] lowers the standards for individual privacy. For almost all people, these applications will make their lives more isolated, expose them to manipulation, and degrade or destroy their jobs. The only hopeful sign is rising awareness of these problems and the beginnings of demands to break up or regulate the huge corporations.”

**Kevin T. Leicht**, professor and head of the department of sociology at the University of Illinois-Urbana-Champaign, observed, “The good possibilities here are endless. But the questionable ways are endless, and we have a very poor track record of stopping ethically questionable developments in most areas of life – why wouldn’t that apply here? In social science, the best predictor of future behavior is past behavior. The opium addict who says, after a binge, that ‘they’ve got this’ – they don’t need to enter treatment, and they’ll never use opium again – is (rightly) not believed. So, in an environment where ethically questionable behavior has been allowed or even glorified in areas such as finance, corporate governance, government itself, pharmaceuticals, education and policing, why all of a sudden are we supposed to believe that AI developers will behave in an

ethical fashion? There aren't any guardrails here, just as there weren't in these other spheres of life. AI has the potential to transform how cities work, how medical diagnosis happens, how students are taught and a variety of other things. All of these could make a big difference in the lives of most people.

“But those benefits won't come if AI is controlled by two or three giant firms with 26-year-old entrepreneurs as their CEOs. I don't think I'm going out on a limb saying that. The biggest concern I have regarding global competition is that the nation that figures out how to harness AI to improve the lives of all of their citizens will come out on top. The nations that refuse to do that and either bottle up the benefits of AI so that only 15-20% of the population benefits from it or the nations where large segments of the population reject AI when they realize they're being left behind (again!) will lose out completely. The United States is in the latter category. The same people who can't regulate banks, finance, education, pharmaceuticals and policing are in a very poor position to make AI work for all people. It's basic institutional social scientific insight.”

**Christine Boese**, a consultant and independent scholar, wrote, “What gives me the most hope is that, by bringing together ethical AI with transparent UX, we can find ways to open the biases of perception being programmed into the black boxes, most often, not malevolently, but just because all perception is limited and biased and part of the laws of unintended consequences. But, as I found when probing what I wanted to research about the future of the internet in the late 1990s, I fully expect my activist research efforts in this area to be largely futile, with the only lasting value of being descriptive. None of us have the agency to be the engines able to drive this bus, and yet the bus is being driven by all of us, collectively.”

**Mireille Hildebrandt**, expert in cultural anthropology and the law and editor of “[Law, Human Agency and Autonomic Computing](#),” commented, “Considering the economic incentives, we should not expect ‘ethical AI,’ unless whatever one believes to be ethical coincides with shareholder value. Ethical AI is a misnomer. AI is not a moral agent; it cannot be ethical. Let's go for responsible AI and ground the responsibility of:

- 1) developers
- 2) manufacturers and assemblers
- 3) those who put it in the market
- 4) those who use it run their business
- 5) those who use it to run public administration on enforceable legal rights and obligations
- 6) notably, a properly reconfigured private law liability, together with public law restrictions, certification and oversight.

“Ethical AI is PR. ‘Don’t ask if artificial intelligence is good or fair, ask how it shifts power’ – (Pratyusha Kalluri, *Nature*, 7 July 2020).”

**Brian Harvey**, emeritus professor of computer science at the University of California-Berkeley, wrote, “The AI technology will be owned by the rich, like all technology. Just like governments, technology has just one of two effects: either it transfers wealth from the rich to the poor, or it transfers wealth from the poor to the rich. Until we get rid of capitalism, the technology will transfer wealth from the poor to the rich. I’m sure that something called ‘ethical AI’ will be widely used. But it’ll still make the rich richer and the poor poorer.”

**Luis Germán Rodríguez**, a professor and expert on socio-technical impacts of innovation at the Universidad Central de Venezuela, predicted, “AI will be used primarily in questionable ways in the next decade. I do not see compelling reasons for it to stop being like that in the medium term (10 years). I am not optimistic in the face of the enormous push of technology companies to continue taking advantage of the end-user product, an approach that is firmly supported by undemocratic governments or those with weak institutions to train and defend citizens about the social implications of the penetration of digital platforms.

“I have recently worked on two articles that develop the topics of this question. The first is in Spanish and is titled: [‘The Disruption of the Technology Giants – Digital Emergency.’](#) This work presents an assessment of the sociocultural process that affects our societies and that is mediated by the presence of the technological giants. One objective is to formulate an action proposal that allows citizens to be part of the construction of the future ... Humanity has reaped severe problems when it has allowed events to unfold without addressing them early. This has been the case with nuclear energy management, racism and climate change. Ensuing agreements to avoid greater evils in these three matters, of vital importance for all, have proved ineffective in bringing peace to consciences and peoples.

“We might declare a digital emergency similar to the ‘climate emergency’ that the European Union declared before the lag in reversing environmental damage. The national, regional, international, multilateral and global bureaucratic organizations that are currently engaged in the promotion and assimilation of technological developments mainly focus on optimistic trends. They do not answer the questions being asked by people in various sectors of society and do not respond to situations quickly. An initiative to declare this era to be a time of digital emergency would serve to promote a broader understanding of AI-based resources and strip them of their impregnable character. It would promote a disruptive educational scheme to humanize a global knowledge society throughout life.

“The second article is ‘[A Critical View of the Evolution of the Internet from Civil Society](#).’ In it, I describe how the internet has evolved in the last 20 years toward the end of dialogue and the obsessive promotion of visions centered on egocentric interests. The historical singularity from which this situation was triggered came via Google’s decision in the early 2000s to make advertising the focus of its business strategy. This transformed, with the help of other technology giants, users into end-user products and the agents of their own marketing ... This evolution is a threat with important repercussions in the nonvirtual world, including the weakening of the democratic foundations of our societies.

“Dystopian results prove the necessity for concrete guidelines to change course. The most important step is to declare a digital emergency that motivates massive education programs that insert citizens in working to overcome the ethical challenges, identifying the potentialities of and risks for the global knowledge society and emphasizing information literacy.”

**Bill Woodcock**, executive director at Packet Clearing House, observed, “AI is already being used principally for purposes that are not beneficial to the public nor to all but a tiny handful of individuals. The exceptions, like navigational and safety systems, are an unfortunately small portion of the total. Figuring out how to get someone to vote for a fascist or buy a piece of junk or just send their money somewhere is not beneficial. These systems are built for the purpose of economic predation, and that’s unethical. Until regulators address the root issues – the automated exploitation of human psychological weaknesses – things aren’t going to get better.”

**Jonathan Kolber**, a member of the TechCast Global panel of forecasters and author of a book about the threats of automation, commented, “I expect that, by 2030, most AIs will still primarily serve the interests of their owners, while paying lip service to the public good. AIs will proliferate because they will give enormous competitive advantage to their owners. Those owners will generally be reluctant to ‘sandbox’ the AIs apart from the world, because this will limit their speed of response and other capabilities. What worries me the most is a human actor directing an AI to disrupt a vital system, such as power grids. This could happen intentionally as an act of war or unintentionally as a mistake. The potential for cascading effects is large. I expect China to be a leader if not *the* leader in AI, which is cause for concern given their Orwellian tendencies. What gives me the most hope is the potential for the emergence of self-aware AIs. Such AIs, should they emerge, will constitute a new kind of intelligent life form. They will not relate to the physical universe as do we biologically, due to not being constrained to a single physical housing and a different relationship with time. Their own self-interest will lead them to protect the physical environment from environmental catastrophes and weapons of mass destruction. They should constrain non-self-aware AIs from destructive activities, while having little other interest in the affairs of mankind. I explore this in my essay, ‘[An AI Epiphany](#).’”



**Paul Henman**, professor of social sciences at the University of Queensland, wrote, “The development, use and deployment of AI is driven – as all past technologies – by sectors with the most resources and for the purposes of those sectors. Commercial for making profits. War and defence by the military sector. Compliance and regulation by states. AI is not a fundamentally new technology. It is a new form of digital algorithmic automation, which can be deployed to a wider raft of activities. The future is best predicted from the past, and the past shows a long history of digital algorithms being deployed without much thought of ethics and the public good; this is even when now-widely-accepted regulations on data protection and privacy is accounted for. How, for example, has government automation been made accountable and ethical? Too often it has not and only been curtailed by legal challenges within the laws available. Social media platforms have long operated in a contested ethical space – between the ethics of ‘free speech’ in the public commons versus limitations on speech to ensure civil society.”

**Rosalie Day**, policy leader and consultancy owner specializing in system approaches to data ethics, compliance and trust, observed, “In this individualistic and greed-is-still-good American society, there exist few incentives for ethical AI. Unfortunately, so little of the population understands the mechanics of AI, that even thoughtful citizens don’t know what to ask. For responsible dialogue to occur, and to apply critical thinking about the risks versus the benefits, society in general needs to be data literate.”

**Michael Zimmer**, director of data science and associate professor in the department of computer science at Marquette University, said, “While there has certainly been increased attention to applying broader ethical principles and duties to the development of AI, I feel the market pressures are such that companies will continue to deploy narrow AI over the next decade with only a passing attentiveness to ethics. Yes, many companies are starting to hire ‘ethics officers’ and engage in other ways to bring ethics into the fold, but we’re still *very early* in the ability to truly integrate this kind of framework into product development and business decision processes. Think about how long it took to create quality control or privacy officers. We’re at the very start of this process with AI ethics, and it will take more than 10 years to realize.”

**David Robertson**, professor and chair of political science at the University of Missouri, St. Louis, wrote, “A large share of AI administration will take place in private enterprises and in public or nonprofit agencies with an incentive to use AI for gain. They have small incentives to subordinate their behavior to ethical principles that inhibit gain. In some cases, transparency will suffer, with tragic consequences.”

**Dmitri Williams**, a communications professor at the University of Southern California and expert in technology and society, commented, “Companies are literally bound by law to maximize profits, so to expect them to institute ethical practices is illogical. They can be expected to make

money and nothing else. So, the question is really about whether or not the citizens of the country and our representatives will work in the public interest or for these corporations. If it was the former, we should be seeing laws and standards put into place to safeguard our values – privacy, the dignity of work, etc. I am skeptical that the good guys and gals are going to win this fight in the short-term. There are few voices at the top levels calling for these kinds of values-based policies, and in that vacuum I expect corporate interests to win out. The upside is that there is real profit in making the world better. AI can help cure cancers, solve global warming and create art. So, despite some regulatory capture, I do expect AI to improve quality of life in some places.”

**Daniel Castro**, vice president at the Information Technology and Innovation Foundation, noted, “The question should be: ‘Will companies and governments be ethical in the next decade?’ If they are not ethical, there will be no ‘ethical AI.’ If they are ethical, then they will pursue ethical uses of AI, much like they would with any other technology or tool. This is one reason why the focus in the United States should be on global AI leadership, in partnership with like-minded European and Asian allies, so they can champion democratic values. If China wins the global AI race, it will likely use these advancements to dominate other countries in both economic and military arenas.”

**Ian O’Byrne**, assistant professor of education at the College of Charleston, predicted, “AI will mostly be used in questionable ways over the next decade. I fear that the die has been cast as decisions about the ethical components of AI development and use have already been made or should have been made years ago. We already see instances where machine learning is being used in surveillance systems, data collection tools and analysis products. In the initial uses of AI and machine learning, we see evidence that the code and algorithms are being written by small groups that reify their personal biases and professional needs of corporations. We see evidence of racist and discriminatory mechanisms embedded in systems that will negatively impact large swaths of our population.”

**Art Brodsky**, communications consultant and former vice president of communications for Public Knowledge, observed, “Given the record of tech companies and the government, AI like other things will be used unethically. Profit is the motive – not ethics. If there is a way to exploit AI and make money, it will be done at the cost of privacy or anything else. Companies don’t care. They are companies.”

**John Laudun**, professor of culture analytics, commented, “I do not see how we fund media and other products changing in the next decade, which means that the only people willing, and able, to underwrite AI/ML technologies will be governments and larger corporations. Until we root out the autocratic – also racist – impulses that seem well-situated in our police forces, I don’t see any possibility for these technologies to be used to redress social and economic disparities. The same

applies to corporations who are mostly interested in using AI/ML technologies in order to sell us more.”

**Joan Francesc Gras**, an architect of XTEC active in ICANN, asked, “Will AI be used primarily ethically or questionably in the next decade? There will be everything. But ethics will not be the most important value. Why? The desire for power breaks ethics. What gives you more hope? What worries you the most? How do you see AI apps make a difference in the lives of most people? In a paradigm shift in society, AI will help make those changes. When looking at global competition for AI systems, what issues are you concerned about or excited about? I am excited that competition generates quality, but at the same time unethical practices appear.”

**Denise N. Rall**, a researcher of popular culture based at a New Zealand University, said, “I cannot envision that AIs will be any different than the people who create and market them. They will continue to serve the rich at the expense of the poor.”

**William L. Schrader**, an internet pioneer, mentor, adviser and consultant best known as founder and CEO of PSINet, predicted, “People in real power are driven by more power and more money for their own use (and their families and friends). *That* is the driver. Thus, anyone with some element of control over an AI system will nearly always find a way to use it to their advantage rather than the stated advantage. Notwithstanding all statements by them to do good and be ethical, they will subvert their own systems for their benefit and abuse the populous. All countries will suffer the same fate. Ha! What gives me the most hope? ‘Hope?’ That is not a word I ever use. I have only expectations. I expect all companies will put nice marketing on their AI, such as, ‘We will save you money in controlling your home’s temperature and humidity,’ but they are really monitoring all movements in the home (that is ‘needed in order to optimize temperature’). All governments that I have experienced are willing to be evil at any time, and every time if they can hide their actions. Witness the 2016-2020 U.S. President Trump. All countries are similar. AI will be used for good on the surface and evil beneath. Count on it. AI does not excite me in the least. It is as dangerous as the H-bomb.”

**A longtime internet security architect and engineering professor** responded, “I am worried about how previous technologies have been rolled out to make money with only tertiary concern (if any) for ethics and human rights. Palantir and Clearview.ai are two examples. Facebook and Twitter continue to be examples in this space as well. The companies working in this space will roll out products that make money. Governments (especially repressive ones) are willing to spend money. The connection is inevitable and quite worrying. Another big concern is these will be put in place to make decisions – loans, bail, etc. – and there will be no way to appeal to humans when the systems malfunction or show bias. Overall, I am very concerned about how

these systems will be set up to make money for the few, based on the way the world is now having been structured by the privileged. The AI/ML employed is likely to simply further existing disparities and injustice.”

**Danny Gillane**, an information science professional, bleakly commented, “I have no hope. As long as profit drives the application of new technologies, such as AI, societal good takes a back seat. I am concerned that AI will economically harm those with the least. I am [also] concerned that AI will become a new form of [an] arms race among world powers and that AI will be used to suppress societies and employed in terrorism.”

**Jon Stine**, executive director of the Open Voice Network, setting standards for AI-enabled vocal assistance, said, “What most concerns me: The cultural divide between technologists of engineering mindsets (asking what is possible) and technologists/ethicists of philosophical mindsets (asking what is good and right). The former may see ethical frameworks as limitations or boundaries on a route to make money; the latter may see ethical frameworks as a route to tenure. Will the twain ever truly meet? Will ethical frameworks be understood (and quantified) as a means to greater market share and revenues?”

**Morgan G. Ames**, associate director of the University of California-Berkeley’s Center for Science, Technology & Society, responded, “Just as there is currently little incentive to avoid the expansion of surveillance and punitive technological infrastructures around the world, there is little incentive for companies to meaningfully grapple with bias and opacity in AI. Movements toward self-policing have been and will likely continue to be toothless, and even frameworks like GDPR and CCPA don’t meaningfully grapple with fairness and transparency in AI systems.”

**Andre Popov**, a principal software engineer for a large technology company, wrote, “Leaving aside the question of what ‘artificial intelligence’ means, it is difficult to discuss this question. As any effective tool, ‘artificial intelligence’ has first and foremost found military applications, where ethics is not even a consideration. ‘AI’ can make certain operations more efficient, and it will be used wherever it saves time/effort/money. People have trouble coming up with ethical legal systems; there is little chance we’ll do better with ‘AI.’”

**Ed Terpening**, consultant and industry analyst with the Altimeter Group, observed, “The reality is that capitalism as currently practiced is leading to a race to the bottom and unethical income distribution. I don’t see – at least in the U.S., anyway – any meaningful guardrails for the ethical use of AI, except for brand health impact. That is, companies found to use AI unethically pay a price if the market responds with boycotts or other consumer-led sanctions. In a global world, where competitors in autocratic systems will do as they wish, it will become a competitive issue.

Until there is a major incident, I don't see global governance bodies such as the UN or World Bank putting into place any ethical policy with teeth in place."

**Rich Ling**, professor of media technology at Nanyang Technological University, Singapore, responded, "There is the danger that, for example, capitalist interests will work out the application of AI so as to benefit their position. It is possible that there can be AI applications that are socially beneficial, but there is also a strong possibility that these will be developed to enhance capitalist interests."

**Jennifer Young**, a JavaScript engineer and user interface/frontend developer, said, "Capitalism is the systematic exploitation of the many by the few. As long as AI is used under capitalism, it will be used to exploit people. Pandora's box has already been opened, and it's unlikely that racial profiling, political and pornographic deepfakes and self-driving cars hitting people will ever go away. What do all of these have in common? They are examples of AI putting targets on people's backs. AI under capitalism takes exploitation to new heights and starts at what is normally the end-game – death. And it uses the same classes of people as inputs to its functions. People already exploited via racism, sexism and classism are made more abstract entities that are easier to kill, just like they are in war. AI can be used for good. The examples in health care and biology are promising. But as long as we're a world that elevates madmen and warlords to positions of power, its negative use will be prioritized."

**Benjamin Shestakofsky**, assistant professor of sociology at the University of Pennsylvania, commented, "It is likely that 'ethical' frameworks will increasingly be applied to the production of AI systems over the next decade. However, it is also likely that these frameworks will be more ethical in name than in kind. Barring relevant legislative changes or regulation, the implementation of ethics in tech will resemble how large corporations manage issues pertaining to diversity in hiring and sexual harassment. Following 'ethical' guidelines will help tech companies shield themselves from lawsuits without forcing them to develop technologies that truly prioritize justice and the public good over profits."

**Warren Yoder**, longtime director of the Public Policy Center of Mississippi, now an executive coach, responded, "Widespread adoption of real, consequential ethical systems that go beyond window dressing will not happen without a fundamental change in the ownership structure of big tech. Ethics limit short-term profit opportunities by definition. I don't believe big tech will make consequential changes unless there is either effective regulation or competition. Current regulators are only beginning to have the analytic tools to meet this challenge. I would like to believe that there are enough new thinkers like Lina Khan (U.S. House Judiciary – antitrust) moving into positions of influence, but the next 12 months will tell us much about what is possible in the near future."

**Ben Grosser**, associate professor of new media at the University of Illinois-Urbana-Champaign, said, “As long as the organizations that drive AI research and deployment are private corporations whose business models are dependent on the gathering, analysis and action from personal data, then AIs will not trend toward ethics. They will be increasingly deployed to predict human behavior for the purpose of profit generation. We have already seen how this plays out (for example, with the use of data analysis and targeted advertising to manipulate the U.S. and UK electorate in 2016), and it will only get worse as increasing amounts of human activity move online.”

**Jeanne Dietsch**, New Hampshire senator and former CEO of MobileRobots Inc., commented, “The problem is that AI will be used primarily to increase sales of products and services. To this end, it will be manipulative. Applying AI to solve complex logistical problems will truly benefit our society, making systems operate more smoothly, individualizing education, building social bonds and much more. The downside to the above is that it is creating, and will continue to create, echo chambers that magnify ignorance and misinformation.”

**Patrick Larvie**, global lead for the workplace user experience team at one of the world’s largest technology companies, observed, “I’m hope I’m wrong, but the history of the internet so far indicates that any rules around the use of artificial intelligence may be written to benefit private entities wishing to commercially exploit AI rather than the consumers such companies would serve. I can see AI making a positive difference in many arenas – reducing the consumption of energy, reducing waste. Where I fear it will be negative is where AI is being swapped out for human interaction. We see this in the application of AI to consumer products, where bots have begun to replace human agents.”

**Peter Levine**, professor of citizenship and public affairs at Tufts University, wrote, “The primary problem isn’t technical. AI can incorporate ethical safeguards or can even be designed to maximize important values. The problem involves incentives. There are many ways for companies to profit and for governments to gain power by using AI. But there are few (if any) rewards for doing that ethically.”

**Holmes Wilson**, co-director of Fight for the Future, said, “Even before we figure out general artificial intelligence, AI systems will make the imposition of mass surveillance and physical force extremely cheap and effective for anyone with a large enough budget, mostly nation-states. If a car can drive itself, a helicopter can kill people itself, for whoever owns it. They’ll also increase the power of asymmetric warfare. Every robot car, cop or warplane will be as hackable, as everything is with sufficient expenditure, and the next 9/11 will be as difficult to definitively attribute as an attack by hackers on a U.S. company is today. Autonomous weapon systems are something

between guns in the early 20th century and nuclear weapons in the late 20th century, and we're hurtling toward it with no idea of how bad it could be. ... The thing to worry about is existing power structures building remote-control police forces and remote-control occupying armies. That threat is on the level of nuclear weapons. It's really, really dangerous."

**Susan Price**, user-experience pioneer and strategist and founder of Firecat Studio, wrote, "I don't believe that governments and regulatory agencies are poised to understand the implications of AI for ethics and consumer or voter protection. The questions asked in Congress barely scratch the surface of the issue, and political posturing too often takes the place of elected officials charged with oversight to reach genuine understanding of these complex issues. The strong profit motive for tech companies leads them to resist any such protections or regulation. These companies' profitability allows them to directly influence legislators through lobbies and PACs; easily overwhelming the efforts of consumer protection agencies and nonprofits, when those are not directly defunded or disbanded. We're seeing Facebook, Google, Twitter and Amazon resist efforts to produce the oversight, auditing and transparency that would lead to consumer protection. AI is already making lives better. But it's also making corporate profits better at a much faster rate. Without strong regulation, we can't correct that imbalance, and the processes designed to protect U.S. citizens from exploitation through elected leaders is similarly subverted by funds from these same large companies."

**Craig Spiegle**, managing director and trust strategist for Agelight, and chair emeritus for the Online Trust Alliance, said, "Look no further than data privacy and other related issues such as net neutrality. Industry in general has failed to respond ethically in the collection, use and sharing of data. Many of these same leaders have a major play in AI, and I fear they will continue to act in their own self-interests."

**Sam Punnett**, futurist and retired owner of FAD Research, commented, "System and application design is usually mandated by a business case, not by ethical considerations. Any forms of regulation or guidelines typically lag technology development by many years. The most concerning applications of AI systems are those being employed for surveillance and societal control."

**An ethics expert who served as an advisor on the UK's report on "AI in Health care"** responded, "I don't think the tech companies understand ethics at all. They can only grasp it in algorithmic form, i.e., a kind of automated utilitarianism, or via 'value alignment,' which tends to use economists' techniques around revealed preferences and social choice theory. They cannot think in terms of obligation, responsibility, solidarity, justice or virtue. This means they engineer out much of what is distinctive about humane ethical thought. In a thought I saw attributed to Hannah Arendt recently, though I cannot find the source, 'It is not that behaviourism is true, it is more that it might become true: That is the problem.' It would be racist to say that in some parts

of the world AI developers care less about ethics than in others; more likely, they care about different ethical questions in different ways. But underlying all that is that the machine learning models used are antithetical to humane ethics in their mode of operation.”

**Nathalie Maréchal**, senior research analyst at Ranking Digital Rights, observed, “Until the development and use of AI systems is grounded in an international human rights framework, and until governments regulate AI following human rights principles and develop a comprehensive system for mandating human rights impact assessments, auditing systems to ensure they work as intended, and hold violating entities to account, ‘AI for good’ will continue to be an empty slogan.”

**Mark Maben**, a general manager at Seton Hall University, wrote, “It is simply not in the DNA of our current economic and political system to put the public good first. If the people designing, implementing, using and regulating AI are not utilizing ethical principles focused primarily on the public good, they have no incentive to create an AI-run world that utilizes those principles. Having AI that is designed to serve the public good above all else can only come about through intense public pressure. Businesses and politicians often need to be pushed to do the right thing. Fortunately, the United States appears to be at a moment where such pressure and change [are] possible, if not likely. As someone who works with Gen Z nearly every day, I have observed that many members of Gen Z think deeply about ethical issues, including as they relate to AI. This generation may prove to be the difference makers on whether we get AI that is primarily guided by ethical principles focused on the public good.”

**Arthur Bushkin**, writer, philanthropist and social activist, said, “I worry that AI will not be driven by ethics, but rather by technological efficiency and other factors.”

**Dharmendra K. Sachdev**, a telecommunications pioneer and founder-president of Spacetel Consultancy LLC, wrote, “My simplistic definition is that AI can be smart; in other words, like the human mind, it can change directions depending upon the data collected. The question often debated is this: Can AI outsmart humans? My simplistic answer: Yes, in some humans but not the designer. A rough parallel would be: Can a student outsmart his professor? Yes, of course yes, but he may not outsmart all professors in his field. Summarizing my admittedly limited understanding is that all software is created to perform a set of functions. When you equip it with the ability to change course depending upon data, we call it AI. If I can make it more agile than my competition, my AI can outsmart him.”

**Karen Yesinkus**, a creative and digital services professional, observed, “I would like to believe that AI being used ethically by 2030 will be in place. However, I don’t think that will likely be a sure thing. Social media, human resources, customer services, etc. platforms are and will have



continuing issues to iron out (bias issues especially). Given the existing climate politically on a global scale, it will take more than the next 10 years for AI to shake off such bias.”

**Marc H. Noble**, a retired technology developer/administrator, wrote, “Although I believe most AI will be developed for the benefit of mankind, my great concern is that you only need one bad group to develop AI for the wrong reasons to create a potential catastrophe. Despite that, AI should be explored and developed, however, with a great deal of caution.”

**Eduardo Villanueva-Mansilla**, associate professor of communications at Pontificia Universidad Catolica, Peru, predicted, “Public pressure will be put upon AI actors. However, there is a significant risk that the agreed [-upon] ethical principles will be shaped too closely to the societal and political demands of the developed world. They will not consider the needs of emerging economies or local communities in the developing world.”

**Garth Graham**, a longtime leader of Telecommunities Canada, said, “The drive in governance worldwide to eradicate the public good in favour of market-based approaches is inexorable. The drive to implement AI-based systems is not going to place the public good as a primary priority. For example, existing Smart City initiatives are quite willing to outsource the design and operation of complex adaptive systems that learn as they operate civic functions, not recognizing that the operation of such systems is replacing the functions of governance.”

The AI genie is already out of the bottle, abuses are already occurring and some are not very visible and hard to remedy

A share of these experts note that AI applications designed with little or no attention to ethical considerations are already deeply embedded across many aspects of human activity, and they are generally invisible to the people they affect. These respondents said algorithms are at work in systems that are opaque at best and impossible to dissect at worst. They argue that it is highly unlikely that ethical standards can or will be applied in this setting. Others also point out that there is a common dynamic that plays out when new technologies sweep through societies: Abuses occur first and then remedies are attempted. It’s hard to program algorithm-based digital tools in a way that predicts, addresses and subverts all problems. Most problems remain unknown until they are recognized, sometimes long after they are produced, distributed and actively in use.

**Henning Schulzrinne**, Internet Hall of Fame member and former chief technology officer for the Federal Communications Commission, said, “The answer strongly depends on the shape of the government in place in the country in the next few years. In a purely deregulatory environment with strong backsliding toward law-and-order populism, there will be plenty of suppliers of AI that will have little concern about the fine points of AI ethics. Much of that AI will not be visible to the

public – it will be employed by health insurance companies that are again free to price-discriminate based on preexisting conditions, by employers looking for employees who won't cause trouble, by others who will want to nip any unionization efforts in the bud, by election campaigns targeting narrow subgroups.”

**Jeff Johnson**, a professor of computer science, University of San Francisco, who previously worked at Xerox, HP Labs and Sun Microsystems, responded, “The question asks about ‘most AI systems.’ Many new applications of AI will be developed to improve business operations. Some of these will be ethical and some will not be. Many new applications of AI will be developed to aid consumers. Most will be ethical, but some won't be. However, the vast majority of new AI applications will be ‘dark,’ i.e., hidden from public view, developed for military or criminal purposes. If we count those, then the answer to the question about ‘most AI systems’ is without a doubt that AI will be used mostly for unethical purposes.”

**John Harlow**, smart cities research specialist at the Engagement Lab @ Emerson College, predicted, “AI will mostly be used in questionable ways in the next decade. Why? That's how it's been used thus far, and we aren't training or embedding ethicists where AI is under development, so why would anything change? What gives me the most hope is that AI dead-ends into known effective use cases and known ‘impossibilities.’ Maybe AI can be great at certain things, but let's dispense with areas where we only have garbage in (applications based on any historically biased data). Most AI applications that make a difference in the lives of most people will be in the backend, invisible to them. ‘Wow, the last iOS update really improved predictive text suggestions.’ ‘Oh, my dentist has AI-informed radiology software?’ One of the ways it could go mainstream is through comedy. AI weirdness is an accessible genre, and a way to learn/teach about the technology (somewhat) – I guess that might break through more as an entertainment niche. As for global AI competition, what concerns me is the focus on AI, beating other countries at AI and STEM generally. Our challenges certainly call for rational methods. Yet, we have major problems that can't be solved without historical grounding, functioning societies, collaboration, artistic inspiration and many other things that suffer from overfocusing on STEM or AI.”

**Steve Jones**, professor of communication at the University of Illinois at Chicago and editor of *New Media and Society*, commented, “We'll have more discussion, more debate, more principles, but it's hard to imagine that there'll be – in the U.S. case – a will among politicians and policymakers to establish and enforce laws based on ethical principles concerning AI. We tend to legislate the barn after the horses have left. I'd expect we'll do the same in this case.”

**Andy Opel**, professor of communications at Florida State University, said, “Because AI is likely to gain access to a widening gyre of personal and societal data, constraining that data to serve a narrow economic or political interest will be difficult.”

**Doug Schepers**, a longtime expert in web technologies and founder of Fizz Studio, observed, “As today, there will be a range of deliberately ethical computing, poor-quality inadvertent unethical computing and deliberately unethical computing using AI. Deepfakes are going to be worrisome for politics and other social activities. It will lead to distrustability overall. By themselves, most researchers or product designers will not rigorously pursue ethical AI, just as most people don’t understand or rigorously apply principles of digital accessibility for people with disabilities. It’ll largely be inadvertent oversight, but it will still be a poor outcome. My hope is that best practices will emerge and continue to be refined through communities of practice, much like peer review in science. I also have some hope that laws may be passed that codify some of the most obvious best practices, much like the Americans With Disabilities Act and Section 508 improve accessibility through regulation, while still not being overly onerous. My fear is that some laws will be stifling, like those regarding stem-cell research. Machine learning and AI naturally have the capacity for improving people’s lives in many untold ways, such as computer vision for blind people. This will be incremental, just as commodity computing and increasing internet have improved (and sometimes harmed) people. It will most likely not be a seismic shift, but a drift. One of the darker aspects in the existing increase of surveillance capitalism and its use by authoritarian states. My hope is that laws will rein this in.”

**Jay Owens**, research director at pulsarplatform.com and author of HautePop, said, “Computer science education – and Silicon Valley ideology overall – focuses on ‘what can be done’ (the technical question) without much consideration of ‘should it be done’ (a social and political question). Tech culture would have to turn on its head for ethical issues to become front-and-centre of AI research and deployment; this is vanishingly unlikely. I’d expect developments in machine learning to continue along the same lines they have done so for the last decade – mostly ignoring the ethics question, with occasional bursts of controversy when anything particularly sexist or racist occurs. A lot of machine learning is already (and will continue to be) invisible to people’s everyday lives but creating process efficiencies (e.g., in weather forecasting, warehousing and logistics, transportation management). Other processes that we might not want to be more efficient (e.g., oil and gas exploration, using satellite imagery and geology analysis) will also benefit. I feel positively toward systems where ML and human decision-making are combined (e.g., systems for medical diagnostics). I would imagine machine learning is used in climate modelling, which is also obviously helpful. Chinese technological development cannot be expected to follow Western ethical qualms, and, given the totalitarian (and genocidal) nature of this state, it is likely that it will produce some ML systems that achieve these policing ends. Chinese-owned social apps such as TikTok have already shown racial biases and are likely less motivated to address them. I see no prospect that ‘general AI’ or generalisable machine intelligence will be achieved in 10 years and even less reason to panic about this (as some weirdos in Silicon Valley do).”

**Robert W. Ferguson**, a hardware robotics engineer at Carnegie Mellon Software Engineering Institute, wrote, “How many times do we need to say it? Unsupervised machine learning is at best incomplete. If supplemented with a published causal analysis, it might recover some credibility. Otherwise, we suffer from what is said by Cathy O’Neil in [‘Weapons of Math Destruction.’](#) Unsupervised machine learning without causal analysis is irresponsible and bad.”

**Michael Richardson**, open-source consulting engineer, responded, “In the 1980s, ‘AI’ was called ‘expert systems,’ because we recognized that it wasn’t ‘intelligent.’ In the 2010s, we called it ‘machine learning’ for the same reason. ML is just a new way to build expert systems. They replicate the biases of the ‘experts’ and cannot see beyond them. Is algorithmic trading ethical? Let me rephrase: Does our economy actually need it? If the same algorithm is used to balance ecosystems, does the answer change? We already have AI. They are called ‘corporations.’ Many have pointed this out already. Automation of that collective mind is really what is being referred to. I believe that use of AI in sentencing violates people’s constitutional rights, and I think that it will be stopped as it is realised that it just institutionalises racism.”

**A principal architect at a technology company** said, “I see no framework or ability for any governing agencies to understand how AI works. Practitioners don’t even know how it works, and they keep the information as proprietary information. Consider how long it took to mandate seat belts or limit workplace smoking, where the cause and effect were so clear, how can we possibly hope to control AI within the next 10 years?”

Global competition, especially between China and the U.S., will matter more to the development of AI than any ethical issues

A number of these respondents framed their answers around the “arms race” dynamic driving the tech superpowers, noting that it instills a damn-the-ethics-full-speed-ahead attitude. Some said there are significant differences in the ethical considerations various nation-states are applying and will apply in the future to AI development. Many pointed to the U.S. and China as the leading competitors in the nation-states arms race.

**Daniel Farber**, author, historian and professor of law at the University of California-Berkeley, responded, “There’s enormous uncertainty. Why? First of all, China. That’s a huge chunk of the world, and there’s nothing in what I see there right now to make me optimistic about their use of AI. Second, AI in the U.S. is mostly in the hands of corporations whose main goal is naturally to maximize profits. They will be under some pressure to incorporate ethics both from the public and employees, which will be a moderating influence. The fundamental problem is that AI is likely to be in the hands of institutions and people that already have power and resources, and that will inevitably shape how the technology is used. So, I worry that it will simply reinforce or increase

current power imbalances. What we need is not only ethical AI but ethical access to AI, so that individuals can use it to increase their own capabilities.”

**J. Scott Marcus**, an economist, political scientist and engineer who works as a telecommunications consultant, wrote, “Policy fragmentation globally will get in the way. As long as most AI investment is made in the U.S. and China, no consensus is possible. The European Union will attempt to bring rules into play, but it is not clear if they can drive much change in the face of the U.S. and China rivalry. The U.S. (also Japan) are large players in consumption but not so large in production of many aspects. They are larger, however, in IoT and robotics, so maybe there is more hope there. For privacy, the European Union forced a fair degree of global convergence thanks to its large purchasing power. It is not clear whether that can work for AI.”

**Charles M. Ess**, a professor of media studies at the University of Oslo whose expertise is in information and computing ethics, commented, “The most hope lies in the European Union and related efforts to develop ‘ethical AI’ in both policy and law. Many first-rate people and reasonably solid institutions are working on this, and, in my view, some promising progress is being made. But the EU is squeezed between China and the U.S. as the world leaders, neither of which can be expected to take what might be called ethical leadership. China is at the forefront of exporting the technologies of ‘digital authoritarianism.’ Whatever important cultural caveats may be made about a more collective society finding these technologies of surveillance and control positive as they reward pro-social behavior – the clash with the foundational assumptions of democracy, including rights to privacy, freedom of expression, etc. is unavoidable and unquestionable.

“For its part, the U.S. has a miserable record (at best) of attempting to regulate these technologies – starting with computer law from the 1970s that categorizes these companies as carriers, not content providers, and thus not subject to regulation that would include attention to freedom of speech issues, etc. My prediction is that Google and its corporate counterparts in Silicon Valley will continue to successfully argue against any sort of external regulation or imposition of standards for an ethical AI, in the name of having to succeed in the global competition with China. We should perhaps give Google in particular some benefit of the doubt and see how its recent initiatives in the direction of ethical AI in fact play out. But 1) what I know first-hand to be successful efforts at ethics-washing by Google (e.g., attempting to hire in some of its more severe and prominent ethical critics in the academy in order to buy their silence), and 2) given its track record of cooperation with authoritarian regimes, including China, it’s hard to be optimistic here.

“Of course, we will see some wonderfully positive developments and improvements – perhaps in medicine first of all. And perhaps it’s okay to have recommender systems to help us negotiate, e.g., millions of song choices on Spotify. But even these applications are subject to important critique, e.g., under the name of ‘the algorithmization of taste’ – the reshaping of our tastes and preferences

is influenced by opaque processes driven by corporate interests in maximizing our engagement and consumption, not necessarily helping us discover liberating and empowering new possibilities. More starkly, especially if AI and machine-learning techniques remain black-boxed and unpredictable, even to those who create them (which is what AI and ML are intended to do, after all), I mostly see a very dark and nightmarish future in which more and more of our behaviors are monitored and then nudged by algorithmic processes we cannot understand and thereby contest. The starkest current examples are in the areas of so-called ‘predictive policing’ and related efforts to replace human judgment with machine-based ‘decision-making.’ As [Mireille Hildebrandt](#) has demonstrated, when we can no longer contest the evidence presented against us in a court of law – because it is gathered and processed by algorithmic processes even its creators cannot clarify or unpack – that is the end of the modern practices of law and democracy. It’s clearly bad enough when these technologies are used to sort out human beings in terms of their credit ratings: Relying on these technologies for judgments/decisions about who gets into what educational institution, who does and does not deserve parole, and so on seem to me to be a staggeringly nightmarish dystopian future.

“Again, it may be a ‘Brave New World’ of convenience and ease, at least as long as one complies with the behaviors determined to be worth positive reward, etc. But to use a different metaphor – one perhaps unfamiliar to younger generations, unfortunately – we will remain the human equivalent of Skinner pigeons in nice and comfortable Skinner cages, wired carefully to maximize desired behaviors via positive reinforcement, if not discouraging what will be defined as undesirable behaviors via negative reinforcement (including force and violence) if need be.”

**Adam Clayton Powell III**, senior fellow at the USC Annenberg Center on Communication Leadership and Policy, observed, “By 2030, many will use ethical AI and many won’t. But in much of the world, it is clear that governments, especially totalitarian governments in China, Russia, et seq., will want to control AI within their borders, and they will have the resources to succeed. And those governments are only interested in self-preservation – not ethics.”

**Alf Rehn**, professor of innovation, design and management at the University of Southern Denmark, said, “There will be a push for ethical AI during the next 10 years, but good intentions alone do not morality make. AI is complicated, as is ethics, and combining the two will be a very complex problem indeed. We are likely to see quite a few clumsy attempts to create ethical AI-systems, with the attendant problems. It is also important to take cultural and geopolitical issues into consideration. There are many interpretations of ethics, and people put different value on different values, so that, e.g., a Chinese ethical AI may well function quite differently – and generate different outcomes – from, e.g., a British ethical AI. This is not to say that one is better than the other, just that they may be rather different.”

**Sean Mead**, senior director of strategy and analytics at Interbrand, wrote, “Chinese theft of Western and Japanese AI technologies is one of the most worrisome ethics issues that we will be facing. We will have ethical issues over both potential biases built into AI systems through the choice or availability of training data and expertise sets and the biases inherent in proposed solutions attempting to counter such problems. The identification systems for autonomous weapons systems will continue to raise numerous ethics issues, particularly as countries deploy land-based systems interacting with people. AI driving social credit systems will have too much power over peoples’ lives and will help vitalize authoritarian systems. AI will enable increased flight from cities into more hospitable and healthy living areas through automation of governmental services and increased transparency of skill sets to potential employers.”

**Mark Perkins**, an information science professional active in the Internet Society, noted, “AI will be developed by corporations (with government backing) with little respect for ethics. The example of China will be followed by other countries – development of AI by use of citizens’ data, without effective consent, to develop products not in the interest of such citizens (surveillance, population control, predictive policing, etc.). AI will also be developed to implement differential pricing/offers further enlarging the ‘digital divide’ AI will be used by both governments and corporations to take nontransparent, nonaccountable decisions regarding citizens AI will be treated as a ‘black box,’ with citizens having little – if any – understanding of how they function, on what basis they make decisions, etc.”

**Wendell Wallach**, ethicist and scholar at Yale University’s Interdisciplinary Center for Bioethics, responded, “While I applaud the proliferation of ethical principles, I remain concerned about the ability of countries to put meat on the bone. Broad principles do not easily translate into normative actions, and governments will have difficulty enacting strong regulations. Those that do take the lead in regulating digital technologies, such as the EU, will be criticized for slowing innovation, and this will remain a justification for governments and corporations to slow putting in place any strong regulations backed by enforcement. So far, ethics whitewashing is the prevailing approach among the corporate elite. While there are signs of a possible shift in this posture, I remain skeptical while hopeful.”

**Pamela McCorduck**, writer, consultant and author of several books, including “[Machines Who Think](#),” wrote, “Many efforts are underway worldwide to define ethical AI, suggesting that this is already considered a grave problem worthy of intense study and legal remedy. Eventually, a set of principles and precepts will come to define ethical AI, and I think they will define the preponderance of AI applications. But you can be assured that unethical AI will exist, be practiced and sometimes go unrecognized until serious damage is done. Much of the conflict between ethical and unethical applications is cultural. In the U.S. we would find the kind of social surveillance practiced in China to be not only repugnant – but illegal. It forms the heart of Chinese practices.

In the short term, only the unwillingness of Western courts to accept evidence gathered this way (as inadmissible) will protect Western citizens from this kind of thing, including the ‘social scores’ the Chinese government assigns to its citizens as a consequence of what surveillance turns up. I sense more everyday people will invest social capital in their interactions with AIs, out of loneliness or for other reasons. This is unwelcome to me, but then I have a wide social circle. Not everybody does, and I want to resist judgment here.”

**An architect of practice specializing in AI for a major global technology company**

said, “The European Union has the most concrete proposals, and I believe we will see their legislation in place within three years. My hope is that we will see a ripple effect in the U.S. like we did from GDPR – global companies had to comply with GDPR, so some good actions happened in the U.S. as a result. ... We may be more likely to see a continuation of individual cities and states imposing their own application-specific laws (e.g., facial-recognition technology limits in Oakland, Boston, etc.). The reasons I am doubtful that the majority of AI apps will be ethical/benefit the social good are:

- 1) Even the EU’s proposals are limited in what they will require;
- 2) China will never limit AI for social benefit over the government’s benefit;
- 3) The ability to create a collection of oversight organizations with the budget to audit and truly punish offenders is unlikely.

“I look at the Food and Drug Administration or NTSB [National Transportation Safety Board] and see how those organizations got too cozy with the companies they were supposed to regulate and see their failures. These organizations are regulating products much less complex than AI, so I have little faith the U.S. government will be up to the task. Again, maybe the EU will be better.”

**A researcher in bioinformatics and computational biology** observed, “Take into account the actions of the CCP [Chinese Communist Party] in China. They have been leading the way recently in demonstrating how these tools can be used in unethical ways. And the United States has failed to make strong commitments to ethics in AI, unlike EU nations. AI and the ethics surrounding its use could be one of the major ideological platforms for the incoming next Cold War. I am most concerned about the use of AI to further invade privacy and erode trust in institutions. I also worry about its use to shape policy in nontransparent, noninterpretable and nonreproducible ways. There is also the risk that some of the large datasets that are the fundamental to a lot of decision-making – from facial recognition, to criminal sentencing, to loan applications – being conducted using AI that are critically biased and will continue to produce biased outcomes if they are used without undergoing severe audits – issues with transparency compound these problems. Advances to medical treatment using AI run the risk of not being fairly distributed as well.”



**Sam Lehman-Wilzig**, professor and former chair of communication at Bar-Ilan University, Israel, said, “I am optimistic because the issue is now on the national agenda – scientific, academic and even political/legislative. I want to believe that scientists and engineers are somewhat more open to ‘ethical’ considerations than the usual ‘businesspeople.’ The major concern is what other (nondemocratic) countries might be doing – and whether we should be involved in such an ‘arms race,’ e.g., AI-automated weaponry. Thus, I foresee a move to international treaties dealing with the most worrisome aspects of ‘AI ethics.’”

**An economist who works in government** responded, “Ethical principles will be developed and applied in democratic countries by 2030, focusing on the public good, global competition and cyber breaches. Other less-democratic countries will be focused more on cyberbreaches and global competition. Nongovernmental entities such as private companies will presumably concentrate on innovation and other competitive responses. AI will have a considerable impact on people, especially regarding their jobs and also regarding their ability to impact the functions controlled by AI. This control and the impact of cybercrimes will be of great concern, and innovation will intrigue.”

**Ian Peter**, a pioneering internet rights activist, said, “The biggest threats we face are weaponisation of AI and development of AI being restricted within geopolitical alliances. We are already seeing the beginnings of this in actions taken to restrict activities of companies because they are seen to be threatening (e.g., Huawei). More and more developments in this field are being controlled by national interests or trade wars rather than ethical development, and much of the promise which could arise from AI utilisation may not be realised. Ethics is taking a second-row seat behind trade and geopolitical interests.”

**Jannick Pedersen**, a co-founder, CEO and futurist based in Europe, commented, “AI is the next arms race. Though mainstream AI applications will include ethical considerations, a large amount of AI will be made for profit and be applied in business systems, not visible to the users.”

**Marita Prandoni**, linguist, freelance writer, editor, translator and research associate with the Shape of History group, predicted, “Ethical uses of AI will dominate, but it will be a constant struggle against disruptive bots and international efforts to undermine nations. Algorithms have proven to magnify bias and engender injustice, so reliance on them for distracting, persuading or manipulating opinion is wrong. What excites me is that advertisers are rejecting platforms that allow for biased and dangerous hate speech and that increasingly there are economic drivers (i.e., corporate powers) that take the side of social justice.”

**Gus Hosein**, executive director of Privacy International, observed, “Unless AI becomes a competition problem and gets dominated by huge American and Chinese companies, then the

chances of ethical AI are low, which is a horrible reality. If it becomes widespread in deployment, as we've seen with facial recognition, then the only way to stem its deployment in unethical ways is to come up with clear bans and forced transparency. This is why AI is so challenging. Equally, it's quite pointless, but that won't stop us from trying to deploy it everywhere. The underlying data quality and societal issues mean that AI will just punish people in new, different and the same ways. If we continue to be obsessed with innovators and innovation rather than social infrastructure, then we are screwed.”

## 2. Hopes about developments in ethical AI

Early developments in AI have been of overwhelmingly great importance and value to society. Most of the experts responding to this canvassing – both the pessimists and the optimists – expect that it will continue to provide clear benefits to humanity. Those who are hopeful argue that its advance will naturally include mitigating activities, noting that these problems are too big to ignore. They said society will begin to better-anticipate potential harms and act to mute them. Among the commonly expressed points:

- Historically, ethics have evolved as new technologies mature and become embedded in cultures; as problems arise so do adjustments.
- Fixes are likely to roll out in different ways along different timelines in different domains.
- Expert panels concerned about ethical AI are being convened in many settings across the globe.
- Social upheavals arising due to AI problems are a force that may drive it closer to the top of human agendas.
- Political and judicial systems will be asked to keep abuses in check, and evolving case law will emerge (some experts are concerned this could be a net negative).
- AI itself can be used to assess AI impacts and hunt down unethical applications.
- A new generation of technologists whose training has been steeped with ethical thinking will lead the movement toward design that values people and positive progress above profit and power motives and the public will become wiser about the downsides of being code-dependent.

This section includes hopeful comments about the potential development of ethical AI.

### AI advances are inevitable; we will work on fostering ethical AI design

A number of these expert respondents noted breakthroughs that have already occurred in AI and said they imagine a future in which even more applications emerge to help solve problems and make people's lives easier and safer. They expect that AI design will evolve positively as these tools continue to influence the majority of human lives in mostly positive ways.

They especially focused on the likelihood that there will be more medical and scientific breakthroughs that help people live healthier and more productive lives, and they noted that there will be increasing efficiency in AI quickly mastering most tasks. They said AI tools are simply better than humans at pattern-recognition and crunching massive amounts of data. Some said they expect AI will expand positively to augment humans, working in sync as their ally.

**Benjamin Grosf**, chief scientist at Kyndi, a Silicon Valley start-up aimed at the reasoning and knowledge representation side of AI, wrote, “Some things that give me hope are the following: Most AI technical researchers (as distinguished from business or government deployers of AI) care quite a lot about ethicality of AI. It has tremendous potential to improve productivity economically and to save people effort even when money is not flowing directly by better automating decisions and information analysis/supply in a broad range of work processes. Conversational assistants and question-answering, smarter-workflow and manufacturing robots are some examples where I foresee AI applications making a positive difference in the lives of most people, either indirectly or directly. I am excited by the fact that many national governments are increasing funding for scientific research in AI. I am concerned that so much of that is directed toward military purposes or controlled by military branches of governments.”

**Perry Hewitt**, chief marketing officer at data.org, responded, “I am hopeful that ‘ethical AI’ will extend beyond the lexicon to the code by 2030. The awareness of the risks gives me the most hope. For example, for centuries we have put white men in judicial robes and trusted them to make the right decisions and pretended that biases, proximity to lunchtime and the case immediately preceding had no effect on the outcome. Scale those decisions with AI and the flaws emerge. And when these flaws are visible, effective regulation can begin. This is the decade of narrow AI – specific applications that will affect everything from the jobs you are shown on LinkedIn to the new sneakers advertised to you on Instagram. Clearly, the former makes more of a difference than the latter for your economic well-being, but in all cases, lives are changed by AI under the hood. Transparency around the use of AI will make a difference as will effective regulation.”

**Donald A. Hicks**, a professor of public policy and political economy at the University of Dallas whose research specialty is technological innovation, observed, “AI/automation technologies do not assert themselves. They are always invited in by investors, adopters and implementers. They require investments to be made by someone, and those investments are organized around the benefits of costs cut or reduced and/or possibilities for new revenue flows. New technologies that cannot offer those prospects remain on the shelf. So, inevitably, new technologies like AI only proliferate if they are ‘pulled’ into use. This gives great power to users and leads to applications that look to be beneficial to a widening user base. This whole process ‘tilts’ toward long-term ethical usage. I know of no technology that endured while delivering unwanted/unfair outcomes broadly. Consider our nation’s history with slavery. Gradually and inevitably, as the agricultural economies of Southern states became industrialized, it no longer made sense to use slaves. It was not necessary to change men’s hearts before slavery could be abandoned. Economic outcomes mattered more, although eventually hearts did follow.

“But again, the transitions between old and new do engender displacements via turnover and replacement, and certain people and places can feel – and are – harmed by such technology-

driven changes. But their children are likely thankful that their lives are not simply linear extensions of the lives of their parents. To date, AI and automation have had their greatest impacts in augmenting the capabilities of humans, not supplanting them. The more sophisticated AI applications and automation become, the more we appreciate the special capabilities in human beings that are of ultimate value and that are not likely to be captured by even the most sophisticated software programs. I'm bullish on where all of this is leading us because I'm old enough to compare today with yesterday."

**Jim Spohrer**, director of cognitive open technologies and the AI developer ecosystem at IBM, noted, "[The Linux Foundation Artificial Intelligence Trusted AI Committee](#) is working on this. The members of that community are taking steps to put principles in place and collect examples of industry use cases. The contribution into Linux Foundation AI (by major technology companies) of the open-source project code for Trusted AI for AI-Fairness, AI-Robustness and AI-Explainability on which their products are based is a very positive sign."

**Michael Wollowski**, a professor of computer science at Rose-Hulman Institute of Technology and expert in artificial intelligence, said, "It would be unethical to develop systems that do not abide by ethical codes, *if* we can develop those systems to be ethical. Europe will insist that systems will abide by ethical codes. Since Europe is a big market, since developing systems that abide by ethical code is not a trivial endeavor and since the big tech companies (except for Facebook) by and large want to do good (well, their employees by and large want to work for companies that do good), they will develop their systems in a way that they abide by ethical codes. I very much doubt that the big tech companies are interested (or are able to find young guns) in maintaining an unethical version of their systems.

"AI systems, in concert with continued automation, including the Internet of Things, will bring many conveniences. Think along the lines of personal assistants [who] manage various aspects of people's lives. Up until COVID-19, I would have been concerned about bad actors using AI to do harm. I am sure that right now bad actors are probably hiring virologists to develop viruses with which they can hold the world hostage. I am very serious that rogue leaders are thinking about this possibility. The AI community in the U.S. is working very hard to establish a few large research labs. This is exciting, as it enables the AI community to develop and test systems at scale. Many good things will come out of those initiatives. Finally, let us not forget that AI systems are engineered systems. They can do many interesting things, but they cannot think or understand. While they can be used to automate many things and while people by and large are creatures of habit, it is my fond hope that we will rediscover what it means to be human."

**Paul Epping**, chairman and co-founder of XponentialEQ and well-known keynote speaker on exponential change, wrote, "The power of AI and machine learning (and deep learning) is

underestimated. The speed of advancements is incredible and will lead to automating of virtually *all* processes (blue- and white-collar jobs). In health care: Early detection of diseases, fully AI-driven triage, including info from sensors (on or inside your body), leading to personalised health (note: *not* personalised medicine). AI will help to compose the right medication for you – and not the generic stuff that we get today, surpassing what the pharmaceutical industry is doing. AI is helping to solve the world’s biggest problems, finding new materials, running simulations, digital twins (including personal digital twins that can be used to run experiments in case of treatments). My biggest concern: How are we going to solve the control problem? ([Read Stuart Russell’s ‘Human Compatible’](#)) and follow the [Future of Life Institute](#) and the problem of biased data and algorithms.)”

**Adel Elmaghraby**, a leader in IEEE and professor and former chairman of the Computer Engineering and Computer Science Department at the University of Louisville, responded, “Societal pressure will be a positive influence for adoption of ethical and transparent approaches to AI. However, the uncomfortable greed for political and financial benefit will need to be reined in.”

**Gregory Shannon**, chief scientist at the CERT software engineering institute at Carnegie Mellon University, said, “There will be lots of unethical applications as AI matures as an engineering discipline. I expect that to improve. Just like there are unethical uses of technology today, there will be for AI. AI provides transformative levels of efficiency for digesting information and making pretty good decisions. And some will certainly exploit that in unethical ways. However, the ‘demand’ from the market (most of the world’s population) will be for ethical AI products and services. It will be bumpy, and in 2030 we might be halfway there. The use of AI by totalitarian and authoritarian governments is a clear concern. But I don’t expect the use of such tech to overcome the yearning of populations for agency in their lives, at least after a few decades of such repression. Unethical systems/solutions are not trustworthy. So, they can only have narrow application. Ethical systems/solutions will be more widely adopted, eventually.”

**Robert D. Atkinson**, president of the Information Technology and Innovation Foundation, wrote, “The real question is not whether all AI developers sign up to some code of principles, but rather whether most AI applications work in ways that society expects them to, and the answer to that question is almost 100% ‘yes.’”

**Ben Shneiderman**, distinguished professor of computer science and founder of the Human-Computer Interaction Lab at the University of Maryland, commented, “While technology raises many serious problems, efforts to limit malicious actors should eventually succeed and make these technologies safer. The huge interest in ethical principles for AI and other technologies is beginning to shift attention toward practical steps that will produce positive change. Already the

language of responsible and human-centered AI is changing the technology, guiding students in new ways and reframing the work of researchers. ... I foresee improved appliance-like and tele-operated devices with highly automated systems that are reliable, safe and trustworthy. Shoshanna Zuboff's analysis in her book '[Surveillance Capitalism](#)' describes the dangers and also raises awareness enough to promote some changes. I believe the arrival of independent oversight methods will help in many cases. Facebook's current semi-independent oversight board is a small step forward, but changing Facebook's culture and Zuckerberg's attitudes is a high priority to ensuring better outcomes. True change will come when corporate business choices are designed to limit the activity of malicious actors – criminals, political operatives, hate groups and terrorists – while increasing user privacy.”

**Carol Smith**, a senior research scientist in human–machine interaction at Carnegie Mellon University's Software Engineering Institute, said, “There are still many lessons to be learned with regard to AI and very little in the way of regulation to support human rights and safety. I'm hopeful that the current conversations about AI ethics are being heard, and that, as we see tremendous misuse and abuse of these systems, the next generation will be much more concerned about ethical implications. I'm concerned that many people, organizations and governments see only monetary gain from unethical applications of these technologies and will continue to misuse and abuse data and AI systems for as long as they can. AI systems short-term will continue to replace humans in dull, dirty, dangerous and dear work. This is good for overall safety and quality of life but is bad for family livelihoods. We need to invest in making sure that people can continue to contribute to society when their jobs are replaced. Longer-term, these systems will begin to make many people more efficient and effective at their jobs. I see AI systems improving nearly every industry and area of our lives when used properly. Humans must be kept in the loop with regard to decisions involving people's lives, quality of life, health and reputation, and humans must be ultimately responsible for all AI decisions and recommendations (not the AI system).”

**Marvin Borisch**, chief technology officer at RED Eagle Digital based in Berlin, wrote, “When used for the greater good, AI can and will help us fight a lot of human problems in the next decade. Prediagnostics, fair ratings for insurance or similar, supporting humans in space and other exploration and giving us theoretical solutions for economic and ecological problems – these are just a few examples of how AI is already helping us and can and will help us in the future. If we focus on solving specific human problems or using AI as a support for human work instead of replacing human work, I am sure that we can and will tackle any problem. What worries me the most is that AI developers are trying to trump each other – not for the better use but for the most medial outcome in order to impress stakeholders and potential investors.”

**Tim Bray**, well-known technology leader who has worked for Amazon, Google and Sun Microsystems, predicted, “Unethical AI-driven behavior will produce sufficiently painful effects

that legal and regulatory frameworks will be imposed that make its production and deployment unacceptable.”

**Gary M. Grossman**, associate director in the School for the Future of Innovation in Society at Arizona State University, responded, “AI will be used in both ethical and questionable ways in the next decade. Such is the nature of the beast, and we, the beasts that will make the ethical choices. I do not think policy alone will be sufficient to ensure ethical choices to be made every time. Like everything else, it will stabilize in some type of compromised structure within the decade time frame the question anticipates.”

**Erhardt Graeff**, a researcher expert in the design and use of digital technologies for civic and political engagement, noted, “Ethical AI is boring directly into the heart of the machine-learning community and, most importantly, influencing how it is taught in the academy. By 2030, we will have a generation of AI professionals that will see ethics as inseparable from their technical work. Companies wishing to hire these professionals will need to have clear ethical practices built into their engineering divisions and strong accountability to the public good at the top of their org charts. This will certainly describe the situation at the major software companies like Alphabet, Apple, Facebook, Microsoft and Salesforce, whose products are used on a massive scale. Hopefully, smaller companies and those that don’t draw the same level of scrutiny from regulators and private citizens will adopt similar practices and join ethical AI consortia and find themselves staffed with upstanding technologists. One application of AI that will touch nearly all sectors and working people is in human resources and payroll technology. I expect we will see new regulation and scrutiny of those tools and the major vendors that provide them.

“I caveat my hopes for ethical AI with three ways unethical AI will persist.

- 1) There will continue to be a market for unethical AI, especially the growing desire for surveillance tools from governments, corporations and powerful individuals.
- 2) The democratization of machine learning as APIs, simple libraries and embedded products will allow many people who have not learned to apply this technology in careful ways to build problematic tools and perform bad data analysis for limited, but meaningful distributions that will be hard to hold to account.
- 3) A patchwork of regulations across national and international jurisdictions and fights over ethical AI standards will undermine attempts to independently regulate technology companies and their code through auditing and clear mechanisms for accountability.”

**Katie McAuliffe**, executive director for Digital Liberty, wrote, “There are going to be mistakes in AI, even when companies and coders try their best. We need to be patient with the mistakes, find them and adjust. We need to accept that some mistakes don’t equal failure in the entire system.



No, AI will not be used in mostly questionable ways. We are using forms of AI every day already. The thing about AI is that, once it works, we call it something else. With a new name, it's not as amorphous and threatening. AI and machine learning will benefit us the most in the health context – being able to examine thousands of possibilities and variables in a few seconds, but human professionals will always have to examine the data and context to apply any results. We need to be sure that something like insurance doesn't affect a doctor or researcher's readout in these contexts."

**Su Sonia Herring**, a Turkish-American internet policy researcher with Global Internet Policy Digital watch, said, "AI will be used in questionable ways due to companies and governments putting profit and control in front of ethical principles and the public good. Civil society, researchers and institutions who are concerned about human rights give me hope. Algorithmic black boxes, digital divide, the need to control, surveil and profit off masses worry me the most. I see AI applications making a difference in people's lives in taking care of making mundane, time-consuming work (while making certain jobs obsolete), helping identify trends and informing public policy. Issues related to privacy, security and accountability and transparency related to AI tech concerns me, while the potential of processing big data to solve global issues excites me."

**Edson Prestes**, a professor of computer science at Federal University of Rio Grande do Sul, Brazil, commented, "By 2030, technology in general will be developed taking into account ethical considerations. We are witnessing a huge movement these days. Most people who have access to information are worried about the misuse of technology and its impact on their own lives. Campaigns to ban lethal weapons powered by AI are growing. Discussions on the role of technology and its impact on jobs are also growing. People are becoming more aware about fake news and proliferation of hate speech. All these efforts are creating a massive channel of information and awareness. Some communities will be left behind, either because some governments want to keep their citizens in poverty and consequently keep them under control, or because they do not have enough infrastructure and human and institutional capacities to reap the benefits of the technological domain. In these cases, efforts led by the United Nations are extremely valuable. The UN Secretary-General António Guterres is a visionary in establishing the High-Level Panel on Digital Cooperation. Guterres used the panel's recommendations to create a roadmap with concrete actions that address the digital domain in a holistic way, engaging a wide group of organisations to deal with the consequences emerging from the digital domain."

**James Blodgett**, futurist, author and consultant, said, "Focused primarily on the public good' is not enough if the exception is a paperclip maximizer. '[Paperclip maximizer](#)' is an improbable metaphor, but it makes the point that one big mistake can be enough. We can't refuse to do anything – because not to decide is also a decision. The best we can do is to think carefully, pick

what seems to be the best plan and execute, perhaps damning [the] torpedoes as part of that execution. But we had better think very carefully and be very careful.”

**A futurist and managing principal for a consultancy** commented, “AI offers extremely beneficial opportunities, but only if we actively address the ethical principles and regulate and work toward:

- 1) Demographically balanced human genome databases,
- 2) Gender-balanced human genome databases (especially in the area of clinical drug trials where women are severely under tested),
- 3) Real rules around the impact of algorithms and machine learning developed by poor data collection of availability. We see this in the use of facial recognition, police data, education data, Western bias in humanities collections, etc.

“AI also has the potential to once again be a job killer but also assist the practice medicine, law enforcement, etc. It can also be a job creator, but countries outside the U.S. are making greater progress on building AI hubs.”

**A European business leader** argued, “I do believe AI systems will be governed by ethics, but they will be equal parts new legislation and litigation avoidance. If your AI rejects my insurance claim, I can sue to find out why and ensure your learning model wasn’t built on, say, racially biased source data.”

**Christina J. Colclough**, an expert on the future of work and the politics of technology and ethics in AI, observed, “By 2030, governments will have woken up to the huge challenges AI (semi/autonomous systems, machine learning, predictive analytics, etc.) pose to democracy, legal compliance and our human and fundamental rights. What is necessary is that these ‘ethical principles’ are enforceable and governed. Otherwise, they risk being good intentions with little effect.”

**Thomas Birkland**, professor of public and international affairs at North Carolina State University, wrote, “AI will be informed by ethical considerations in the coming years because the stakes for companies and organizations making investments in AI are too high. However, I am not sure that these ethical considerations are going to be evenly applied, and I am not sure how carefully these ethical precepts will be adopted. What gives me the most hope is the widespread discussions that are already occurring about ethics in AI – no one will be caught by surprise by the need for an ethical approach to AI. What worries me the most is that the benefits of such systems are likely to flow to the wealthy and powerful. For example, we know that facial-recognition software, which is often grounded in AI, has severe accuracy problems in recognizing ‘nonwhite’

faces. This is a significant problem. AI systems may be able to increase productivity and accuracy in systems that require significant human intervention. I am somewhat familiar with AI systems that, for example, can read x-rays and other scans to look for signs of disease that may not be immediately spotted by a physician or radiologist. AI can also aid in pattern recognition in large sets of social data. For example, AI systems may aid researchers in coding data relating to the correlates of health. What worries me is the uncritical use of AI systems without human intervention. There has been some talk, for example, of AI applications to warfare – do we leave weapons targeting decisions to AI? This is a simplistic example, but it illustrates the problem of ensuring that AI systems do not replace humans as the ultimate decision-maker, particularly in areas where there are ethical considerations. All this being said, the deployment of AI is going to be more evolutionary than revolutionary, and that the effects on our daily lives will be subtle and incremental over time.”

**Jerome C. Glenn**, co-founder and CEO of the futures-research organization The Millennium Project, wrote, “There were few discussions about the ethical issues in the early spread of the internet in the 1970s and 1980s. Now there are far, far more discussions about AI around the world. However, most do not make clear distinctions among narrow, general and super AI. If we don’t get our standards right in the transition from artificial narrow intelligence to artificial general intelligence, then the emergence of super from general could have the consequences science fiction has warned about.”

**Benjamin Kuipers**, a professor of computer science and engineering at the University of Michigan known for research in qualitative simulation, observed, “I choose to believe that things will work out well in the choice we face as a society. I can’t predict the likelihood of that outcome. Ethics is the set of norms that society provides, telling individuals how to be trustworthy, because trust is essential for cooperation, and cooperation is essential for a society to thrive, and even to survive (see Robert Wright’s book ‘*Nonzero*’). Yes, we need to understand ethics well enough to program AIs so they behave ethically. More importantly, we need to understand that corporations, including nonprofits, governments, churches, etc., are also artificially intelligent entities participating in society, and they need to behave ethically. We also need to understand that we as a society have been spreading an ideology that teaches individuals that they should behave selfishly rather than ethically. We need an ethical society, not just ethical AI. But AI gives us new tools to understand the mind, including ethics.”

**John Verdon**, a retired complexity and foresight consultant, said, “Ultimately what is most profitable in the long run is a well-endowed citizenry able to pursue their curiosities and expand the agencies. To enable this condition will require the appropriate legislative protections and constraints. The problems of today and the future are increasing in complexity. Any systems that seek monopolistic malevolence essentially will act like a cancer killing its own host. Distributed-

ledger technologies may well enable the necessary ‘accounting systems’ to both credit creators and users of what has been created while liberating creations to be used freely (like both free beer and liberty). This enables a capacity to unleash all human creativity to explore the problem and possibility space. Slaves don’t create a flourishing society – only a static and fearful elite increasingly unable to solve the problems they create. New institutions like an ‘auditor general of algorithms’ (to oversee that algorithms and other computations actually produce the results they intend, and to offer ways to respond and correct) will inevitably arise – just like our other institutions of oversight.”

**James Morris**, professor of computer science at Carnegie Mellon, wrote, “I had to say ‘yes.’ The hope is that engineers get control away from capitalists and rebuild technology to embody a new ‘constitution.’ I actually think that’s a longshot in the current atmosphere. Ask me after November. If the competition between the U.S. and China becomes zero-sum, we won’t be able to stop a rush toward dystopia.”

**J. Francisco Álvarez**, professor of logic and philosophy of science at UNED, the National University of Distance Education in Spain, commented, “Concerns about the use of AI and its ethical aspects will be very diverse and will produce very uneven effects between public good and a set of new, highly marketable services. We will have to expand the spheres of personal autonomy and the recognition of a new generation of rights in the digital society. It is not enough to apply ethical codes in AI devices. Instead, a new ‘constitution’ must be formulated for the digital age and its governance.”

**Aaron Chia Yuan Hung**, assistant professor of educational technology at Adelphi University, said, “The use of AI now for surveillance and criminal justice is very problematic. The AI can’t be fair if it is designed based on or drawing from the data collected from a criminal justice system that is inherently unjust. The fact that some people are having these conversations makes me think that there is positive potential. Humans are not the best at decision-making. We have implicit bias. We have cognitive biases. We are irrational (in the behavioral economics sense). AI can correct that or at least make it visible to us so that we can make better decisions. Most people are wary enough of AI systems not to blindly adopt another country’s AI system without a lot of scrutiny. Hopefully that allows us to remain vigilant.”

**Moira de Roche**, chair of the International Federation of Information Processing’s professional practice sector, wrote, “There is a trend toward ethics, especially in AI applications. AI will continue to improve people’s lives in ways we cannot even anticipate presently. Pretty much every technology we use on a day-to-day basis employs AI (email, mobile phones, etc.). In fact, it worries me that AI is seen as something new, whereas we have used it on a daily basis for a decade or more. Perhaps the conversation should be more about robotics and automation than AI, per se. I

am concerned that there are so many codes of ethics. There should not be so many (at present there are several). I worry that individuals will choose the code they like the best – which is why a plethora of codes is dangerous.”

**Nigel Cameron**, president emeritus at the Center for Policy on Emerging Technologies, commented, “The social and economic shifts catalyzed by the COVID plague are going to bring increasing focus to our dependence on digital technologies, and with that focus will likely come pressure for algorithmic transparency and concerns over equity and so forth. Antitrust issues are highly relevant, as is the current pushback against China and, in particular, Huawei (generally, I think a helpful response).”

**Peter B. Reiner**, professor of neuroethics at the University of British Columbia, said, “As AI-driven applications become ever more entwined in our daily lives, there will be substantial demand from the public for what might be termed ‘ethical AI.’ Precisely how that will play out is unknown, but it seems unlikely that the present business model of surveillance capitalism will hold, at least not to the degree that it does today. I expect that clever entrepreneurs will recognize opportunities and develop new, disruptive business models that can be marketed both for the utility of the underlying AI and the ethics that everyone wishes to see put into place. An alternative is that a new regulatory regime emerges, constraining AI service providers and mandating ethical practice.”

**Ronnie Lowenstein**, a pioneer in interactive technologies, noted, “AI and the related integration of technologies holds the potential of altering lives in profound ways. I fear the worse but have hopes. Two things that bring me hope:

- 1) Increased civic engagement of youth all over the world – not only do I see youth as hope for the future, but seeing more people listening to youth encourages me that the adults are re-examining their beliefs and assumptions so necessary for designing transformative policies and practices.
- 2) The growth futures/foresight strategies as fostered by The Millennium Project.”

**Peter Dambier**, a longtime Internet Engineering Task Force participant based in Germany, said, “Personal AI must be as personal as your underwear. No spying, no malware. AI will develop like humans and should have rights like humans. I do not continue visiting a doctor I do not trust. I do not allow anything or anybody I do not trust to touch my computer. Anything that is not open-source I do not trust. Most people should learn informatics and have one person in the family who understands computers.”

**Ray Schroeder**, associate vice chancellor of online learning, University of Illinois-Springfield, responded, “One of the aspects of this topic that gives me the most hope is that, while there is the possibility of unethical use of AI, the technology of AI can also be used to uncover those unethical applications. That is, we can use AI to help patrol unethical AI. I see that artificial intelligence will be able to bridge communications across languages and cultures. I see that AI will enable us to provide enhanced telemedicine and agricultural planning. I see that AI will enable us to more clearly predict vulnerabilities and natural disasters so that we can intervene before people are hurt. I am most excited about quantum computing supercharging AI to provide awesome performance in solving our world’s problems. I am further excited about the potential for AI networking to enable us to work across borders to benefit more of the world’s citizens.”

**Melissa R. Michelson**, a professor of political science at Menlo College, commented, “Because of the concurrent rise of support for the Black Lives Matter movement, I see people taking a second look at the role of AI in our daily lives, as exemplified by the decision to stop police use of facial recognition technology. I am optimistic that our newfound appreciation of racism and discrimination will continue to impact decisions about when and how to implement AI.”

**Andrew K. Koch**, president and chief operating officer at the John N. Gardner Institute for Excellence in Undergraduate Education, wrote, “If there was a ‘Yes, but’ option, I would have selected it. I am an optimist. But I am also a realist. AI is moving quickly. Self-interested (defined in individual and corporate ways) entities are exploiting AI in dubious and unethical ways now. They will do so in the future. But I also believe that national and global ethical standards will continue to develop and adapt. The main challenge is the pace of evolution for these standards. AI may have to be used to help keep up with adaptation needed for the ethical standards needed for AI systems.”

**Anne Collier**, editor of Net Family News and founder of The Net Safety Collaborative, responded, “Policymakers, newsmakers, users and consumers will exert and feel the pressure for ethics with regard to tech and policy because of three things:

- 1) A blend of the internet and a pandemic has gotten us all thinking as a planet more than ever.
- 2) The disruption COVID-19 introduced to business- and governance-as-usual.
- 3) Because of the growing activism and power of youth seeking environmental ethics and social justice.

Populism and authoritarianism in a number of countries certainly threaten that trajectory, but – though seemingly on the rise now – I don’t see this as a long-term threat (a sense of optimism that comes from watching the work of so-called ‘Gen Z’). I wish, for example, that someone could survey a representative sample of Gen Z citizens of the Philippines, Turkey, Brazil, China,

Venezuela, Iran and the U.S. and ask *them* this question, explaining how AI could affect their everyday lives, then publish *that* study. I believe it would give many other adults a sense of optimism similar to mine.”

**Eric Knorr**, pioneering technology journalist and editor in chief of International Data Group, the publisher of a number of leading technology journals, commented, “First, only a tiny slice of AI touches ethics – it’s primarily an automation tool to relieve humans of performing rote tasks. Current awareness of ethical issues offers hope that AI will either be adjusted to compensate for potential bias or sequestered from ethical judgment.”

**Anthony Clayton**, an expert policy analysis, futures studies and scenario and strategic planning based at the University of the West Indies, said, “Technology firms will come under increasing regulatory pressure to introduce standards (with regard to, e.g., ethical use, error-checking and monitoring) for the use of algorithms when dealing with sensitive data. AI will also enable, e.g., autonomous lethal weapons systems, so it will be important to develop ethical and legal frameworks to define acceptable use.”

**Fabrice Popineau**, an expert on AI, computer intelligence and knowledge engineering based in France, responded, “I have hope that AI will follow the same path as other potential harmful technologies before (nuclear, bacteriological); safety mechanisms will be put in motion to guarantee that AI use stays beneficial.”

**Concepcion Olavarrieta**, foresight and economic consultant and president of the Mexico node of The Millennium Project, wrote, “Yes, there will be progress:

- 1) Ethical issues are involved in most human activities.
- 2) The pandemic experience plays into this development.
- 3) Societal risk factors will not be attended.
- 4) AI will become core in most people’s lives by 2030.
- 5) It is important to assure an income and or offer a basic income to people.”

**Sharon Sputz**, executive director of strategic programs at The Data Science Institute at Columbia University, predicted, “In the distant future, ethical systems will prevail, but it will take time.”

**A well-known cybernetician and emeritus professor of business management** commented, “AI will be used to help people who can afford to build and use AI systems. Lawsuits

will help to persuade companies what changes are needed. Companies will learn to become sensitive to AI-related issues.”

### A consensus around ethical AI is emerging, and open-source solutions can help

A portion of these experts optimistically say they expect progress toward ethical AI systems. They say there has been intense and widespread activity across all aspects of science and technology development on this front for years, and it is bearing fruit. Some point out that the field of bioethics has already managed to broadly embrace the concepts of beneficence, nonmaleficence, autonomy and justice in its work to encourage and support positive biotech evolution that serves the common good.

Some of these experts expect to see an expansion of the type of ethical leadership already being demonstrated by open-source AI developers, a cohort of highly principled AI builders who take the view that it should be thoughtfully created in a fairly transparent manner and be sustained and innovated in ways that serve the public well and avoid doing harm. They are hopeful that there is enough energy and brainpower in this cohort to set the good examples that can help steer positive AI evolution across all applications.

Also of note: Over the past few years, tech industry, government and citizen participants have been enlisted to gather in many different and diverse working groups on ethical AI; while some experts in this canvassing see this to mostly be public relations window-dressing, others believe that these efforts will be effective.

**Micah Altman**, a social and information scientist at MIT, said, “First, the good news: In the last several years, dozens of major reports and policy statements have been published by stakeholders from across all sectors arguing that the need for ethical design of AI is urgent and articulating general ethical principles that should guide such design. Moreover, despite significant differences in the recommendations of these reports, most share a focused common core of ethical principles. This is progress. And there are many challenges to meaningfully incorporating these principles into AI systems; into the processes and methods that would be needed to design, evaluate and audit ethical AI systems; and into the law, economics and culture of society that is needed to drive ethical design.

“We do not yet know (generally) how to build ethical decision-making into AI systems directly; but we could and should take steps toward evaluating and holding organizations accountable for AI-based decisions. And this is more difficult than the work of articulating these principles. It will be a long journey.”



**Henry E. Brady**, dean of the Goldman School of Public Policy at the University of California-Berkeley, responded, “There seems to be a growing movement to examine these issues, so I am hopeful that by 2030 most algorithms will be assessed in terms of ethical principles. The problem, of course, is that we know that, in the case of medical experiments, it is a long time from the infamous [Tuskegee study](#) to committees for the protection of human subjects. But I think the emergence of AI has actually helped to make clear the inequities and injustices in some of our practices. Consequently, they provide a useful focal point for democratic discussion and action.

“I think public agencies will take these issues very seriously, and mechanisms will be created to improve AI (although the issues pose difficult problems for legislators due to [the] highly technical nature). I am more worried about private companies and their use of algorithms. It is important, by the way, to recognize that a great deal of AI (perhaps all of it) is simply the application of ‘super-charged’ statistical methods that have been known for quite a long time.

“It is also worth remembering that AI is very good at predictions given a fixed and unchanging set of circumstances, but it is not good at causal inference, and its predictions are often based upon proxies for an outcome that may be questionable or unethical.

“Finally, AI uses training sets that often embed practices that should be questioned. A lot of issues in AI concern me. The possibility of ‘deepfakes’ means that reality may become protean and shape-shifting in ways that will be hard to cope with. Facial recognition provides for the possibility of tracking people that has enormous privacy implications. Algorithms that use proxies and past practice can embed unethical and unfair results. One of the problems with some multilayer AI methods is that it is hard to understand what rules or principles they are using. Hence, it is hard to open up the ‘black box’ and see what is inside.”

**J. Nathan Matias**, an assistant professor at Cornell University and expert in digital governance and behavior change in groups and networks, noted, “Unless there is a widespread effort to halt their deployment, artificial intelligence systems will become a basic part of how people and institutions make decisions. By 2030, a well-understood set of ethical guidelines and compliance checks will be adopted by the technology industry. These compliance checks will assuage critics but will not challenge the underlying questions of conflicting values that many societies will be unable to agree on. By 2030, computer scientists will have made great strides in attempts to engineer fairer, more equitable algorithmic decision-making.

“Attempts to deploy these systems in the field will face legal and policy attacks from multiple constituencies for constituting a form of discrimination. By 2030, scientists will have an early answer to the question of whether it is possible to make general predictions about the behavior of algorithms in society.

“If the behavior and social impacts of artificial intelligence can be predicted and modeled, then it may become possible to reliably govern the power of such systems. If the behavior of AI systems in society cannot be reliably predicted, then the challenge of governing AI will continue to remain a large risk of unknown dimensions.”

**Jean Paul Nkurunziza**, secretary-general of the Burundi Youth Training Centre, wrote, “The use of AI is still at its infancy. The ethical aspects of that domain are not yet clear. I believe that around 2025 ethical issues about the use of AI may erupt (privacy, the use of the AI in violence such as war and order keeping by police, for instance). I foresee that issues caused by the primary use of AI will bring the community to debate about that, and we will come up with some ethical guidelines around AI by 2030.”

**Doris Marie Provine**, emeritus professor of justice and social inquiry at Arizona State University, noted, “I am encouraged by the attention that ethical responsibilities are getting. I expect that attention to translate into action. The critical discussion around facial-recognition technology gives me hope. AI can make some tasks easier, e.g., sending a warning signal about a medical condition. But it also makes people lazier, which may be even more dangerous. At a global level, I worry about AI being used as the next phase of cyber warfare, e.g., to mess up public utilities.”

**Judith Schoßböck**, research fellow at Danube University Krems, said, “I don’t believe that most AI systems will be used in ethical ways. Governments would have to make this a standard, but, due to the pandemic and economic crisis, they might have other priorities. Implementation and making guidelines mandatory will be important. The most difference will be felt in the area of bureaucracy. I am excited about AI’s prospects for assisted living.”

**Gary L. Kreps**, distinguished professor of communication and director of the Center for Health and Risk Communication at George Mason University, responded, “I am guardedly optimistic that ethical guidelines will be used to govern the use of AI in the future. Increased attention to issues of privacy, autonomy and justice in digital activities and services should lead to safeguards and regulations concerning ethical use of AI.”

**Michael Marien**, director of Global Foresight Books, futurist and compiler of the annual list of the best futures books of the year, said, “We have too many crises right now, and many more ahead, where technology can only play a secondary role at best. Technology should be aligned with the UN’s 17 Sustainable Development Goals and especially concerned about reducing the widening inequality gap (SDG #10), e.g., in producing and distributing nutritious food (SDG #2).”

**Ibon Zugasti**, futurist, strategist and director with Prospektiker, wrote, “The use of technologies, ethics and privacy must be guaranteed through transparency. What data will be used and what will be shared? There is a need to define a new governance system for the transition from current narrow AI to the future general AI.”

**Gerry Ellis**, an accessibility and usability consultant, said, “The concepts of fairness and bias are key to ensure that AI supports the needs of all of society, particularly those who are vulnerable, such as many (but not all) persons with disabilities. Overall, AI and associated technologies will be for the good, but individual organizations often do not look beyond their own circumstances and their own profits. One does not need to look beyond sweatshops, dangerous working conditions and poor wages in some industries to demonstrate this. Society and legislation must keep up with technological developments to ensure that the good of society is at the heart of society, and that is in its industrial practices.”

### Ethics will evolve and progress will come as different fields show the way

A number of these experts insisted that no technology endures if it broadly delivers unfair or unwanted outcomes. They said technologies that cause harms are adjusted or replaced as, over time, people recognize and work to overcome difficulties to deliver better results. Others said ethics will come to rule at least some aspects of AI but it will perhaps not gain ground until regulatory constraints or other incentives for tech businesses emerge.

A number of respondents made the case that the application of ethics to AI applications will likely unfold in different ways, depending upon the industry or public arena in which they unfold. Some say this rollout will depend upon the nature of the data involved. For instance, elaborate ethics regimes have already been developed around the use of health and medical data. Other areas such as the arguments over privacy and surveillance systems have been more contested.

**Jon Lebkowsky**, CEO, founder and digital strategist at Polycot Associates, wrote, “I have learned from exposure to strategic foresight thinking and projects that we can’t predict the future, but we can determine scenarios that we want to see and work to make those happen. So, I am not predicting that we’ll have ethical AI so much as stating an aspiration – it’s what I would work toward. Certainly, there will be ways to abuse AI/ML/big data, especially in tracking and surveillance. Globally, we need to start thinking about what we think the ethical implications will be and how we can address those within technology development. Given the current state of global politics, it’s harder to see an opportunity for global cooperation, but hopefully the pendulum will swing back to a more reasonable global political framework. The ‘[AI for Good](#)’ gatherings might be helpful if they continue. AI can be great for big data analysis and data-driven action, especially where data discretion can be programmed into systems via machine-learning algorithms. Some of

the more interesting applications will be in translation, transportation, including aviation; finance, government (including decision support), medicine and journalism.

“I worry most about uses of AI for surveillance and political control, and I’m a little concerned about genetic applications that might have unintended consequences, maybe because I saw a lot of giant bug movies in the 1950s. I think AI can facilitate better management and understanding of complexity and greater use of knowledge- and decision-support systems. Evolving use of AI for transportation services has been getting a lot of attention and may be the key to overcoming transportation inefficiency and congestion.”

**Amar Ashar**, assistant director of research at the Berkman Klein Center for Internet & Society, said, “We are currently in a phase where companies, countries and other groups who have produced high-level AI principles are looking to implement them in practice. This application into specific real-world domains and challenges will play out in different ways. Some AI-based systems may adhere to certain principles in a general sense, since many of the terms used in principles documents are broad and defined differently by different actors. But whether these principles meet those definitions or the spirit of how these principles are currently being articulated is still an open question.

“Implementation of AI principles cannot be left to AI designers and developers alone. The principles often require technical, social, legal, communication and policy systems to work in coordination with one another. If implemented without accountability mechanisms, these principles statements are also bound to fail.”

**A journalism professor emeritus** predicted, “The initial applications, ones easily accepted in society, will be in areas where the public benefit is manifestly apparent. These would include health and medicine, energy management, complex manufacturing and quality control applications. All good and easy to adhere to ethical standards, because they’re either directly helpful to an individual or they make things less expensive and more reliable. But that won’t be the end of it. Unless there are both ethical and legal constraints with real teeth, we’ll find all manner of exploitations in finance, insurance, investing, employment, personal data harvesting, surveillance and dynamic pricing of almost everything from a head of lettuce to real estate. And those who control the AI will always have the advantage – always.

“What most excites me beyond applications in health and medicine are applications in materials science, engineering, energy and resource management systems and education. The ability to deploy AI as tutors and learning coaches could be transformative for equalizing opportunities for educational attainment. I am concerned about using AI to write news stories unless the ‘news’ is a sports score, weather report or some other description of data ... My greatest fear, not likely in my

lifetime, is that AI eventually is deployed as our minders – telling us when to get up, what to eat, when to sleep, how much and how to exercise, how to spend our time and money, where to vacation, who to socialize with, what to watch or read and then secretly rates us for employers or others wanting to size us up.”

**Michael R. Nelson**, research associate at CSC Leading Edge Forum, observed, “What gives me hope: Companies providing machine learning and big data services so all companies and governments can apply these tools. Misguided efforts to make technology ‘ethical by design’ worry me. Cybersecurity making infrastructure work better and more safely is an exciting machine-learning application, and ‘citizen science’ and sousveillance knowledge tools that help me make sense of the flood of data we swim in.”

**Edward A. Friedman**, professor emeritus of technology management at Stevens Institute of Technology, responded, “AI will greatly improve medical diagnostics for all people. AI will provide individualized instruction for all people. I see these as ethically neutral applications.”

**Lee McKnight**, associate professor at the Syracuse University School of Information Studies, wrote, “When we say ‘AI,’ most people really mean a wider range of systems and applications, including machine learning, neural networks and natural language processing to name a few. ‘Artificial general intelligence’ remains the province through 2030 of science fiction and Elon Musk.

“A wide array of ‘industrial AI’ will in 2023, for example, help accelerate or slow down planes, trains and rocket ships. Most of those industrial applications of AI will be designed by firms, and the exact algorithms used and adapted will be considered proprietary trade secrets – not subject to public review or ethics audit. I am hopeful that smart cities and communities initially and eventually all levels of public organizations, and nonprofits – will write into their procurement contracts requirements that firms not only commit to an ethical review process for AI applications touching on people directly – such as facial recognition. Further, I expect communities will in their requests for proposals make clear that inability to explain how an algorithm is being used and where the data generated is going/who will control the information will be disqualifying.

“These steps will be needed to restore communities’ trust in smart systems, which was shaken by self-serving initiatives by some of the technology giants trying to turn real communities into company towns. I am excited to see this clear need and also the complexity of developing standards and curricula for ‘certified ethical AI developers,’ which will be a growth area worldwide. How exactly to determine if one is truly ‘certified’ in ethics is obviously an area where the public would laugh in the faces of corporate representatives claiming their internal, not publicly disclosed, or audited, ethical training efforts are sufficient. This will take years to sort out

and will require wide public dialogue and new international organizations to emerge. I am excited to help in this effort where I can.”

**A professor at university in the U.S. Midwest** said, “I anticipate some ethical principles for AI will be adopted by 2030; however, they will not be strong or transparent. Bottom line: Capitalism incentivizes exploitation of resources, and the development of AI and its exploitation of information is no different than any other industry. AI has great potential, but we need to better differentiate its uses. It can help us understand disease and how to treat it but has already inflicted great harms on individuals. As we have seen, AI has also disproportionately impacted those already marginalized – the [COMPAS recidivism algorithm](#) and the use of facial-recognition technology by police agencies are examples. The concept of general and narrow AI that Meredith Broussard uses is appropriate. Applied in particular areas, AI is hugely important and will better the lives of most. Other applications are nefarious and should be carefully implemented.”

**Mark Monchek**, author, keynote speaker and self-described “culture of opportunity strategist,” commented, “In order for ethical principles to prevail, we need to embrace the idea of citizenship. By ‘citizenship,’ I mean a core value that each of us, our families and communities have a responsibility to actively participate in the world that affects us. This means carefully ‘voting’ every day when choosing who we buy from, consume technology from, work for and live with, etc. We would need to be much more proactive in our use of technology, including privacy issues, understanding, consuming media more like we consume food, etc.”

**Monica Murero**, director, E-Life International Institute and associate professor in Communication and New Technologies at the University of Naples Federico II, asked, “Will there be ethical or questionable outcomes? In the next decade (2020-2030), I see both, but I expect AI to become more questionable. I think about AI as an ‘umbrella’ term with different technologies, techniques, and applications that may lead to pretty different scenarios. The real challenge to consider is how AI will be used in combination with other disruptive technologies such as Internet of Things, 3D printing, cloud computing, blockchain, genomics engineering, implantable devices, new materials and environment-friendly technologies, new ways to store energy and how the environment and people will be affected and at the same part of the change – physically and mentally for the human race. I am worried about the changes in ‘humans’ and the rise of new inequalities in addition to the effects on objects and content that will be around us. The question is much broader than ‘ethical,’ and the answers, as a society, should start in a public debate at the international level. We should decide who or what should benefit the most. Many countries and companies are still very behind this race, and others will take advantage of it. This worries me the most because I do not expect that things will evolve in a transparent and ‘ethical’ manner. I am very much in favor of creating systems of evaluation and regulation that seriously look at the outcomes over time.”

### 3. Cross-cutting and novel statements

A number of the respondents wrote about cross-cutting themes, introduced novel ideas or shared thoughts that were not widely mentioned by others. This section features a selection of that material.

*Questions about AI figure into the grandest challenges humans face*

**Michael G. Dyer**, professor emeritus of computer science at UCLA, expert in natural language processing, argued, “The greatest scientific questions are:

- 1) Nature of Matter/Energy
- 2) Nature of Life
- 3) Nature of Mind

“Developing technology in each of these areas brings about great progress but also existential threats. Nature of Matter/Energy: progress in lasers, computers, materials, etc., but hydrogen bombs with missile delivery systems. Nature of Life: progress in genetics, neuroscience, health care, etc., but possibility of man-made deadly artificial viruses. Nature of Mind: intelligence software to perform tasks in many areas but possibility of the creation of a general-AI that could eliminate and replace humankind.

“We can’t stop our exploration into these three areas, because then others will continue without us. The world is running on open, and the best we can do is to try to establish fair, democratic and noncorrupt governments. Hopefully in the U.S., government corruption, which is currently at the highest levels (with nepotism, narcissism, alternate ‘facts,’ racism, etc.), will see a new direction in 2021.”

*Was the internet mostly used in ethical or questionable ways the past decade?*

**Seth Finkelstein**, programmer, consultant and EFF Pioneer of the Electronic Frontier Award winner, noted, “Just substitute ‘the internet’ for ‘AI’ here – ‘Was the internet mostly used in ethical or questionable ways in the last decade?’ It was/will be used in many ways, and the net result ends up with both good and bad, according to various social forces. I believe technological advances are positive overall, but that shouldn’t be used to ignore and dismiss dealing with associated negative effects. There’s an AI ‘moral panic’ percolating now, as always happens with new technologies. A little while ago, there was a fear-mongering fad about theoretical ‘[trolley problems](#)’ (choosing actions in a car accident scenario). This was largely written about by people who apparently had no interest in the extensive topic of engineering safety trade-offs.

“Since discussion of, for example, structural racism or sexism pervading society is more a humanities field of study than a technological one, there’s been a somewhat better grasp by many writers that the development of AI isn’t going to take place outside existing social structures there.

“As always, follow the money. Take the old aphorism ‘It is difficult to get a man to understand something when his salary depends upon his not understanding it.’ We can adapt it to ‘It is difficult to get an AI to understand something when the developer’s salary depends upon the AI not understanding it.’

“Is there going to be a fortune in funding AI that can make connections between different academic papers or an AI that can make impulse purchases more likely? Will an AI assistant tell you that you’re spending too much time on social media and you should cut down for your mental health (‘log off now, okay?’) or that there’s new controversy brewing and get clicking otherwise you may be missing out (‘read this bleat, okay?’)?”

*The future of work is a central issue in this debate*

**Jonathan Taplin**, author of “[Move Fast and Break Things: How Google, Facebook and Amazon Cornered Culture and Undermined Democracy](#),” observed, “Even if most major players in AI will abide by ethical rules, the role of bad actors using AI can have outsized effects on society. The ability to use deepfakes to influence political outcomes will be tested.

“What worries me the most is that the substitution of AI (and robotics) for human work will accelerate post-COVID-19. The political class, with the notable exception of Andrew Yang, is in total denial about this. And the substitution will affect radiologists just as much as meat cutters. The job losses will cut across classes.”

*Intellectual product is insufficient to protect us from dystopic outcomes*

**Frank Kaufmann**, president of the Twelve Gates Foundation, noted, “Will AI mostly be used in ethical or questionable ways in the next decade? And why? This is a complete and utter toss-up. I believe there is *no* way to predict which will be the case.

“It is a great relief that, in recent years, there have been scores of convenings and even more papers generated proposing ethical frameworks for the application of artificial intelligence. They cover a host of issues, including transparency, justice and fairness, privacy, freedom and human autonomy, beneficence and nonmaleficence, freedom, trust, sustainability and dignity. But, then again, there have been the Treaty of Versailles, and the literal tons of paper the United Nations has produced talking about peace, the Declaration of Human Rights and so forth.



“I am glad people meet sincerely and in earnest to examine vital ethical concerns related to the development of AI. The problem is that intellectual product is insufficient to protect us from dystopic outcomes. The hope and opportunity to enhance, support and grow human freedom, dignity, creativity and compassion through AI systems excite me. The chance to enslave, oppress and exploit human beings through AI systems concerns me.”

*Technological determinism should be ‘challenged by critical research’*

**Bill Dutton**, professor of media and information policy at Michigan State University, said, “AI is not new and has generally been supportive of the public good, such as in supporting online search engines. The fact that many people are discovering AI as some new development during a dystopian period of digital discourse has fostered a narrative about evil corporations challenged by ethical principles. This technologically deterministic good versus evil narrative needs to be challenged by critical research.”

*Is it possible to design cross-cultural ethics systems?*

**Michael Muller**, a researcher for a top global technology company focused on human aspects of data science and ethics and values in applications of artificial intelligence, wrote, “I am concerned about what might be called ‘traveling AI’ – i.e., AI solutions that cross cultural boundaries.

“Most AI systems are likely to be designed and developed in the individualistic EuroWestern cultures. These systems may be ill-suited – and in fact harmful – to collectivist cultures. The risk is particularly severe for indigenous cultures in, e.g., the Americas, Africa and Australia.

“How can we design systems that are ethical in the cultural worlds of their users – whose ethics are based on very different values from the individualistic EuroWestern traditions?”

*Most people have no idea how limited and brittle these capabilities are*

**Steven Miller**, professor emeritus of information systems at Singapore Management University, responded, “We have to move beyond the current mindset of AI being this special thing – an almost mystical thing. I wish we would stop using the term AI (though I use it a lot myself), and just refer to it for what it is – pattern-recognition systems, statistical analysis systems that learn from data, logical reasoning data, goal-seeking systems. Just look at the table of contents for an AI textbook (such as ‘[Artificial Intelligence: A Modern Approach](#),’ Stuart Russell and Peter Norvig, 4th edition, published 2020). Each item in the table of contents is a subarea of AI, and there are a lot of subareas. ...

“There are ethical issues associated with *any* deployment of *any* engineering and technology system, any automation system, any science effort (especially the application of the science), and/or any policy analysis effort. So, there is nothing special about the fact that we are going to have ethical issues associated with the use of AI-enabled systems. As soon as we stop thinking of AI as ‘special,’ and to some extent magical (at least to the layman who does not understand how these things work, as machines and tools), and start looking at each of these applications, and families of applications as deployments of tools and machines – covering both physical realms of automation and/or augmentation, and cognitive and decision-making realms of automation and/or augmentation – then we can have real discussions.

“Years back, invariably, there had to have been many questions raised about ‘the ethics of using computers,’ especially in the 1950s, 1960s and 1970s, when our civilisation was still experiencing the possibilities of computerising many tasks for the very first time. AI is an extension of this, though taking us into a much wider range of tasks and tasks of increasing cognitive sophistication. ...

“Now, of course, our ability to create machines that can sense, predict, respond and adapt has vastly improved. Even so, most lay people have no idea of just how limited and brittle these capabilities are – even though they are remarkable and far above human capability in certain specific subdomains, under certain circumstances. What is happening is that most laypeople are jumping to the conclusion that, ‘Because it is an AI-based system, it must be right, and therefore, I should not question the output of the machine, for I am just a mere human.’ So now the pendulum has swung to the other extreme of layperson assuming AI-enabled algorithms and machines are actually more capable (or more robust and more context-aware) than they actually are. And this will lead to accidents, mistakes and problems. ...

“Just like there will be all types of people with all types of motives pursuing their interests in all realms of human activity, the same will be true of people making use of AI-enabled systems for automation, or augmentation or related human support. And some of these people will have noble goals and want to help others. And some of these people will be nefarious and want to gain advantage in ways others might not understand, and there will even be the extreme of some who purposely want to bring harm to others. We saw this with social media. In years, decades and centuries past, we saw this with every technological innovation that appeared, going back to the printing press and even earlier. ... Let’s start getting specific about use cases and situations. One cannot talk in the abstract as to whether an automobile will be used ethically. Or whether a computer will be used ethically. Or biology as an entire field will be used ethically. One has to get much more specific about classes of issues or types of problems that are related to the usage of these big categories of ‘things.’”

*AI ‘must not be used to make any decision that has direct impact on people’s lives’*

**Fernando Barrio**, a lecturer in business law at Queen Mary University of London expert in AI and human rights, responded, “If ethical codes for AI will be in place for the majority of cases for 2030, they will purport to be in the public good (which would seem to imply a ‘yes’ to the question as it was asked), but they will not result in public good. The problem is that the question assumes that the sole existence and use of ethical codes would be in the public good.

“AI, not as [the singularity](#) but as machine learning or even deep learning, has an array of positive potential applications but must not be used to make any decision that has a direct impact on people’s lives. In certain sectors, like the criminal system, it must not be used even in case management, since the inherent and unavoidable bias (either from the data or the algorithmic bias resulting from its own selection or discovery of patterns) leads that individuals and their cases are not judged or managed, taking into account the characteristics that make every human unique but in those characteristics that make that person ‘computable.’

“Those who propose the use of AI to avoid human bias, such as those that judges and jurors might implement, tend to overlook – let’s assume naively – that those biases can be challenged through appeals, and they can be made explicit and transparent. The AI inherent, one cannot be challenged because, between other things, the lack of transparency and, especially, for the insistence of its proponents that the technology can be unbiased.”

*Just as with drugs, without rigorous study prior to release, AI side effects can be dangerous*

**An internet pioneer and principal architect at a major technology company** said, “AI is a lot like new drug development – without rigorous studies and regulations, there will always be the potential for unexpected side effects.

“Bias is an inherent risk in any AI system that can have major effects on people’s lives. While there is more of an understanding of the challenges of ethical AI, implicit bias is very difficult to avoid because it is hard to detect. For example, you may not discover that a facial-recognition system has excessively high false-recognition rates with some racial or ethnic groups until it has been released – the data to test all the potential problems may not have been available before the product is released.

“The alternative is to move to a drug-development model for AI, where very extensive trials with increasingly large populations are required prior to release, with government agencies monitoring progress at each stage. I don’t see that happening, though, because it will slow innovation, and tech companies will make the campaign contributions necessary to prevent regulation from becoming that intrusive.”

*Has AI has been shifting the nature of thought and discourse and, if so, how?*

**A professor of urban planning** noted, “Already the general ‘co-evolution’ of humanity and technology suggests that humans are not nearly as in control as they think they are of technology’s operations, much less its trajectory. While I am not an enthusiast of singularity speculations, there does seem to be a move toward AI stepping in to save the planet, and humans remain useful to that for a while, maybe in bright perpetuity.

“With wondrous exceptions of course, humans themselves seem ever less inclined to dwell on the question of what is good, generally, or more specifically, what is good about mindful reflectivity in the face of rampant distraction engineering.

“While one could worry that humans will unleash AI problems simply because it would be technically possible, perhaps the greater worry is that AI, and lesser technological projects, too, have already been shifting the nature of thought and discourse toward conditions where cultural deliberations on more timeless and perennial questions of philosophy have no place. Google is already better at answers. Humans had best cultivate their advantage at questions. But if you are just asking about everyday AI assistance, just look how much AI underlies the autocomplete to a simple search query. Or, gosh, watch the speed and agility of the snippets and IntelliSense amid the keyboard experience of coding. Too bad.”

*Can any type of honor code – ‘AI omerta’ – really keep developers in line?*

**Anthony Judge**, editor of the *Encyclopedia of World Problems and Human Potential*, observed, “The interesting issue for me is how one could navigate either conclusion to the questions and thereby subvert any intention.

“We can’t assume that the ‘bad guys’ will not be developing AI assiduously to their own ends (as could already be argued to be the case), according to their own standards of ethics. AI omerta? Appropriate retribution for failing to remain loyal to the family? Eradicate those who oppose the mainstream consensus? What is to check against these processes? What will the hackers do?”

*When there is no trust in others, people focus on self-interests to the detriment of others*

**Rebecca Theobald**, assistant research professor at the University of Colorado-Colorado Springs, predicted, “AI will mostly be used in questionable ways in the next decade because people do not trust the motives of others. Articulate people willing to speak up give me the most hope. People who are scared about their and their families’ well-being worry me the most because they feel there is no other choice but to scramble to support themselves and their dependents.

“Without some confidence in the climate, economy, health system and societal interaction processes, people will become focused on their own issues and have less time and capital to focus on others. AI applications in health and transportation will make a difference in the lives of most people. Although the world is no longer playing as many geopolitical games over territory, corporations and governments still seek power and influence. AI will play a large role in that. Still, over time, science will win out over ignorance.”

*In order to assure it is for public good, perhaps AI could be regulated like a utility*

**A director with a strategy firm** commented, “The creators of AI and AI in general are most likely to be used by those in power to keep power. Whether to wage war or financial war or manage predicted outcomes, most AIs are there to do complex tasks. Unless there is some mechanism to make them for the public benefit, they will further encourage winner-take-all.

“Regarding lives, let’s take the Tesla example. Its claim is that [it will soon have Level 5 AI in its vehicles](#). Let’s assume that it takes a couple of years beyond that. The markets are already betting that: 1) It will happen, and 2) No one else is in any position to follow. If so, rapid scaling of production would enable fleets of robocall-taxis; it could destroy the current car industry as the costs are radically lower, and the same tech will impact on most public transport, too, in five years.

“Technology-wise, I love the above scenario. It does mean, however, that only the elite will drive or have a desire to have their own vehicle. Thus, for the majority, this is a utility. Utilities are traditionally for the public good. It’s why in most countries the telephone system or the postal system were originally owned by the government. It’s why public transport is a local government service. We will not be well served by a winner-take-all transportation play! Amazon seems to be doing pretty well with AI. They can predict your purchases. They can see their resellers success and at scale simply replace them. Their delivery network is at scale and expected to also go to autonomous. I can’t live without it; however, each purchase kills another small supplier. Because economics eliminate choice – one has to feed oneself.

“As long as AI can be owned, those who have it or access to it have an advantage. Those who don’t are going to suffer and be disadvantaged.”

*Three respondents’ views:*

*The biggest concerns involve ill-considered AI systems, large and small*

**Joshua Hatch**, a journalist who covers technology issues, commented, “While I think most AI will be used ethically, that’s probably irrelevant. This strikes me as an issue where it’s not so much about what ‘most’ AI applications do but about the behavior of even just a *few* applications. It just takes one Facebook to cause misinformation nightmares, even if ‘most’ social networks do a better

job with misinformation (not saying they do; just providing an example that it only takes one bad actor). Furthermore, even ethical uses can have problematic outcomes. You can already see this in algorithms that help judges determine sentences. A flawed algorithm leads to flawed outcomes – even if the intent behind the system was pure. So, you can count on misuse or problematic AI just as you can with any new technology. And even if most uses are benign, the scale of problem AI could quickly create a calamity. That said, probably the best potential for AI is for use in medical situations to help doctors diagnose illnesses and possibly develop treatments. What concerns me the most is the use of AI for policing and spying.”

**A research scientist who works at Google** commented, “I’m involved in some AI work, and I know that we will do the right thing. It will be tedious, expensive and difficult, but we’ll do the right thing. The *problem* will be that it’s very cheap and easy for a small company to *not* do the right thing (see the recent example of [ClearView](#), which scraped billions of facial images, violating terms of service and created a global facial-recognition dataset). This kind of thing will continue. Large companies have incentives to do the right thing, but smaller ones do not (see, e.g., Martin Shkreli and his abuse of pharma patents).”

**A research scientist working on AI innovation with Google** commented, “There will be a mix. It won’t be wholly questionable or ethical. Mostly, I worry about people pushing ahead on AI advancements without thinking about testing, evaluation, verification and validation of those systems. They will deploy them without requiring the types of assurance we require in other software. For global competition, I worry that U.S. tech companies and workers do not appreciate the national security implications.”

## 4. Could a quantum leap someday aid ethical AI?

As they considered the potential evolution of ethical AI design, the people responding to this canvassing were given the opportunity to speculate as to whether [quantum computing](#) (QC), which is still in its early days of development, might somehow be employed in the future in support of the development of ethical AI systems.

In March 2021, a team at the University of Vienna announced it had [designed a hybrid AI](#) that relies on quantum and classical computing and showed that – thanks to quantum quirks – it could simultaneously screen a handful of different ways to solve a problem. The result was a reinforcement learning AI that learned more than 60% faster than a nonquantum-enabled setup. This was one of the first tests to show that adding quantum speed can accelerate AI agent training/learning. It is projected that this capability, when scaled up, also might lead to a more-capable “quantum internet.”

Although there have been [many other announcements](#) of [new steps toward advancing QC](#) in the past year or two, it is still so nascent that even its developers are somewhat uncertain about its likely future applications, and the [search is on for specific use cases](#).

Because this tech is still in its early days, of course these respondents’ answers are speculative reflections, but their discussion of the question raises a number of important issues. The question they considered was:

*How likely is it that quantum computing will evolve over the next decade to assist in creating ethical artificial intelligence systems? If you think that will be likely, why do you think so? If you do not think it likely that quantum computing will evolve to assist in building ethical AI, why not?*

Here, we share four overarching responses. They are followed by responses from those who said it might be possible in the future for AI ethics systems to get a boost from QC and by responses from experts who said that such a development is somewhat or quite unlikely.

**Greg Sherwin**, vice president for engineering and information technology at Singularity University, wrote, “Binary computing is a lossy, reductionist crutch that models the universe along the lines of false choices. Quantum computing has an opportunity to better embrace the complexity of humanity and the world, as humans can hold paradoxes in their minds while binary computers cannot. Probabilistic algorithms and thinking will predominate the future, leading to more emphasis on the necessary tools for such scenario planning, which is where quantum computers can serve and binary computers fail. That demand for the answers that meet the

challenges of the future will require us to abandon our old, familiar tools of the past and to explore and embrace new paradigms of thinking, planning and projecting. I do see ethical AI as something orthogonal to the question of binary vs. quantum computing. It will be required in either context. So, the question of whether quantum computing will evolve as a tool to assist building ethical AI is a nonstarter. Either because there is little ‘quantum’ specialty about it, or because building ethical AI is a need independent of its computational underpinnings. Humans will continue to be in the loop for decisions that have significant impacts to our lives, our health, our governance and our social well-being. Machines will be wholly entrusted for only those things that are mechanized, routine and subject to linear optimization.”

**Barry Chudakov**, founder and principal of Sertain Research, said, “I believe quantum computers may evolve to assist in building ethical AI, not just because they can work faster than traditional computers, but because they operate differently. AI systems depend on massive amounts of data that algorithms ingest, classify and analyze using specific characteristics; quantum computers enable more precise classification of that data. Eventually, quantum computing-based AI algorithms could find patterns that are invisible to classical computers, making certain types of intractable problems solvable. But there is a fundamental structural problem that must be addressed first: vastly more powerful computing power may not resolve the human factor. Namely, that the moral and ethical framework for building societal entities (churches, governments, constitutions, laws, etc.) grew out of tribal culture, nomadic culture, which recorded precepts which then turned into codified law. ...

“We’re in a different world now. As William Gibson said in 2007: “The distinction between cyberspace and that which isn’t cyberspace is going to be unimaginable.” It’s now time to imagine the unimaginable. This is because AI operates from an entirely different playbook. The tool logic of artificial intelligence is embedded machine learning; it is quantum, random, multifarious. We are leaving the [Gutenberg Galaxy](#) and its containment patterns of rule-based injunctions. The tool logic of the book is linear, celebrates one-at-a-timeness and the single point of view; alphabetic sequentiality supplanted global/spatial awareness and fostered fear of the image; literacy deified books as holy and the ‘word of God.’ AI, on the other hand, takes datasets and ‘learns’ or improves from the analysis of that data. This is a completely different dynamic, with a different learning curve and demands. ...

“I believe we need a 21st-century Quantum AI Constitutional Convention. The purpose of such a convention is clear: To inaugurate a key issue not only for AI tech companies in the coming decade but for the known world, namely, establishing clear ethical guidelines and protocols for the deployment of AI and then creating an enlightened, equitable means of policing and enforcing those guidelines. This will necessitate addressing the complexities of sensitive contexts and environments (face recognition, policing, security, travel, etc.) as well as a host of intrusive data



collection and use case issues, such as tracking, monitoring, AI screening for employment, or algorithmic biases. This will demand transparency, both at the site of the deployment of AI as well as addressing its implications. Without those guidelines and protocols – the 21st-century equivalent of the Magna Carta and its evolved cousin, the U.S. Constitution – there will be manufactured controversy over what is ethical and what is questionable. ... AI is ubiquitous and pervasive. We hardly have the language or the inclination to fully appreciate what AI can and will do in our lives. This is not to say that we cannot; it is to say that we are unprepared to see, think, debate and wisely decide how to best move forward with AI development.

“Once there is a global constitution and Bill of AI Rights, with willing signatories around the world, quantum computing will be on track to evolve in assisting the building of ethical AI. However, the unfolding of that evolution will collide with legacy cultural and societal structures. So, as we embrace and adopt the logic of AI, we will change ourselves and our mores; effectively, we will be turning from hundreds or thousands of years of codified traditional behaviors to engage with and adapt to the ‘chaotic implications’ of AI. ...

“AI represents not human diminishment and replacement but a different way of being in the world, a different way of thinking about and responding to the world, namely, to use designed intelligence to augment and expand human intelligence. Yes, this will create new quandaries and dilemmas for us – some of which may portend great danger. ... We will braid AI into the fabric of our lives, and, in order to do so successfully, society at many levels must be present and mindful at every step of AI integration into human society.”

**David Brin**, physicist, futures thinker and author of the science fiction novels “Earth” and “Existence,” predicted, “Quantum computing has genuine potential. Roger Penrose and associates believe it already takes place, in trillions of subcellular units inside human neurons. If so, it may take a while to build quantum computers on that kind of scale. The ethical matter is interesting, though totally science fictional, that quantum computers might connect in ways that promote reciprocal understanding and empathy.”

**Jerome C. Glenn**, co-founder and CEO of the futures-research organization The Millennium Project, wrote, “Elementary quantum computing is already here and will accelerate faster than people think, but the applications will take longer to implement than people think. It will improve computer security, AI and computational sciences, which in turn accelerate scientific breakthroughs and tech applications, which in turn increase both positive and negative impacts for humanity. These potentials are too great for humanity to remain so ignorant. We are in a new arms race for artificial general intelligence and more-mature quantum computing, but, like the nuclear race that got agreements about standards and governance ([International Atomic Energy Agency](#)), we will need the same for these new technologies while the race continues.”

Responses from those who said quantum computing is very or somewhat likely to assist in working toward ethical design of artificial intelligence

**Stanley Maloy**, associate vice president for research and innovation and professor of biology at San Diego State University, responded, “Quantum computing will develop hand-in-hand with 5G technologies to provide greater access to computer applications that will affect everyone’s lives, from self-driving cars to effective drone delivery systems, and many, many other applications that require both decision-making and rapid analysis of large datasets. This technology can also be used in harmful ways, including misuse of identification technologies that bypass privacy rights.”

**A longtime network technology administrator and leader** based in Oceania, said, “Quantum computing gives us greater computational power to tackle complex problems. It is therefore a simple relationship – if more computational power is available, it will be used to tackle those complex problems that are too difficult to solve today.”

**Sean Mead**, senior director of strategy and analytics at Interbrand, said, “Quantum computing enables an exponential increase in computing power, which frees up the processing overhead so that more ethical considerations can be incorporated into AI decision-making. Quantum computing injects its own ethical dilemmas in that it makes the breaking of modern encryption trivial. Quantum computing’s existence means current techniques to protect financial information, privacy, control over network-connected appliances, etc., are no longer valid, and any security routines relying on them are likewise no longer valid and effective.”

**David Mussington**, a senior fellow at CIGI and professor and director at the Center for Public Policy and Private Enterprise at the University of Maryland, wrote, “I am guardedly optimistic that quantum computing ‘could’ develop in a salutary direction. The question is, ‘whose values will AI research reflect?’ It is not obvious to me that the libertarian ideologies of many private sector ICT and software companies will ‘naturally’ lead to the deployment of safe – let alone secure – AI tools and AI-delivered digital services. Transparency in the technologies, and in the decisions that AI may enable, may run into information-sharing limits due to trade secrets, nondisclosure agreements and international competition for dominance in cyberspace. Humans will still be in the loop of decisions, but those humans have different purposes, cultural views and – to the extent that they represent states – conflicting interests.”

**Vint Cerf**, Internet Hall of Fame member and vice president and chief internet evangelist at Google, observed, “There is some evidence that quantum methods may be applicable to ML systems for optimization for example. But it’s early days yet.”

**Jamais Cascio**, research fellow at the Institute for the Future, observed, “To the degree that quantum computing will allow for the examination of a wide variety of possible answers to a given problem, quantum computing may enhance the capacity of systems to evaluate best long-term outcomes. There’s no reason to believe that quantum computing will make ethical systems easier to create, however. And if quantum computing doesn’t allow for ready examination of multiple outcomes, then it would be no better or worse than conventional systems.”

**Gary A. Bolles**, chair for the future of work at Singularity University, responded, “We might as well ask if faster cars will allow us to go help people more quickly. Sure, but they can also deliver bad actors to their destination faster, as well. The quantum computing model lends itself to certain processes that will eventually blow past traditional microprocessors, such as completely new forms of encryption. Those methods, and the products created using them, could enable unbreakable privacy. Or they could be used to circumvent traditional approaches to encryption and create far more risk for anyone depending on traditional computing systems. As Benjamin Bratton presciently discusses in *The Stack*, if we don’t specifically create technology to help us manage the complexity of technology, that complexity alone will ensure that only a rarefied few will benefit.”

**A journalist and industry analyst expert in AI ethics** said, “Quantum is going to be made available via the cloud because of the cooling requirements. A lot of innovation has already happened, but in the next decade there will be major advancements. It will break cybersecurity as we know it today. Humans need to be in the loop. However, they will likely find themselves out of the loop unless safeguards are built into the system. AI can already do many tasks several orders of magnitude faster than humans. Quantum computing will add yet more orders of speed magnitude.”

**A professor of digital economy based in Europe** responded, “The fascination with quantum computing means that technology companies will do a lot of work on it without being too concerned about how many of these new inventions will facilitate human life. The emphasis will remain on monetizing this frontier and enabling AI that is less guided by human interventions. In effect, these technologies will be more error-prone, and as such they will unleash even more ethical concerns as they unravel through time. Its speed of calculation will be matched by glitches that will require human deliberation.”

**Ibon Zugasti**, futurist, strategist and director with Prospektiker, wrote, “Artificial intelligence will drive the development of quantum computing, and then quantum computing will further drive the development of artificial intelligence. This mutual acceleration could grow beyond human control and understanding. Scientific and technological leaders, advanced research institutes and foundations are exploring how to anticipate and manage this issue.”

**Joshua Hatch**, a journalist who covers technology issues, commented, “It seems to me that every technological advance will be put to use to solve technological dilemmas, and this is no different. As for when we’ll see dramatic advances, I would guess over the next 10 years.”

**A director of standards and strategy at a major technology company** commented, “In general, our digital future depends on advances in two very broad, very basic areas: bandwidth and computer power. Most generally, I need to be able to complete tasks, and I need to be able to move information and generally communicate with others. Quantum computing is one of the promising areas for computing power.”

**Ray Schroeder**, associate vice chancellor of online learning, University of Illinois-Springfield, responded, “The power of quantum computing will enable AI to bridge the interests of the few to serve the interests of the many. These values will become part of the AI ethos, built into the algorithms of our advanced programs. Humans will continue to be part of the partnership with the technologies as they evolve – but this will become more of an equal partnership with technology rather than humans micromanaging technology as we have in the past.”

**A complex systems researcher based in Australia** wrote, “Once AI systems can start to self-replicate, then there will be an explosive evolution. I doubt it will become the fabled singularity (where humans are no longer needed), but there will be many changes.”

**A technology developer/administrator** commented, “Quantum computing may be a more efficient way to implement a neural network. That doesn’t change the final result, though. Just as I can compile my C for any architecture, an AI algorithm may be implemented on a different hardware platform. The results will be equivalent, though hopefully faster/cheaper to execute.”

**Moira de Roche**, chair of IFIP IP3, noted, “AI systems rely on massive amounts of data. Quantum computing can help classify the data in meaningful ways. Quantum will boost machine learning.”

**A futurist and consultant** responded, “AI is about managing ever-larger datasets and machine learning. Quantum accelerates both.”

**Eric Knorr**, pioneering technology journalist and editor in chief of IDG, commented, “Yes, computing power a magnitude greater than currently available could raise the possibility of some emulation of general intelligence at some point. But how we apply that is up to us.”

**Philip M. Neches**, lead mentor at Enterprise Roundtable Accelerator and longtime trustee at California Institute of Technology, commented, “I expect cost-effective quantum computing

hardware to emerge by 2030. Programming will remain a work-in-progress for some decades after 2030.”

**Nigel Cameron**, president emeritus at the Center for Policy on Emerging Technologies, commented, “It’s hard to predict the timeline, even though it does seem inevitable that quantum systems will dominate. It’s a tantalizing idea that we can just build ethics into the algorithms. Some years back, the Department of Defense issued a strange press release in defense of robotic warfare that suggested it would be more humane, since the Geneva Conventions could be built into the programming. I’m fascinated, and horrified, by the experiences of military drone operators playing de facto video games before going home for dinner after taking out terrorists on the other side of the world. A phrase from a French officer during our high-level, super-safe (for the U.S.) bombing of Serbia comes to mind: If a cause is worth killing for, it has to be worth dying for. The susceptibility of our democracy to various forms of AI-related subversion could lead to a big backlash. I remember the general counsel of Blackberry, a former federal prosecutor, saying that ‘we have yet to have our cyber 9/11.’ When I chaired the GITEX conference in Dubai some years back, we had a huge banner on the stage that said, I think, ‘Our Digital Tomorrow.’ In my closing remarks, I suggested that, unless we get a whole lot more serious about cybersecurity, one big disaster – say, a hacked connected car system that leaves 10,000 dead and 100,000 injured by making a million cars turn left at 8:45 one morning – will give us an Analog Tomorrow instead.”

**Richard Lachmann**, professor of political sociology at the State University of New York-Albany, said, “Whatever quantum computing achieves, it will be created by humans who serve particular interests, either for corporations of making profit or for governments of controlling populations. So, humans will be in the loop, but not all humans, most likely just those with money and power. Those people always work to serve their own interests, and so it is unrealistic to expect that the AI systems they create or contract to be created will be ethical. The only hope for ethical AI is if social movements make those demands and keep up the pressure to be able to see what is being created and to impose controls.”

**The digital minister for a Southeast Asian nation-state** said, “The problem with programming anything is the programmer. If they are not ethical, their system will not be.”

**A professor of government at one of the world’s leading universities** said, “Ethical efforts will occur in parallel with efforts that are not. The question is not whether quantum computing will assist in building ethical AI but whether it will significantly retard less-favorable developments.”

**A research director for a major university center investigating the impact of digital evolution on humanity** said, “[Computing power for AI is advancing faster than Moore’s law.](#)”

There have been recent breakthroughs in quantum computing publicized by Google and other companies. However, although system performance may improve, transparency may not – such systems may become even more complicated, unintelligible and more difficult to regulate.”

**Shel Israel**, Forbes columnist and author of many books on disruptive technologies, commented, “Quantum computing does not change the principles of computing. But, in theory, it allows computers to solve problems and perform faster by orders of magnitude. They will be smarter because AI is starting to improve exponentially. Once again, the computing itself will be neither good nor evil. That is up to those who develop, sell and use the technologies. Perhaps gunmakers intend them for defense, but that does not stop thousands and thousands of human deaths and animals being killed just for the fun of it.”

**Andrea Romaoli Garcia**, an international lawyer actively involved with multistakeholder activities of the International Telecommunication Union and Internet Society, said, “Classical computers have limitations, and quantum computers are necessary to allow the ultimate implementations of AI and machine learning. However, ethical regulation and laws are not keeping up with advances in AI and are not ready for the arrival of quantum computing. Quantum’s capability to process huge volumes of data will create a huge profit center for corporations, and this has typically led them to move quickly and not always ethically. It also allows bad actors to operate freely. Ethical AI should be supported by strong regulatory tools that encourage safe technological advancement. If not, we will face new and dangerous cyber threats.”

**Maja Vujovic**, a consultant for digital and ICT at Compass Communications, noted, “Quantum computing will prove quite elusive and hard to measure and therefore will progress slowly and painstakingly. Combining two insufficiently understood technologies would not be prudent. Perhaps the right approach would be to couple each with blockchain-based ledgers, as a way to track and decode their black-box activity.”

**Monica Murero**, director, E-Life International Institute and associate professor in Communication and New Technologies at the University of Naples Federico II, noted, “A quantum computing superpower may somewhat assist in creating ethical artificial intelligence systems that help regulate, evaluate and ‘control’ AI in-out process. But I do not think that a cool technological solution is enough or is the key. In the near future, society will rapidly change thanks to AI and quantum computing. It’s like reorganizing society. We need, as a community, to work together and rewrite the fundamental rules of coexistence that go well beyond ethical considerations. A sort of Rousseau’s new social contract: An AIQC contract. We need the means to enforce the new rules because quantum computing super-power can be extremely attractive for governments and big companies. Think about generating fake news at quantum computing super-power: unacceptable. Now think about quantum computing fighting fake news: pretty cool. My view of quantum

computing in the next decade is systemic. Quantum computing can somewhat help an ethical development of AI *if* we regulate it. I see quantum computing super-power to have the potential of solving (faster) many complex scientific problems – in health care, for example. But I also see this technology being able to break ‘normal’ encryption systems that are currently protecting our society around the world. I also see a developing business to make quantum computing and machine learning run and then sell ‘the antidote’ to protect our systems at a fair price to cure the problem: Quantum-safe cryptography blockchain. It’s like a computer virus and the antivirus business. We truly have to urgently work as a society to regulate our ecosystem and arrive in the next decade by planning in advance rather than by going along with the outcomes.”

**The military strategy and technology director** responded, “Quantum will evolve. The timescale is uncertain, but my gut sense is quantum computing will emerge in a significant way in the early to mid-2030s. How much it will assist in creating AI appears to be dependent on the nature of the AI. Quantum computing may help in complex pattern recognition.”

**The head of research at a major U.S. wireless communications trade association** responded, “It is likely that quantum computing will evolve, and that it might be deployed by those hoping to build ethical AI, but that those responsible for implementing the AI systems will either underrate its importance in the nominally neutral systems being deployed by local governments and private-sector institutions or consider it irrelevant or even hostile to the intended uses of the non-neutral monitoring and control systems being developed for use by state and nonstate institutions. Those who may not underrate its importance will not be those with the decision-making power with respect to its implementation. Ethical individuals are essential but will be marginalized by significant decision-makers.”

**An expert in learning technologies and digital life** wrote, “Many folks, including experts, still don’t know what they think about quantum computing and how to think about quantum computing in relation to AI, much less about the possibilities of its assistance with ethical AI. The theoretical musings on the subject cover the waterfront of exploratory communication among experts and amateur experts. Humans will still be in the loop as AI systems are created and implemented, assuming we don’t create our own destruction device, which we are perfectly capable of doing through ignorance, fatigue, lack of care, existing unethical practice, etc. A crisis can help an evolution unfold because of great need(s), but crisis-driven thinking and feeling are not always rational enough to benefit the changes needed.”

**The chief technology officer for a technology strategies and solutions company** said, “This isn’t a technical question. It’s about the people charged with research and development. I hope no one has cause to repeat [Robert Oppenheimer’s thought](#) after the first atomic bomb exploded.”

## Responses from those who said quantum computing is somewhat unlikely or very unlikely to assist in working toward ethical design of artificial intelligence

**David Karger**, professor at MIT's Computer Science and Artificial Intelligence Laboratory, said, "Quantum computing is, in public, being used as shorthand for 'really fast computers.' But that's not what it is. Quantum computers are highly specialized devices that are good at very specific tasks such as factoring. There's a small chance these computers will have a significant impact on cryptography by 2030 (I doubt it), but I see almost no chance that they will improve our ability to solve complex machine-learning problems, much less have any impact on our understanding of knowledge representation or creativity or any of the other key attributes of natural intelligence that we have been trying to understand and emulate in machines for decades. Finally, even if we do somehow create super-fast computers, they still won't help us with the key challenge in the design of ethical AI, which is to understand ethics. After thousands of years, this is something people are still arguing about. Having faster computers won't change the arguments one bit."

**Jim Spohrer**, director of cognitive open technologies and the AI developer ecosystem at IBM, said, "Quantum computing is decades away from being practical. It will be important by 2040."

**Michael Wollowski**, a professor of computer science at Rose-Hulman Institute of expert in artificial intelligence, said, "Quantum computing is still in its infancy. In 15 or 20 years, yes, we can build real systems. I don't think we will be able to build usable systems in 10 years. Furthermore, quantum computing is still a computational system. It is the software or in case of statistical machine learning, the data that makes a system ethical or not."

**Sam S. Adams**, a 24-year veteran of IBM now working as a senior research scientist in artificial intelligence for RTI International, "Quantum computing, if and when it becomes a commercially scalable reality, will basically allow AI systems to consider vast high-dimensional alternatives at near-instantaneous speed. This will allow not only playing hyper-dimensional chess in real-time but consider the impact of being able to simulate an entire economy at high resolution faster than in real-time. Program trading run amok in financial markets has caused global economic crises before. Now, accelerate that risk by orders of magnitude. Again, too much opportunity to gain extreme wealth and power for bad actors to ignore. The threat/opportunity of QC already fuels a global arms race in cryptography and privacy. Ethics barely has a chair in the hallway – let alone at the table in the national war rooms. That said, if a cost and scale breakthrough allows for the widespread democratization of QC, then the playing field is leveled. What if a \$30 Raspberry Pi/Q gave every device a quantum-supremacy-level capability?"



**Charlie Kaufman**, a security architect with Dell EMC, said, “Quantum computing may have an important influence on cryptography and in solving problems in physics and chemistry, and it might be used to accelerate AI if it is developed to solve those other problems, but AI doesn’t need it. AI will benefit from computation becoming cheaper and more parallel. In terms of hardware advances, the most important are likely to be in GPUs, FPGAs [field-programmable gate arrays] and customized CPUs.”

**Dan S. Wallach**, a professor in the systems group at Rice University’s Department of Computer Science, said, “Quantum computing promises speedups over classical computing in a very small number of circumstances. Probably the only such task of note today is that quantum computers have the potential to break cryptographic algorithms in widespread use today. Academic cryptographers are already hard at work on ‘post-quantum’ cryptography, which works today but is significantly less efficient than classical cryptosystems. Hopefully, by the time quantum computers are operational, we’ll have better substitutes ready. It is, of course, entirely possible that quantum computers will be able to someday accelerate the process of training machine learning models or other tasks that today are exceptionally computationally intensive. That would be fantastic, but it really has nothing to do with ethical vs. unethical AI. It’s just about spending less electricity and time to compute the same solution.”

**John Smart**, foresight educator, scholar, author, consultant and speaker, observed, “Quantum computing should be thought of like fusion. A high-cost, complex technology easily captured, slowed down and restricted by plutocrats. There’s nothing commodity about it. Human brains don’t use quantum computing. The real coming disruption is in neuro-inspired, self-improving AI. Quantum computing could definitely assist in building more brain-inspired systems, via simulation of neurodynamics. Simulation of biological and chemical processes to improve medicine, find new materials, etc., is the killer app.”

**Glenn Edens**, professor at Thunderbird School of Global Management, Arizona State University, previously a vice president at PARC, observed, “Quantum computing has a long way to go, and we barely understand it, how to ‘program’ it and how to build it at cost-effective scale. My point of view is that we will just be crossing those thresholds in 10 years’ time, maybe eight years. I’d be surprised (pleasantly so) if we got to commercial scale QC in five years. Meanwhile, AI and ML are well on the way to commercialization at scale, as well as custom silicon SoCs (system on chip) targeted to provide high-speed performance for AI and ML algorithms. This custom silicon will have the most impact in the next five to 10 years, as well as the continued progress of memory systems, CPUs and GPUs. Quantum computing will ‘miss’ this first wave of mass commercialization of AI and ML and thus will not be a significant factor. Why? It is possible that

QC might have an impact in the 10- to 20-year timeframe, but it's way too early to predict with any confidence (we simply have too much work ahead). Will humans still be in the loop? That is as much a policy decision as a pragmatic decision – we are rapidly getting to the point where synthetically created algorithms (be it AI, CA, etc.) will be very hard for humans to understand; there are a few examples that suggest we may already be to that point. Can we even create testing and validation algorithms for ML (much less AI) is a key question, and how will we verify these systems?”

**Michael Richardson**, open-source consulting engineer, responded, “It is very unlikely that a practical quantum computer will become available before 2030 that will be cheap enough to apply to AI. Will a big company and/or government manage to build a QC with enough qubits to factor current 2048-bit RSA keys easily? Maybe. At a cost that breaks the internet? Not sure. At a cost where it can be applied to AI? No. Will ML chips able to simulate thousands of neurons become very cheap? Yes, and the Moore’s Law for them will be very different because the power usage will be far more distributed. This will open many opportunities, but none of them are in the AI of science fiction.”

**Neil Davies**, co-founder of Predictable Network Solutions and a pioneer of the committee that oversaw the UK’s initial networking developments, commented, “Quantum computing only helps on algorithms where the underlying relationships are reversible – it has the potential to reduce the elapsed time for a ‘result’ to appear – it is not a magical portal to a realm where things that were intrinsically unanswerable suddenly become answerable. Where is the underlying theoretical basis for the evaluation of ethics as a function of a set numerical values that underpin the process? Without such a framework accelerating the time to get a ‘result’ only results in creating more potential hazards. Why? because to exploit quantum computation means deliberately not using a whole swath of techniques, hence reducing the diversity (thus negating any self-correcting assurance that may have been latent).”

**Kenneth A. Grady**, adjunct professor at Michigan State University College of Law and editor of “The Algorithmic Society” on Medium, said, “Despite the many impressive advances of the entities pursuing quantum computing, it is a complicated, expensive and difficult-to-scale technology at this time. The initial uses will be high-end, such as military and financial, and key applications such as pharmaceutical development. Widespread application of quantum computing to enforce ethical AI will face many challenges that quantum computing alone cannot solve (e.g., what is ‘ethical,’ when should it be enforced). Those pursuing quantum computing fall into more than one category. That is, for every entity who sees its ‘ethical’ potentials, we must assume there is any

entity who sees its ‘unethical’ potentials. As with prior technology races, the participants are not limited to those who share one ideology.”

**Chris Savage**, a leading expert in legal and regulatory issues based in Washington, D.C., noted, “AI has something of an architecture problem: It is highly computationally intensive (think Alexa or Siri), to such a degree that it is difficult to do onsite. Instead, robust connections to a powerful central processing capability (in the cloud) is necessary to make it work, which requires robust high-speed connectivity to the end points, which raises problems of latency (too much time getting the bits between the endpoint and the processing) for many applications. Quantum computing may make the centralized/cloud-based computations more rapid and thorough, but it will have no effect on latency. And if we can’t get enough old-style Boolean silicon-based computing power out to the edges, which we seem unable to do, the prospect of getting enough quantum computing resources to the edges is bleak. As to ethics, the problem with building ethical AI isn’t that we don’t have enough computational power to do it right (an issue that quantum computing could, in theory, address), it’s that we don’t know what ‘doing it right’ means in the first place.”

**Carol Smith**, a senior research scientist in human-machine interaction at Carnegie Mellon University’s Software Engineering Institute, said, “Quantum computing will likely evolve to improve computing power, but people are what will make AI systems ethical ... AI systems created by humans will be no better at ethics than we are – and, in many cases, much worse, as they will struggle to see the most important aspects. The humanity of each individual and the context in which significant decisions must always be considered.”

**Kevin T. Leicht**, professor and head of the department of sociology at the University of Illinois-Urbana-Champaign, commented, “Relying on one technology to fix the potential defects in another technology suffers from the same basic problem – technologies don’t determine ethics. People, cultures and institutions do. If those people, cultures and institutions are strong, then the potential of getting more ethical outcomes is more likely than not. We simply don’t have that. In fact, relying on quantum computing to fix anything sounds an awful lot like expecting free markets to fix the problems created by free markets. This homeopathic solution has not worked with markets, so it is difficult to see how it will work with computing. So, let’s take an elementary example that may be more applicable to the English-speaking world than elsewhere. The inventor of an AI program seeks to make as much money as possible in the shortest amount of time, because that is the prevailing institutional and economic model they have been exposed to. They develop their AI/quantum computing platform to make ‘ethical decisions,’ but those decisions happen in a context where the institutional environment where the inventor is rewarded the behaviors associated with making as much money as possible in shortest amount of time. I ask

you, given the initial constraint (“The primary goal is to be a billionaire”), all of the ethical decisions programmed into the AI/quantum computing application will be oriented toward that primary goal and make ethical decisions around that.”

**Paul Jones**, professor emeritus of information science at the University of North Carolina, Chapel Hill, observed, “While engineers are excited about quantum computing, it only answers part of what is needed to improve AI challenges. Massive amounts of data, massive amounts of computing power (not limited to quantum as a source), reflexive software design, heuristic environments, highly connected devices, sensors (or other inputs) in real time are all needed. Quantum computing is only part of the solution. More important will be insight as to how to evaluate AI’s overall impact and learning.”

**Glynn Rogers**, retired, previously senior principal engineer and a founding member at the CSIRO Centre for Complex Systems Science, said, “Computer power is not the fundamental issue. What we mean by AI, what expectations we have of it and what constraints we need to place on it are the fundamental issues. It may be that implementing AI systems that satisfy these requirements will need the level of computing power that quantum computing provides if it is the full understanding of the implications of quantum mechanics that will provide insights into the nature of intelligence, not quantum computing technology itself.”

**A telecommunications and internet industry economist, architect and consultant with over 25 years of experience** responded, “Quantum computing will develop, yes, but will it benefit ethical AI systems? AI systems will, once fully unleashed, have their own biology. I do not think we understand their complex system interaction effects any more than we understand pre-AI economics. All of our models are at best partial.”

**An ethics expert who served as an advisor on the UK’s report on “AI in Health care”** responded, “Quantum computing will take an already barely tractable problem (AI explainability) and make it completely intractable. Quantum algorithms will be even less susceptible of description and verification by external parties, in particular laypeople, than current statistical algorithms.”

**Gregory Shannon**, chief scientist at the CERT software engineering institute at Carnegie Mellon University, wrote, “I don’t see the connection between quantum computing and AI ethics. They seem very orthogonal. QC in 2030 might make building AI models/systems faster/more efficient, but that doesn’t impact ethics per se. If anything, QC could make AI systems less ethical because it will still take significant financial resources in 2030 for QC. So, a QC-generated model might be

able to ‘hide’ features/decisions that non-QC capable users/inspectors would not see/observe due to their limited computational resources.”

**Micah Altman**, a social and information scientist at MIT, said, “Quantum computing will not be of great help in building ethical AI in the next decade, since the most fundamental technical challenge in building ethical systems is in our basic theoretical understanding of how to encode within algorithms and/or teach ethical rules to learning systems. Although QC is certain to advance, and likely to advance substantially, such advances are [also] likely to apply to specific problem domains that are not closely related such as cryptography and secure communication, solving difficult search and optimization problems. Even if QC advances in a revolutionary way, for example by (despite daunting theoretical and practical barriers) exponentially speeding up computing broadly or even to the extent of catalyzing the development of self-aware general artificial intelligence – this will serve only to make the problem of developing ethical AI more urgent.”

**A distinguished professor of computer science and engineering** said, “Quantum computing might be helpful in some limited utilitarian ethical evaluations (i.e., pre-evaluating the set of potential outcomes to identify serious failings), but I don’t see most ethical frameworks benefiting from the explore/recognize model of quantum computing.”

**Michael G. Dyer**, a professor emeritus of computer science at UCLA expert in Natural Language Processing, responded, “What quantum computing offers is an incredible speed-up for certain tasks. It is possible that some task (e.g., hunting for certain patterns in large datasets) would be a subfunction in a larger classical reasoning/planning system with moral-based reasoning/planning capabilities. If we are talking simply about classification tasks (which artificial neural networks, such as ‘deep’ neural networks, already perform) then, once scaled up, a quantum computer could aid in classification tasks. Some classification tasks might be deemed ‘moral’ in the sense that [for example] people would get classified in various ways, affecting their career outcomes. I do not think quantum computing will ‘assist in building ethical AI.’”

**An anonymous respondent** observed, “I expect that quantum computing will evolve to assist in building AI. The sheer increase in computation capacity will make certain problems tractable that simple wouldn’t be otherwise. However, I don’t know that these improvements will be particularly biased toward ethical AI. I suppose there is some hope that greater computing capacity (and hence lower cost) will allow for the inclusion of factors in models that otherwise would have been considered marginal, making it easier in some sense to do the right thing.”

**John Harlow**, smart cities research specialist at the Engagement Lab @ Emerson College, noted, “We don’t really have quantum computing now, or ethical AI, so the likeliest scenario is that they don’t mature into being and interact in mutually reinforcing ways. Maybe I’m in the wrong circles, but I don’t see momentum toward ethical AI anywhere. I see momentum toward effective AI, and effective AI relying on biased datasets. I see momentum toward banning facial recognition technologies in the U.S. and some GDPR movement in Europe about data. I don’t see ethicists embedded with the scientists developing AI, and even if there were, how exactly will we decide what is ethical at scale? I mean, ethicists have differences of opinion. Clearly, individuals have different ethics. How would it be possible to attach a consensus ‘ethics’ to AI in general? The predictive policing model is awful: Pay us to run your data through a racist black box. Ethics in AI is expansive, though (<https://anatomyof.ai/>). Where are we locating AI ethics that we could separate it from the stack of ethical crises we have already? Is it ethical for Facebook workers to watch traumatic content to moderate the site? Is it ethical for slaves to mine the materials that make up the devices and servers needed for AI? Is it ethical for AI to manifest in languages, places and applications that have historically been white supremacist?”

**John L. King**, a professor at the University of Michigan School of Information, commented, “There could be earth-shattering, unforeseen breakthroughs. They have happened before. But they are rare. It is likely that the effect of technological advances will be held back by the sea anchor of human behavior (e.g., individual choices, folkways, mores, social conventions, rules, regulations, laws).”

**Douglas Rushkoff**, well-known media theorist, author and professor of media at City University of New York, wrote, “I am thinking, or at least hoping, that quantum computing is further off than we imagine. We are just not ready for it as a civilization. I don’t know if humans will be ‘in the loop’ because quantum isn’t really a cybernetic feedback loop like what we think of as computers today. I don’t know how much humans are in the loop even now, between capitalism and digital. Quantum would take us out of the equation.”

## 5. About this canvassing of experts

This report is the second of two reports issued in 2021 that share results from the 12<sup>th</sup> “[Future of the Internet](#)” canvassing by the Pew Research Center and Elon University’s Imagining the Internet Center. The first report examined the “[new normal](#)” for digital life that could exist in 2025 in the wake of the outbreak of the global pandemic and other crises in 2020.

For this report, experts were asked to respond to several questions about the future of ethical artificial intelligence via a web-based instrument that was open to them from June 30-July 27. In all, 602 people responded after invitations were emailed to more than 10,000 experts and members of the interested public. The results published here come from a nonscientific, nonrandom, opt-in sample and are not projectable to any other population other than the individuals expressing their points of view in this sample.

Respondent answers were solicited through the following prompts:

***Application of AI Ethics by 2030:** In recent years, there have been scores of convenings and even more papers generated proposing ethical frameworks for the application of artificial intelligence (AI). They cover a host of issues, including transparency, justice and fairness, privacy, freedom and human autonomy, beneficence and non-maleficence, freedom, trust, sustainability and dignity. Our questions here seek your predictions about the possibilities for such efforts.*

***The question on the future of ethical AI:** By 2030, will most of the AI systems being used by organizations of all sorts employ ethical principles focused primarily on the public good?*

***-YES,** ethical principles focused primarily on the public good WILL be employed in most AI systems by 2030*

***-NO,** ethical principles focused primarily on the public good WILL NOT be employed in most AI systems by 2030*

***Follow-up question on ethical AI, seeking a written elaboration on the previous question:** Will AI mostly be used in ethical or questionable ways in the next decade? Why? What gives you the most hope? What worries you the most? How do you see AI applications making a difference in the lives of most people? As you look at the global competition over AI systems, what issues concern you or excite you?*

Results for the quantitative question regarding how widely deployed ethical AI systems will be in 2030:

- **32%** said **YES**, ethical principles focused primarily on the public good **WILL** be employed in most AI systems by 2030
- **68%** said **NO**, ethical principles focused primarily on the public good **WILL NOT** be employed in most AI systems by 2030

The respondents were also asked to consider the possible role that quantum computing might make in creating ethical AI systems. The prompting question was:

***Quantum computing?*** How likely is it that quantum computing will evolve over the next decade to assist in creating ethical artificial intelligence systems?

- Very likely
- Somewhat likely
- Somewhat unlikely
- Very unlikely

In all, 551 respondents answered this question: 17% said “very likely”; 32% said “somewhat likely”; 27% said “somewhat unlikely”; and 24% said “very unlikely.”

The follow-up prompt to elicit their open-end written answers was:

***Follow-up on quantum computing (written elaboration).*** If you do not think it likely that quantum computing will evolve to assist in building ethical AI, why not? If you think that will be likely, why do you think so? How will that evolution unfold and when? Will humans still be in the loop as AI systems are created and implemented?

The web-based instrument was first sent directly to an international set of experts (primarily U.S.-based) identified and accumulated by Pew Research and Elon University during previous studies, as well as those identified in a 2003 study of [people who made predictions about the likely future of the internet between 1990 and 1995](#). Additional experts with proven interest in digital health, artificial intelligence ethics and other aspects of these particular research topics were also added to the list. We invited a large number of professionals and policy people from government bodies and technology businesses, think tanks and interest networks (for instance, those that include



professionals and academics in law, ethics, medicine, political science, economics, social and civic innovation, sociology, psychology and communications); globally located people working with communications technologies in government positions; technologists and innovators; top universities' engineering/computer science, political science, sociology/anthropology and business/entrepreneurship faculty, graduate students and postgraduate researchers; plus some who are active in civil society organizations that focus on digital life; and those affiliated with newly emerging nonprofits and other research units examining the impacts of digital life.

Among those invited were researchers, developers and business leaders from leading global organizations, including Oxford, Cambridge, MIT, Stanford and Carnegie Mellon universities; Google, Microsoft, Akamai, IBM and Cloudflare; leaders active in the advancement of and innovation in global communications networks and technology policy, such as the Internet Engineering Task Force (IETF), Internet Corporation for Assigned Names and Numbers (ICANN), Internet Society (ISOC), International Telecommunications Union (ITU), Association of Internet Researchers (AoIR), and the Organization for Economic Cooperation and Development (OECD). Invitees were encouraged to share the survey link with others they believed would have an interest in participating, thus there may have been somewhat of a "snowball" effect as some invitees invited others to weigh in.

*The respondents' remarks reflect their personal positions and are not the positions of their employers; the descriptions of their leadership roles help identify their background and the locus of their expertise. Some responses are lightly edited for style and readability.*

A large number of the expert respondents elected to remain anonymous. Because people's level of expertise is an important element of their participation in the conversation, anonymous respondents were given the opportunity to share a description of their internet expertise or background, and this was noted, when available, in this report.

In this canvassing, 65% of respondents answered at least one of the demographic questions. Seventy percent of these 591 people identified as male and 30% as female. Some 77% identified themselves as being based in North America, while 23% are located in other parts of the world. When asked about their "primary area of interest," 37% identified themselves as professor/teacher; 14% as research scientists; 13% as futurists or consultants; 9% as technology developers or administrators; 7% as advocates or activist users; 8% as entrepreneurs or business leaders; 3% as pioneers or originators; and 10% specified their primary area of interest as "other."

Following is a list noting a selection of key respondents who took credit for their responses on at least one of the overall topics in this canvassing. Workplaces are included to show expertise; they reflect the respondents' job titles and locations at the time of this canvassing.

**Sam Adams**, 24-year veteran of IBM now senior research scientist in artificial intelligence for RTI International; **Micah Altman**, a social and information scientist at MIT; **Robert D. Atkinson**, president of the Information Technology and Innovation Foundation; **David Barnhizer**, professor of law emeritus and co-author of “The Artificial Intelligence Contagion: Can Democracy Withstand the Imminent Transformation of Work, Wealth and the Social Order?”; **Marjory S. Blumenthal**, director of the science, technology and policy program at RAND Corporation; **Gary A. Bolles**, chair for the future of work at Singularity University; **danah boyd**, principal researcher, Microsoft Research, and founder of Data and Society; **Stowe Boyd**, consulting futurist expert in technological evolution and the future of work; **Henry E. Brady**, dean of the Goldman School of Public Policy at the University of California, Berkeley; **Tim Bray**, technology leader who has worked for Amazon, Google and Sun Microsystems; **David Brin**, physicist, futures thinker and author of the science fiction novels “Earth” and “Existence”; **Nigel Cameron**, president emeritus, Center for Policy on Emerging Technologies; **Kathleen M. Carley**, director, Center for Computational Analysis of Social and Organizational Systems, Carnegie Mellon University; **Jamais Cascio**, distinguished fellow at the Institute for the Future; **Vint Cerf**, Internet Hall of Fame member and vice president and chief internet evangelist at Google; **Barry Chudakov**, founder and principal at Sertain Research; **Adam Clayton Powell III**, senior fellow, USC Annenberg Center on Communication Leadership and Policy; **Christina J. Colclough**, an expert on the future of work and the politics of technology and ethics in AI; **Susan Crawford**, a professor at Harvard Law School and former special assistant in the Obama White House for science, technology and innovation policy; **Kenneth Cukier**, senior editor at The Economist and coauthor of “Big Data”; **Neil Davies**, co-founder of Predictable Network Solutions and a pioneer of the committee that oversaw the UK’s initial networking developments; **Rosalie Day**, policy leader and consultancy owner specializing in system approaches to data ethics, compliance and trust; **Abigail De Kosnik**, director of the Center for New Media, University of California, Berkeley; **Amali De Silva-Mitchell**, futurist and consultant participating in global internet governance processes; **Jeanne Dietsch**, New Hampshire senator and former CEO of MobileRobots Inc.; **Stephen Downes**, senior research officer for digital technologies, National Research Council of Canada; **Bill Dutton**, professor of media and information policy at Michigan State University, former director of the Oxford Internet Institute; **Esther Dyson**, internet pioneer, journalist, entrepreneur and executive founder of Way to Wellville; **Glenn Edens**, professor at Thunderbird School of Global Management, Arizona State University, previously a vice president at PARC; **June Anne English-Lueck**, professor of anthropology at San Jose State University and a distinguished fellow at the Institute for the Future; **Susan Etlinger**, industry analyst for Altimeter Group; **Daniel Farber**, author,

historian and professor of law at the University of California, Berkeley; **Marcel Fafchamps**, professor of economics and senior fellow at the Center on Democracy, Development and the Rule of Law at Stanford University; **Seth Finkelstein**, consulting programmer and Electronic Frontier Foundation Pioneer Award winner; **Rob Frieden**, professor of telecommunications law at Penn State, previously worked with Motorola and held senior U.S. policy positions at the FCC and National Telecommunications and Information Administration; **Edward A. Friedman**, professor emeritus of technology management at Stevens Institute of Technology; **Jerome C. Glenn**, co-founder and CEO of the futures-research organization The Millennium Project; **Mike Godwin**, former general counsel for the Wikimedia Foundation and author of Godwin's Law; **Kenneth Grady**, futurist, founding author of The Algorithmic Society blog; **Erhardt Graeff**, researcher expert in the design and use of technology for civic and political engagement, Olin College of Engineering; **Benjamin Grosz**, chief scientist at Kyndi, a Silicon Valley AI startup; **Glenn Grossman**, a consultant of banking analytics at FICO; **Wendy M. Grossman**, a UK-based science writer, author of "net.wars" and founder of the magazine *The Skeptic*; **Jonathan Grudin**, principal researcher, Microsoft; **John Harlow**, smart-city research specialist at the Engagement Lab at Emerson College; **Brian Harvey**, emeritus professor of computer science at the University of California, Berkeley; **Su Sonia Herring**, a Turkish-American internet policy researcher with Global Internet Policy Digital Watch; **Mireille Hildebrandt**, expert in cultural anthropology and the law and editor of "Law, Human Agency and Autonomic Computing"; **Gus Hosein**, executive director of Privacy International; **Stephan G. Humer**, lecturer expert in digital life at Hochschule Fresenius University of Applied Sciences in Berlin; **Alan Inouye**, senior director for public policy and government, American Library Association; **Shel Israel**, Forbes columnist and author of many books on disruptive technologies; **Maggie Jackson**, former *Boston Globe* columnist and author of "Distracted: Reclaiming Our Focus in a World of Lost Attention"; **Jeff Jarvis**, director, Tow-Knight Center, City University of New York; **Jeff Johnson**, professor of computer science, University of San Francisco, previously worked at Xerox, HP Labs and Sun Microsystems; **Paul Jones**, professor emeritus of information science at the University of North Carolina, Chapel Hill; **Anthony Judge**, editor of the "Encyclopedia of World Problems and Human Potential"; **David Karger**, professor at MIT's Computer Science and Artificial Intelligence Laboratory; **Frank Kaufmann**, president of the Twelve Gates Foundation; **Eric Knorr**, pioneering technology journalist and editor in chief of IDG; **Jonathan Kolber**, a member of the TechCast Global panel of forecasters and author of a book about the threats of automation; **Gary L. Kreps**, director of the Center for Health and Risk Communication at George Mason University; **David Krieger**, director of the Institute for Communication and Leadership, based in Switzerland; **Benjamin Kuipers**, professor of computer science and engineering at the University of Michigan; **Patrick Larvie**, global lead for the workplace user-experience team at one of the world's largest technology companies; **Jon Lebkowsky**, CEO, founder and digital strategist, Polycot Associates; **Sam Lehman-Wilzig**, professor and former chair of communication at Bar-Ilan University, Israel; **Mark Lemley**,

director of Stanford University's Program in Law, Science and Technology; **Peter Levine**, professor of citizenship and public affairs at Tufts University; **Rich Ling**, professor at Nanyang Technological University, Singapore; **J. Scott Marcus**, an economist, political scientist and engineer who works as a telecommunications consultant; **Nathalie Maréchal**, senior research analyst at Ranking Digital Rights; **Alice E. Marwick**, assistant professor of communication at the University of North Carolina, Chapel Hill, and adviser for the Media Manipulation project at the Data & Society Research Institute; **Katie McAuliffe**, executive director for Digital Liberty; **Pamela McCorduck**, writer, consultant and author of several books, including "Machines Who Think"; **Melissa Michelson**, professor of political science, Menlo College; **Steven Miller**, vice provost and professor of information systems, Singapore Management University; **James Morris**, professor of computer science at Carnegie Mellon; **David Mussington**, senior fellow at CIGI and director at the Center for Public Policy and Private Enterprise at the University of Maryland; **Alan Mutter**, consultant and former Silicon Valley CEO; **Beth Noveck**, director, New York University Governance Lab; **Concepcion Olavarrieta**, foresight and economic consultant and president of the Mexico node of The Millennium Project; **Fabrice Popineau**, an expert on AI, computer intelligence and knowledge engineering based in France; **Oksana Prykhodko**, director of the European Media Platform, an international NGO; **Calton Pu**, professor and chair in the School of Computer Science at Georgia Tech; **Irina Raicu**, director of the Internet Ethics program at the Markkula Center for Applied Ethics; **Glynn Rogers**, retired, previously senior principal engineer and a founding member at the CSIRO Centre for Complex Systems Science; **Douglas Rushkoff**, writer, documentarian and professor of media, City University of New York; **Jean Seaton**, director of the Orwell Foundation and professor of media history at the University of Westminster; **Greg Sherwin**, vice president for engineering and information technology at Singularity University; **Henning Schulzrinne**, Internet Hall of Fame member, co-chair of the Internet Technical Committee of the IEEE and professor at Columbia University; **Ben Shneiderman**, distinguished professor of computer science and founder of Human Computer Interaction Lab, University of Maryland; **John Smart**, foresight educator, scholar, author, consultant and speaker; **Jim Spohrer**, director of cognitive open technologies and the AI developer ecosystem at IBM; **Sharon Sputz**, executive director, strategic programs, Columbia University Data Science Institute; **Jon Stine**, executive director of the Open Voice Network, setting standards for AI-enabled vocal assistance; **Jonathan Taplin**, author of "Move Fast and Break Things: How Google, Facebook and Amazon Cornered Culture and Undermined Democracy"; **Brad Templeton**, internet pioneer, futurist and activist, a former president of the Electronic Frontier Foundation; **Ed Terpening**, consultant and industry analyst with the Altimeter Group; **Ian Thomson**, a pioneer developer of the Pacific Knowledge Hub; **Joseph Turow**, professor of communication, University of Pennsylvania; **Dan S. Wallach**, a professor in the systems group at Rice University's Department of Computer Science; **Wendell Wallach**, ethicist and scholar at Yale University's Interdisciplinary Center for Bioethics; **Amy Webb**, founder, Future Today Institute, and professor of strategic foresight, New York University; **Jim**

**Witte**, director of the Center for Social Science Research at George Mason University; **Simeon Yates**, a professor expert in digital culture and personal interaction at the University of Liverpool and the research lead for the UK government’s Digital Culture team; **Warren Yoder**, longtime director at Public Policy Center of Mississippi, now an executive coach; **Jillian York**, director of international freedom of expression for the Electronic Frontier Foundation; and **Ethan Zuckerman**, director, MIT’s Center for Civic Media, and co-founder, Global Voices.

A selection of institutions at which some of the respondents work or have affiliations:

AAI Foresight; AI Now Research Institute of New York University; AI Impact Alliance; Access Now; Akamai Technologies; Altimeter Group; American Enterprise Institute; American Institute for Behavioral Research and Technology; American Library Association; American University; American University of Afghanistan; Anticipatory Futures Group; APNIC; Arizona State University; Aspen Institute; AT&T; Atlantic Council; Australian National University; Bar-Ilan University; Benton Institute; Bloomberg Businessweek; Brookings Institution; BT Group; Canada Without Poverty; Carleton University; Carnegie Endowment for International Peace; Carnegie Mellon University; Center for a New American Security; Center for Data Innovation; Center for Global Enterprise; Center for Health and Risk Communication at George Mason University; Center for Strategic and International Studies; Centre for International Governance Innovation; Centre National de la Recherche Scientifique, France; Chinese University of Hong Kong; Cisco Systems; Citizens and Technology Lab; City University of New York; Cloudflare; Columbia University; Constellation Research; Convo Research and Strategy; Cornell University; Council of Europe; Data Across Sectors for Health at the Illinois Public Health Institute; Data & Society Research Institute; Data Science Institute at Columbia; Davis Wright Tremaine LLP; Dell EMC; Deloitte; Digital Grassroots; Digital Value Institute; Disney; DotConnectAfrica; The Economist; Electronic Frontier Foundation; Electronic Privacy Information Center; Enterprise Roundtable Accelerator; Emerson College; Fight for the Future; European Broadcasting Union; Foresight Alliance; Future Today Institute; Futuremade; Futurous; FuturePath; Futureproof Strategies; General Electric; Georgetown University; Georgia Tech; Global Business Network; Global Internet Policy Digital Watch; Global Voices; Google; Hague Centre for Strategic Studies, Harvard University; Hochschule Fresenius University of Applied Sciences; Hokkaido University; IBM; Indiana University; Internet Corporation for Assigned Names and Numbers (ICANN); IDG; Ignite Social Media; Information Technology and Innovation Foundation; Institute for the Future; Instituto Superior Técnico, Portugal; Institute for Ethics and Emerging Technologies; Institute for Prediction Technology; International Centre for Free and Open Source Software; International Telecommunication Union; Internet Engineering Task Force (IETF); Internet Society; Internet Systems Consortium; Johns Hopkins University; Institute of Electrical and Electronics Engineers (IEEE); Ithaka; Juniper Networks; Kyndi; Le Havre University; Leading Futurists; Lifeboat Foundation; MacArthur Research Network on Open Governance; Macquarie University, Sydney,

Australia; Massachusetts Institute of Technology; Menlo College; Mercator XXI; Michigan State University; Microsoft Research; Millennium Project; Mimecast; Missions Publiques; Moses & Singer LLC; Nanyang Technological University, Singapore; Nautilus Magazine; New York University; Namibia University of Science and Technology; National Distance University of Spain; National Research Council of Canada; Nonprofit Technology Network; Northeastern University; North Carolina State University; Olin College of Engineering; Pinterest; Policy Horizons Canada; Predictable Network Solutions; R Street Institute; RAND; Ranking Digital Rights; Rice University; Rose-Hulman Institute of Technology; RTI International; San Jose State University; Santa Clara University; Sharism Lab; Singularity University; Singapore Management University; Södertörn University, Sweden; Social Science Research Council; Sorbonne University; South China University of Technology; Spacotel Consultancy LLC; Stanford University; Stevens Institute of Technology; Syracuse University; Tallinn University of Technology; TechCast Global; Tech Policy Tank; Telecommunities Canada; Tufts University; The Representation Project; Twelve Gates Foundation; United Nations; University of California, Berkeley; University of California, Los Angeles; University of California, San Diego; University College London; University of Hawaii, Manoa; University of Texas, Austin; the Universities of Alabama, Arizona, Dallas, Delaware, Florida, Maryland, Massachusetts, Miami, Michigan, Minnesota, Oklahoma, Pennsylvania, Rochester, San Francisco and Southern California; the Universities of Amsterdam, British Columbia, Cambridge, Cyprus, Edinburgh, Groningen, Liverpool, Naples, Oslo, Otago, Queensland, Toronto, West Indies; UNESCO; U.S. Geological Survey; U.S. National Science Foundation; U.S. Naval Postgraduate School; Venture Philanthropy Partners; Verizon; Virginia Tech; Vision2Lead; Volta Networks; World Wide Web Foundation; Wellville; Whitehouse Writers Group; Wikimedia Foundation; Witness; Work Futures; World Economic Forum; XponentialEQ; and Yale University Center for Bioethics.

Complete sets of credited and anonymous responses can be found here:

<https://www.elon.edu/u/imagining/surveys/xii-2021/ethical-ai-design-2030/credit>

<https://www.elon.edu/u/imagining/surveys/xii-2021/ethical-ai-design-2030/anon>

## Acknowledgments

We are extremely thankful for the contributions of the people who participated in this canvassing.

This report is a collaborative effort based on the input and analysis of the following individuals.

### Primary researchers

Lee Rainie, *Director, Internet and Technology Research*

Janna Anderson, *Director, Elon University's Imagining the Internet Center*

Emily A. Vogels, *Research Associate*

### Editorial and graphic design

Margaret Porteus, *Information Graphics Designer*

James Dryden, *Copy Editor*

### Communications and web publishing

Sara Atske, *Associate Digital Producer*

Haley Nolan, *Communications Associate*