How Emerging Technologies Are Rewiring the Global Order

September 23 and 24, 2019

PERRY WORLD HOUSE
UNIVERSITY of PENNSYLVANIA

ABOUT PERRY WORLD HOUSE

Perry World House is a center for scholarly inquiry, teaching, research, international exchange, policy engagement, and public outreach on pressing global issues.

Perry World House's mission is to bring the academic knowledge of the University of Pennsylvania to bear on some of the world's most pressing global policy challenges, and to foster international policy engagement within and beyond the Penn community.

Located in the heart of campus at 38th Street and Locust Walk, Perry World House draws on the expertise of Penn's 12 schools and numerous globally oriented research centers to educate the Penn community and prepare students to be well-informed, contributing global citizens. At the same time, Perry World House connects Penn with leading policy experts from around the world to develop and advance innovative policy proposals.

Through its rich programming, Perry World House facilitates critical conversations about global policy challenges and fosters interdisciplinary research on these topics. It presents workshops and colloquia, welcomes distinguished visitors, and produces content for global audiences and policy leaders, so that the knowledge developed at Penn can make an immediate impact around the world.

Perry World House—its programs and the building itself—builds on Penn's strengths in teaching, interdisciplinary scholarship, and policy-relevant research. By doing so, it cultivates the broad worldview, critical thinking, and leadership required to address the most urgent issues of global affairs.

REPORT AUTHORS

Christine Carpenter
JD Student, Law

Casey Mahoney
PhD Candidate, Political Science

NEXT STEPS AUTHOR

Christian Ruhl
*Perry World House Program
Associate for Global Order*

› EXECUTIVE SUMMARY

At a time of disruption and uncertainty, advances in emerging technologies like artificial intelligence (AI) have the potential to transform the world. To understand how these technologies are rewiring the global order, Perry World House convened a two-day colloquium on September 23 and 24, 2019. After a brief outline of the issues in the following introduction, the report is divided into three parts.

SCHOLARLY PROGRAM
SEPTEMBER 23, 2019

The first day of the colloquium brought together academics, current and former government officials, industry representatives, diplomats, and journalists to engage in interdisciplinary expert discussions. This section of the report outlines the outcomes of these discussions in four panels. The day's first panel, on "AI, Robotics, and the Future of Technology," set the scene for the colloquium by discussing the state of the art in machine learning and robotics and laying out the promises and perils of AI. The second panel built on this foundation to discuss a wide range of issues in AI and international security, from alliance dynamics to the weaponization of AI. The third panel, on "AI and Global Governance," highlighted the importance of international cooperation and global governance for emerging technologies. And the final panel, on "AI and Human Rights, Surveillance, and Democracy," discussed how emerging technologies may force us to rethink our institutions and our very conceptions of privacy and rights in the digital age.

PUBLIC PROGRAM
SEPTEMBER 24, 2019

The second day of the colloquium brought students, faculty, and members of the public to Perry World House for five high-level discussions with world leaders in diplomacy, defense, and industry. Former U.S. Secretary of State John Kerry held a wide-ranging conversation with *WIRED* editor-in-chief Nicholas Thompson that touched on topics from green tech to the weaponization of emerging technologies. Uber Advanced Technology Group Chief Scientist Raquel Urtasun delivered a hopeful vision of a self-driving future and the importance of explainable technology. A panel of human rights experts discussed technology as a double-edged sword for human rights and freedoms. Former President of Kyrgyzstan Roza Otunbayeva discussed her experience as the first female head of state in Central Asia. Finally, former U.S. Secretary of Defense Ash Carter closed out the colloquium with a wide-reaching conversation with Jennifer Griffin of the *Fox News Channel*.

NEXT STEPS AND POLICY RECOMMENDATIONS

Secretary Carter explained a key take-away of the colloquium: "It's up to us to make the good of these technologies win out over the bad." In pursuit of this goal, this report draws on the discussions, pre-colloquium survey, and written analyses of participants to give directions for future research and policy proposals. These are outlined in the last section of the report with three policy recommendations: (1) prioritizing the interoperability of emerging technologies, (2) fostering interdisciplinarity, and (3) investing in digital public goods and the benefits of AI.

- 45% VERY LIKELY
- 36% SOMEWHAT LIKELY
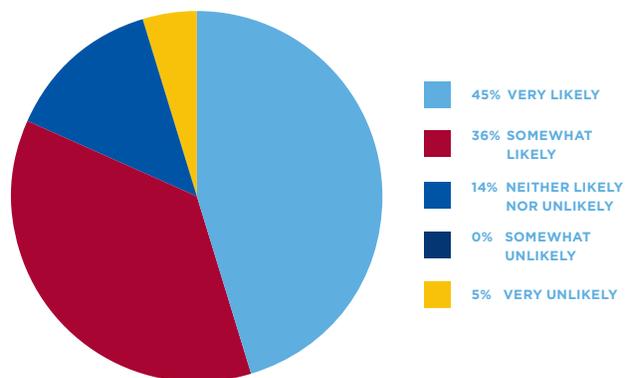- 14% NEITHER LIKELY NOR UNLIKELY
- 0% SOMEWHAT UNLIKELY
- 5% VERY UNLIKELY

Figure 1. How likely is it that advances in AI will have a significant impact on the global order over the next 25 years?

- 14% NEITHER STRONGER NOR WEAKER
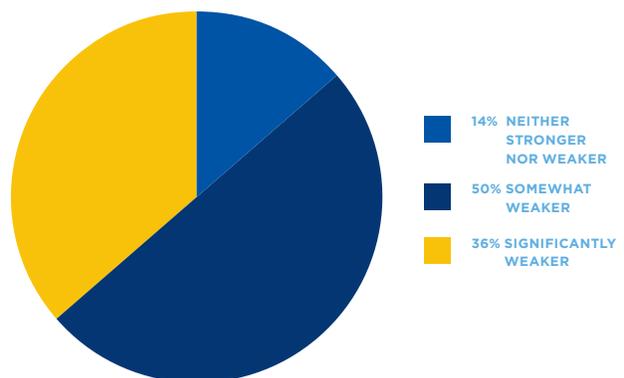- 50% SOMEWHAT WEAKER
- 36% SIGNIFICANTLY WEAKER

Figure 2. Has the global order, as you define it, grown stronger or weaker over the past five years?

› INTRODUCTION:

THE TRANSFORMATIVE POTENTIAL OF EMERGING TECHNOLOGIES

Artificial intelligence (AI) is a general-purpose technology with the potential to shape global politics—from economics to military affairs to human rights. To better understand the intersection of AI and global politics, prior to the colloquium, Perry World House conducted a survey of 22 participating experts, an interdisciplinary group from law and ethics, engineering, and the social sciences.[1] Drawing on this pre-colloquium survey, the two days of discussion, and participants' written analyses, this report frames the issues, catalyzes new scholarly inquiry, and highlights next steps for scholars and policymakers.

---

[1] The survey was conducted online between August 16 and October 14, 2019. Perry World House invited the colloquium's day one participants to complete it as an anonymous assessment of views.

THE AI DOOMSDAY CLOCK: WHAT TIME IS IT?

PARTICIPANTS' ANSWERS

23:58   There is a lack of great power consensus on how to avert global disaster.

22:45   AI capabilities are rapidly expanding, and states / non-state actors are starting to employ these capabilities. There is relatively little regulation of these systems, and there exists a high potential for accidents.

21:00   AI will not cause any doomsday event but the perception of it could.

12:00   AI is not the danger. It's the struggle to power globally and regionally.

06:00   I am much more concerned about nuclear disasters from dirty bombs in the wrong hands.

One key question is whether advances in AI represent a new form of global risk, due to the potential disruption AI could cause. To comprehend global risks in the nuclear weapons setting, since 1947, *Bulletin of the Atomic Scientists* has published a "Doomsday Clock," on which minutes to midnight represent the current risk of nuclear war (and more recently climate-related catastrophe). In the *Bulletin*'s assessment, the Doomsday Clock's hands have never been more than 17 minutes away from midnight—human-made global catastrophe.[2] Imagining an AI Doomsday Clock, with midnight as AI-caused global catastrophe, Perry World House asked participants in the 2019 Global Order Colloquium a simple question: "What time is it?" On average, respondents believed it to be about 2:45 in the afternoon on the AI Doomsday Clock—meaning an AI apocalypse is a long way off.

AI AND THE GLOBAL ORDER

Nonetheless, respondents agreed that emerging technologies are likely to make deep changes in the basic structures of global affairs. Out of those surveyed, 82 percent considered it likely or very likely that advances in AI will have a significant impact on the global order over the next 25 years. Viewed as an enabling technology— a general-purpose tool that can be applied to several domains in international politics—AI will be linked to other recent shifts in the world. AI may destabilize an already weakened global order, as 86 percent of expert respondents said the global order grew somewhat or significantly weaker over the past five years.
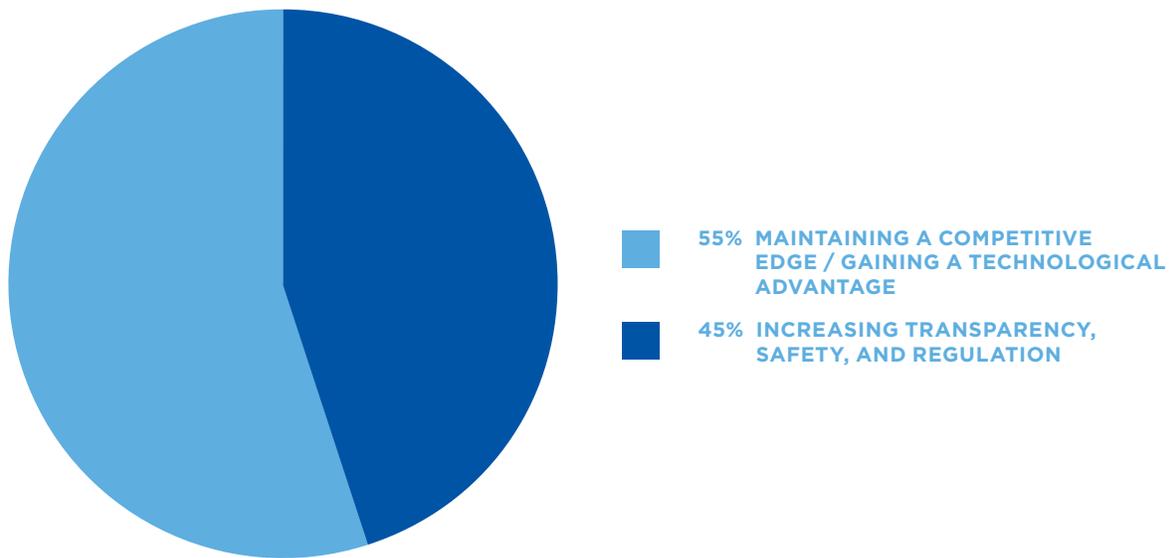
55% MAINTAINING A COMPETITIVE EDGE / GAINING A TECHNOLOGICAL ADVANTAGE

45% INCREASING TRANSPARENCY, SAFETY, AND REGULATION

Figure 3. What do you view as the more pressing concern for states with regards to the development of AI within the context of the current Global Order?

Underscoring the urgency of this colloquium, nearly 88 percent of survey respondents argued that an arms race over military applications of AI in the next 15 years was somewhat or very likely. More broadly, participants recognized a high risk that AI will exacerbate international competition. Only 14 percent of survey respondents considered there to be a greater chance that AI generates more potential for international cooperation than competition. Moreover, this competition may accompany big shifts in the distribution of power and technology. For instance, over two-thirds of survey respondents thought China would overtake the United States in AI research and development. A majority of experts surveyed said that increasing transparency, safety, and regulation is a more pressing concern than maintaining a competitive edge (Figure 3).

With these pressing concerns and the transformative potential of emerging technologies in mind, the following report demonstrates how the 2019 Global Order Colloquium advanced inquiry on the ways that emerging technologies are rewiring the global order.

› DAY ONE

› ARTIFICIAL INTELLIGENCE, ROBOTICS & THE FUTURE OF TECHNOLOGY

Day One brought together experts from industry, academia, and government to discuss AI and other emerging technologies

This panel was tasked with scene-setting for the colloquium and discussing the state of the art and future challenges in AI applications, particularly in transportation and other robotics applications. Missy Cummings, a professor in the Duke University electrical and computer engineering department and the director of the Humans and Autonomy Laboratory, provided introductory remarks, and Vijay Kumar, the Nemirovsky Family Dean of Engineering and a professor at the University of Pennsylvania, moderated conversation among the expert panel.[3]

---

[3] Perry World House conducted academic panels according to Chatham House Rules. All direct quotations are from panelists' written analyses submitted prior to the colloquium.

ASSESSING THE STATE OF THE ART

As an emerging general-purpose technology with many potential applications, from self-driving cars to military robotics, advances in AI research have led to widespread, imaginative predictions about a coming technological revolution. Overly rosy predictions, however, deserve closer scrutiny. The first colloquium panel focused on the limits of many of today's most advanced applications of AI, or the computerized simulation of human intelligence processes like learning, reasoning, and self-correction. Considering AI to run along a sliding scale from rote automation to more complex, probabilistic reasoning, applications toward the truly "intelligent" end of the spectrum remain hampered by their fundamental algorithms' weakness under conditions of high uncertainty.

Moreover, bias in the estimations and predictions that AI systems make about the world is hard—if not impossible—for humans to excise from the creation of those systems. Engineers make countless choices about the questions AI will need to answer; about the construction of statistical models that form the basis of AI; about the source, quality, and form of the data from which AI derives its output; and about the ways AI presents this output to humans in the analog world.

Predicting how even small amounts of bias in such systems, particularly when their construction is opaque to end users, will affect interactions in the social world is equally difficult. Recent, well-publicized examples of these effects, like Amazon's discovery that an AI-based résumé analysis tool mimicked human bias against non-male recruits, are almost certainly only the tip of the iceberg. This, and the shortcomings of narrow algorithms designed to deal with both specific tasks

and uncertainty, led several colloquium participants to the conclusion that AI is further from prime time in high-stakes application areas, such as war or medicine, than corporate developers or international adversaries would have their customers or competitors think.

SOCIETY-RESPONSIVE AI

A key component of any AI system is the human end user. It is vital to minimize AI glitches in safety-critical applications such as autonomous transportation. Considering developments in the automobile industry, colloquium participants discussed the risks that poorly designed human-machine interfaces can pose for drivers and pedestrians alike. New research shows the importance of taking into account individuals' unique, culturally derived understandings of the rules of the road in the design of such interfaces, yet it is unclear whether AI systems being developed in commercial markets adequately account for these differences. For instance, one study from Missy Cummings's Humans and Autonomy Lab of pedestrians distracted by cellphones showed that Asian participants were more likely to attempt risky road crossings and trusted smartphone-based alerts less than other groups.[4] In other areas, such as in internet search engines and social media, end users' lack of knowledge about how AI algorithms select the content they see has troublesome implications for how political or other types of information spread. Recent efforts in academia and industry to develop "explainable AI" need to continue, many panelists agreed, to combat misunderstandings about both the potential and the limits of AI within the societies using it.

---

[4] Mary Cummings, Lixiao Huang, and Michael Clamann, 2019, "HAL2019-01: Development and Evaluation of Vehicle to Pedestrian (V2P) Safety Interventions," Humans and Autonomy Laboratory, Duke University, https://hal.pratt.duke.edu/sites/hal.pratt.duke.edu/files/u35/HAL2019_1.pdf (accessed October 11, 2019).

BALANCING INNOVATION AND RISK-MANAGEMENT

As researchers and developers in industry and academia continue innovating, governments and society have a role to play in identifying domain-specific standards, including for safety, that technologies based on AI need to meet. Leadership from regulators, however important, requires that the public sector, from the Federal Aviation Administration and Rail Administration to the Food and Drug Administration, increase its in-house expertise in AI. Without doing so, the incentives of what Cummings and others identified in their written analyses as a pervasive fake-it-until-you-make-it culture among Silicon Valley and particularly Chinese corporations will decrease the prospects that transparent dialogue among innovators and regulators will strike an appropriate balance between technological progress and social welfare.

The importance of expertise applies in the defense sector, as well. Given that private industry, rather than military-sponsored research centers, drives most developments in AI today, Cummings writes, "it is imperative for governments to monitor developments in military-related artificial intelligence, especially for weapons systems and in cybersecurity [and to] arm themselves with the capabilities to detect inflated or faked claims." Otherwise, misperceptions of national capabilities run the risk of spiraling into competitive dynamics where large public investments in premature commercial technology provide little added capability and fuel insecurity.

QUESTIONS FOR FUTURE RESEARCH

In the design of AI systems, what is the optimal balance between data-gathering sensors and intelligent human agents?

Where do models of industry transparency and self-regulation outside of AI work best?

How can government attract individuals with expertise in AI, and how can industry increase the diversity of its workforce?

Who will provide leadership in determining what society accepts as appropriate levels of safety, fairness, privacy, or data security in how it uses AI technologies?

Participants engaged in lively discussion of the role of emerging technologies in international security.

› ARTIFICIAL INTELLIGENCE & INTERNATIONAL SECURITY

The AI and international security panel discussed the implications of AI for policy and research in international security affairs. Perry World House asked the panel to answer questions like "How well do militaries and/or the international security community understand the military and non-military effects of AI on international security?" and "How are advances in AI likely to shift the trajectory of great power competition?" Missy Cummings (Duke University), Erik Lin-Greenberg (Perry World House), Ashley J. Llorens (Johns Hopkins University Applied Physics Laboratory), Paul Scharre (Center for a New American Security), Nadia Schadlow (Perry World House), and Rebecca Slayton (Cornell University) served as expert commentators. Michael Horowitz, the interim director of Perry World House and a professor of political science at the University of Pennsylvania, moderated the panel.

MILITARY AI: NEW TECH, NEW TACTICS, NEW THINKING?

States have leveraged advances in computing and automation for military purposes since the Cold War, from detection systems for nuclear missiles to early applications of computer vision. Panelists agreed that innovation in AI-based deep learning offers new potential for militaries and intelligence agencies to increase the speed and precision of tasks, from sensor analysis and targeting to overseeing battlefield operations. As in most domains, current AI applications have major limitations. Even so, assuming these can be overcome, AI will have significant impacts on military decision-making. In his written analysis, Paul Scharre drew a parallel from how industrial revolution technologies "off-loaded" significant human labor to stronger machines, to envisioning how AI may enable a large share of basic *cognitive* tasks in the military realm to be transferred to computers. "AI systems could enable military forces to operate faster, more cohesively, and with greater precision and coordination than possible with humans," Scharre writes. "The result could be to accelerate the pace of battle beyond human decision-making." Many panel members agreed that some of AI's greatest implications in the military realm will be how command-and-control structures will need to adapt to integrate such cognitive capacity into decision-making processes. Beyond the challenges this will pose to how *individual* organizations adapt, Erik Lin-Greenberg argued in his written analysis that AI will also pose challenges that could impact coordination across interstate alliances, like the increased need for data synthesis and sharing.

As AI systems become increasingly integrated into military and intelligence settings, strategists and technologists will need to pay greater attention to the risks of increased human-machine teaming. In her written analysis, Rebecca Slayton drew attention to the common errors of AI systems and their vulnerability to deception. Experts need more research to understand how human judgment in symbiotic relations with AI-enabled machines changes, given the risk that increasing the speed of decision-times will decrease how well militaries' strategic choices and tactics fulfill their aims.

GLOBAL COMPETITION AND GREAT-POWER POLITICS

Despite the wide range of unknowns and open questions surrounding the net benefits and costs of AI adoption, many countries, both large and small, are significantly investing in research and development for military
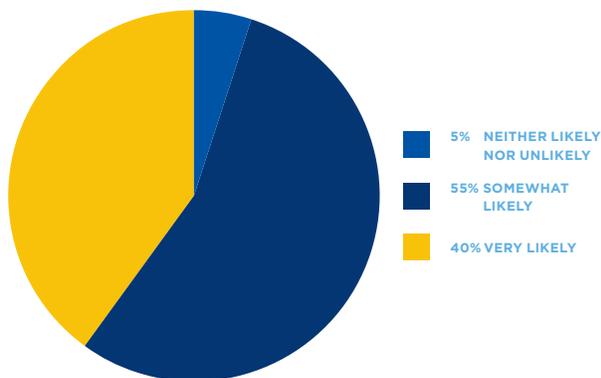
Figure 4. How likely is it that an arms race over specific military applications of AI will occur in the next 15 years?

- 5% NEITHER LIKELY NOR UNLIKELY
- 55% SOMEWHAT LIKELY
- 40% VERY LIKELY

Figure 5. How likely is China to overtake the United States in AI research and development over the next 15 years?

- 36% VERY LIKELY
- 36% SOMEWHAT LIKELY
- 18% NEITHER LIKELY NOR UNLIKELY
- 5% SOMEWHAT UNLIKELY
- 5% VERY UNLIKELY

applications of AI. Existing distributions of military and non-military power and the innovative capacity any given state possesses provide the backdrop for considering the incentives international players face to compete in this arena. A majority of the respondents to the pre-colloquium survey agreed with the prediction that, over the next 15 years, an AI arms race will occur and that China will overtake the United States in AI development (see Figures 4 and 5). Expert opinions expressed during the panel discussion reflected both sides of the debate about how "evolutionary" or "revolutionary" an impact the military adoption of AI will have on the global order.

Several panelists agreed that, despite the risks, pursuing gains in capabilities is necessary to maintain a competitive technological edge and strategic stability against adversaries with unknowable intentions. An alternative route to AI superiority could be to more carefully focus on how to derive a competitive edge from designing organizations capable of optimizing the human-machine teams of the future. In the words of one participant, referencing the U.S.-China competition over basic AI technology: "This is not the space race we think it is. Sometimes it's okay to be number two."

DOMESTIC FOUNDATIONS OF INTERNATIONAL POWER

Ultimately, decisions about the use of power on the international stage—AI-enabled or not—rest with leaders who speak on behalf of domestic societies. Part of the impact of AI on international security will stem from the values of the states that use it. Social movements, such as Computer Professionals for Social Responsibility or the Campaign to Stop Killer Robots, which both protest the development of lethal autonomous weapons systems (LAWS), reflect expert and public desires to engage governments about the implications of computerization and autonomous machines on the
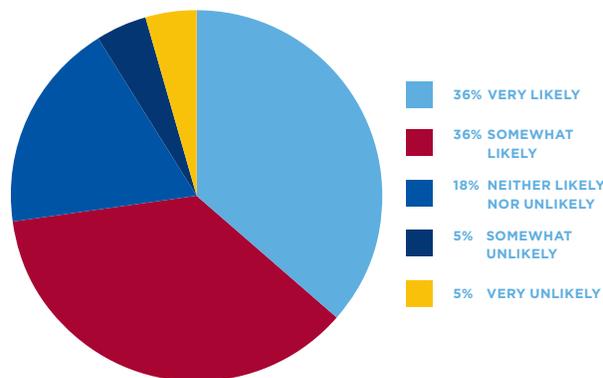
battlefield. The panel discussed the need for dialogue to enhance trust among citizens, leaders, and industry.

Innovation in Silicon Valley or the Pentagon may be breaking ground *and* breaking rules, but social values held deeply among citizens are slower to change. Slayton noted in her written analysis that decision-makers must consider how the broader integration of AI into internet-based communications will impact the information flows among societies that provide states the soft-power tools of persuasion, which often accompany the effective use of coercion in the twenty-first century. As international competition to develop effective military technologies and organizations that integrate AI continues for either demonstrated or presumed strategic benefit, leaders should attend to the unintended, unforeseen consequences of AI blowback at home.

(?)

QUESTIONS FOR FUTURE RESEARCH

What benefits are societies seeking to achieve by developing AI for military or non-military uses? Are these goals contested or consensus views? How do different states' goals conflict or harmonize?

How have militaries and other types of organizations adapted in the past to technologies that have increased the speed of information-transmission and decision-making?

Which models of private-public cooperation can provide lessons for how to ensure policymakers understand the risks and limitations of AI applications in the military sphere?

› ARTIFICIAL INTELLIGENCE & GLOBAL GOVERNANCE

The third panel focused on the global governance of AI and sought to discern whether existing global institutions are prepared to facilitate international cooperation and regulate emerging technologies, what potential strategies for international cooperation surrounding AI there are, and how the intersection of AI's hardware and software requirements influences the prospects for the governance of AI. The University of Pennsylvania's Amy Gadsden moderated the discussion among Jason Matheny (Georgetown University), Elina Noor (Institute of Strategic and International Studies, Malaysia), Ambassador Amandeep Singh Gill (Graduate Institute for International & Development Studies), and Shawn Steene (U.S. Department of Defense).

In considering both the potential risks and benefits of AI, a substantial determinant is its effective incorporation into existing regulatory systems. Viewing AI as both a tool of global governance and an object it seeks to control, the panelists offered myriad suggestions regarding the role of AI in the future of global affairs.

A REGULATORY APPROACH?

Examining AI from a non-military perspective, the colloquium turned to discussions of the civil and legal implications of AI. Many panelists called for increased regulation of machine-learning technologies and emphasized the need for informed officials, well-versed in the mechanics of the field, to play an active role in the development of these technologies in their incorporation into public institutions and governments. American University's Rebecca Hamilton wrote before the session about the need to foster "a new generation of policymakers with the willingness and ability to engage with engineers at the design stage." This model could bake-in values and reduce resistance to the retroactive regulation of these technologies.

Going beyond domestic regulatory institutions, the panel debated whether existing global governance structures are sufficient for global AI regulation. One panelist described this issue as "where technology meets national security, meets international security," and emphasized the unique challenges AI presents for existing notions of global governance that rely on interstate cooperation and export controls to maintain a firewall between the civilian and military uses of so-called "dual-use" technologies. The panelist warned that, "It is more important than ever to look at the civilian side of things, and in fact go beyond the duality in terminology we have been using since the Second World War—civilian and combatant." The panelist concluded, "We can't just handle it using our traditional lens because the involvement of private actors makes this decidedly different." Just as at the domestic level, this suggests the need to update existing models of international governance that incorporate industry actors earlier in regulatory processes.

What is the best approach to global governance of AI, then? One panelist emphasized the need for a clear and unified strategy, based on shared principles, common values, and a concrete vision. Another participant argued for the importance of legal institutions: "Law has to become international on this issue." This panelist argued that industry-wide standards will be necessary to confront the regulatory issues around how these

Ambassador Amandeep Singh Gill and Amy Gadsden discussed the urgent need for international cooperation and the possibilities for global governance of AI and other emerging technologies.

globally operating technologies develop. Rather than using existing paradigms to deal with something like AI, another panelist urged that states and others should focus on building institutions and regulations unique to AI in its specific applications instead.

THINK OF AI IN TERMS OF MEANS AND ENDS

To best coordinate domestic and international policy into systems that function efficiently, policymakers and government officials ought to be intentional in determining their desired outcomes from developments in AI. Doing so matters for the corporations and societies that stand to profit from continued innovation. Global governance can affect both the means and the ends of AI development. The means can be controlled by national governments, industry leaders, and the international community making strategic choices about AI development. Examples of such desired strategies could include: coordination on international immigration and education policy that impact global talent flows related to AI; oversight of financial institutions that monitor and govern international investment in these technologies; cooperative research enterprises; and the development of multinational data sets and standards for data formats. The ends produced as the benefits of an

intentional and strategic global-governance effort, one panelist said, could include: international agreements about applications of AI (e.g., disowned classes of autonomous weapons), limitations on publications of AI tech that could be used maliciously, and agreements to address the consequences (e.g., surveillance and censorship made easier by AI).

Ambassador Amandeep Singh Gill has asked: "How [do we] deliver on the promise of AI while protecting the hard-won tenets of international humanitarian law and respecting the legitimate security and commercial interests of states and industry?"[5] The panelists seemed to concur about the potential rise of a new approach, spearheaded by those well-versed in the actual mechanics of AI technology, that expands on existing international legal tools and demands proactive government involvement. This could involve integrating national regulatory approaches and industry self-regulation into the global-governance mode. The chain of industry self-regulation to national regulatory approaches, to international humanitarian law, panelists agreed, must move in tandem toward standards and best-practices models equipped to confront the challenges and unique questions raised by AI.

---

[5] Amandeep Singh Gill, 2017, "The AI Battlefield," *The Hindu*, November 13, 2017, https://www.thehindu.com/opinion/op-ed/the-ai-battlefield/article20353916.ece (accessed December 20, 2019).

EMPHASIZING TRULY GLOBAL INVOLVEMENT

Panelists emphasized the importance of involving states that have historically been left out of international decision-making in the development of a global-governance approach for AI. Panelist Elina Noor noted in her written analysis that Southeast Asia is becoming the fastest growing internet region—projected to reach nearly half a billion users by 2020. She argued these historically underrepresented regions will be necessary and prominent actors in the AI space, and therefore crucial for the establishment of effective global governance around AI. One panelist pointed out that, between now and 2050, more than half of the population growth will take place outside of the Western world and that most of the world's population will be in Africa. The panelist argued that, if the world's largest populations will be in Asia and Africa, decisions regarding the global governance of AI will have to incorporate those value systems and expectations.

INSTITUTIONS FOR FACILITATING FOR AI COOPERATION

The panelists quickly reached agreement that international cooperation in this realm would be beneficial. How advocates for AI governance should achieve such cooperation naturally became the question occupying much of the discussion. Unpacking the "how" of cooperation, one option that immediately came to mention was the UN's role in setting international standards. A comparison often used was the Nuclear Non-Proliferation Treaty (NPT), as a successful example in which global governance managed to minimize the threat posed by the weaponization of new technologies, while enabling the peaceful sharing of the technology's benefits. However, many noted the difference between AI and nuclear technology—namely, the heavy involvement of private actors in the research, development, and

application of AI. Others, however, did not see this as a stark deviation from the NPT example. One panelist noted that cooperating with competitors even during periods of competition was not unprecedented and that even informal cooperation could be a useful option. As an example, the panelist referenced conferences that brought together nuclear scientists during the Cold War to discuss the need for safety practices.

The involvement of private actors complicates the ability to govern technologies that operate so far outside of national borders and national control, but their involvement may also open up the global-governance conversation to multi-stakeholder approaches. One panelist asked outright: Are nation-states up to this grand task of governing AI? The panelist argued we "do ourselves a disservice by framing this question in such grand and sweeping terms." Another panelist agreed that future inquiry on governance and policy should focus on specific applications of AI to avoid "a level of abstraction that prevents us from making any real progress."

(?)

QUESTIONS FOR FUTURE RESEARCH

How can domain-specific global governance of AI proceed?

What is society trying to solve with AI? To entrench democratic values? To prevent dissemination to bad actors?

How can society use this technology to advance human freedom but not empower authoritarian regimes?

ARTIFICIAL INTELLIGENCE AND HUMAN RIGHTS, SURVEILLANCE & DEMOCRACY

The final panel of the colloquium's expert discussions focused on AI's relationship to human rights, surveillance, and democracy, and addressed questions involving the human rights harms implicated by the misuse of AI—either from intentional misuse or from built-in biases—as well as the ability of existing enforcement mechanisms to regulate the movement and use of private data owned by powerful companies. The University of Pennsylvania Carey Law School's Christopher Yoo moderated the discussion among Zeid Ra'ad Al Hussein (Perry World House); Claire Finkelstein (University of Pennsylvania Carey Law School); Ellen Goodman (Rutgers University); Rebecca Hamilton (American University); and Helen Nissenbaum (Cornell University).

"CONTEXT DETERMINES EVERYTHING"

One panelist opened this segment by comparing the current state of the world to an ailing patient, arguing that innovators are introducing emerging technologies while this patient is already ill from stressors including climate change, the rise of populism, and human rights abuses. "The way I view this is the context determines everything," the panelist argued. If one views the world in the context of the expansion of democracy and human rights, then one might be optimistic about AI's future role in this system. However: "If you, like me, view the world [and see] that it is clear there is a sepsis in place, the introduction of AI may make things worse."

As an example of how new technology may pose a threat to existing human rights, one panelist noted the increasing challenge of striking a balance between supporting freedom of speech and constraining hate speech and expression that incites violence. One panelist pointed out that, rather than having a judge strike that balance, social media companies have implemented reporting tools that can result in hate speech being removed from their platforms. Here, an algorithm might determine what forms of speech ought to be protected or restricted. This occurs without a published opinion, often outside public knowledge. While panelists expressed support for the current pressure on social media companies to clamp down on the incitement of violence on their platforms, many recognized this as a double-edged sword: AI technology is a tool that can both harm and benefit human rights and democracy.

DETERMINING RESPONSIBILITY WITH AUTONOMOUS TECHNOLOGY

The panelists considered the existing international and domestic legal structures for ways to hold private actors developing AI accountable for unanticipated negative impacts of machine-learning technologies. In her written analysis for this colloquium, Ellen Goodman identified numerous areas in existing legal and regulatory systems that serve as useful templates for governing such technologies, including the products-liability regime in tort law and federal environmental and agricultural regulation. Goodman also highlighted in her written analysis ways in which social media and dating applications subvert accountability for social disorder and emotional harms through the free speech rights of the applications' users. This leaves tech powerhouses like Facebook, Twitter, Tinder, and others outside the realm of accountability and regulation

A lively conversation on AI and human rights closed out the afternoon's discussion.

historically relied on by Americans where the actions or oversights of corporations cause tangible harm to an individual.

Conversely, one panelist argued, the concerns expressed around autonomy and accountability with emerging technologies are overstated. The panelist said, "There is always a way to identify a responsible agent when autonomous systems cause harm—from accidents to war crimes—because the autonomy of the car, or the missile, or whatever, *does not break the chain of causation*." The panelists turned to a discussion of liability in tort law. In tort law, there is both the consideration of factual causation (also known as "but for" cause—as in "but for" the actions in question, the resulting injury would not have occurred), and proximate causation (which determines whether or not holding the actor accountable for a certain injury they caused would be fair under the circumstances). In the proximate cause analysis, there is a doctrine surrounding intervening actors (i.e., if a third party takes an unforeseeable, independent action contributing to an injury, it may break the chain of causation set out by the "but for" cause). To the panelist's point, here the autonomy of the car *does not constitute* an intervening actor.

Even where the danger and unpredictability of AI may cause the foreseeability of injuries to be difficult or impossible to determine, this panelist insisted, it did not matter. Using the example of keeping an exotic animal

as a pet, the panelist said, "If you keep a tiger as a pet and release it into the neighborhood, [no one would make the] argument that the tiger is an autonomous actor and therefore you are not responsible." The panelist argued this is even less applicable in the case of AI because, here, *"we have designed the tiger."* As the developers of AI have control over the design, and control over the amount of risk assumed when these products are made available, no liability issue may exist. Another panelist disagreed and argued that cyber technology is more akin to the multiple contributing factors doctrine, as when several factories are polluting a river. Through joint and several liability, multiple actors can all be given a share of the damages. The panel seemed to agree that the cyber globe is more akin to this than the wild animal case and concluded that a major problem in determining accountability of autonomous technology is managing the collective causal influences.

ALGORITHMS AS DECISION MAKERS

One panelist opened this portion of the discussion with an example of a loss in human rights abuse documentation that occurred because YouTube changed its algorithm. The algorithm was unable to distinguish between extremist content and content containing evidence of human rights abuses, and therefore deleted a video it concluded was the latter. The panelist argued this shows what happens as technology becomes ubiquitous:

Certain technological interventions make sense given the goals (as seen here, complying with EU regulatory standards), but are deeply problematic to those traditionally operating in this space. Another panelist responded by emphasizing the subjectivity of AI, as algorithms are created by individuals who bring their prior experiences, values, and viewpoints into every minute decision made in the programming process. Most often, those programming algorithms do not have experience in the policy sphere.

The discussion morphed into a question of cultural fluency in the development of AI, which opened with the example of Facebook introducing its platform in Myanmar. "You have a country that's seen five decades of military rule and a long history of persecution of the Muslim minority population," one panelist explained, "so many people in the field could have warned Facebook that social media could be used as a tool in this genocide." While recognizing that nobody at Facebook had this intention, the consequences of this lack of cultural competence led to Facebook being used as a tool for atrocities.

Returning to the panel's earlier discussion of the applicability of existing legal mechanisms—like tort law—to apply accountability in such situations, the panel again asked: Is it enough? Society needs to incentivize tech companies to employ social scientists, anthropologists, historians, etc.—individuals with multicultural competence—to prevent their complicity in atrocities. To this end, some panelists argued in favor of a strict liability model, which has been applied to ultra-hazardous activities such as blasting, toxic chemicals, and—notably—the keeping of wild animals. In the strict liability doctrine, determining liability does not involve an assessment of the amount of care exercised by the actor—for example, in the case of Facebook, *did* the company go to great lengths to hire experts and try to ensure its platform would not be used to further genocide? Instead, strict liability simply states that because Facebook's platform was in fact used to further persecution of religious minorities, there's sufficient cause for liability. Strict liability may be a useful solution to the accountability question resulting from algorithms acting as decision makers, as it incentivizes the maximum amount of caution and forethought on the part of the programmer, because they would be liable for any resulting harms.

QUESTIONS FOR FUTURE RESEARCH

What existing legal mechanisms are best suited to provide accountability for AI and other recent advances in technology?

How can companies ensure cultural competency in the development and deployment of AI? How can programmers take humans' misuse of technology into account?

How can AI be more effective in curtailing or preventing human rights abuses?

› DAY TWO

› CONVERSATION WITH JOHN KERRY, U.S. SECRETARY OF STATE (2013–2017)

Day 2 convened high-level leaders for a series of panels on the effects of emerging technology on the global order.

Day two of the colloquium brought world leaders together for a day of moderated conversations around the themes of emerging technologies and world order.[6] During an interview with *WIRED*'s editor-in-chief Nicholas Thompson, former U.S. Secretary of State John Kerry considered how new technologies will impact the United States and world politics.

TECHNOLOGY, CLIMATE CHANGE, AND FUTURE GENERATIONS

Thompson opened the interview by explaining that he had asked his young children to choose some questions for Secretary Kerry. Their concern? Climate change. Kerry argued that technological innovation, combined with political will, can help combat this global crisis.

Kerry spoke openly about his unique perspective on this conversation: He worked with Al Gore during his early research on climate change and served in Congress when legislators first discussed the issue. For nearly 40 years, Kerry has participated in debates on how to combat climate change, culminating with his role in bringing the United States to the table with China and other leading global powers to produce the Paris Agreement. For him, these past couple of years have completely reshaped the role the United States has to play in reentering the climate change effort. To Kerry, the Paris Agreement was really about sending a clear message to the private sector: "There are four to five billion consumers worldwide in this agreement, and hopefully

---

[6] Unlike day one, these sessions were open to the public and not held under Chatham House Rules.

that leads the next Bill Gates to tap into the money and innovation that kind of cooperation can support."

Kerry emphasized that multiple green sources of energy are now cheaper than coal. "We have a huge capacity to employ these technologies and we are not," he said. "We need to behave like we are in a war." In combating the adversary of climate change, Kerry explained, innovation is going to be key. He noted the Paris Agreement involved creating Mission Innovation, which called for mass investment in universities, researchers, tech firms, and other entities capable of developing new technology to combat climate change. He urged world leaders, including the America's own president, to convene leaders in the automobile industry, manufacturing industry, utility companies, and so forth and map out how green technology can be quickly incorporated into these industries.

> **"Bottom line, guys, the kids are in the streets asking adults to behave like adults, and they are absolutely right to do so. What we have to do is change the politics."**
>
> *— John Kerry on climate and political change*

TECHNOLOGY AND DEMOCRACY: FRIEND OR FOE?

Thompson, pointing out that the number of liberal democracies in the world is now declining and the number of authoritarian regimes is increasing, asked if technology was a force for authoritarianism. Kerry rejected this outright. "I do not believe that technology is causally related," he said. "And I believe we need to be clear about that. Technology is a tool. … It doesn't speak to you on its own; it doesn't organize. … It opened up a way for people to do what they've always done in new ways. … What it's done is magnified and sped up these processes, … and there has not been a sufficiently thoughtful counter to that." But what is to be done about authoritarian regimes' coopting of new technology to serve oppressive ends? Kerry favors implementing a similar approach as was taken in response to nuclear proliferation. "What I want to see is negotiations on the same level as nuclear," he said, "the same goals for writing the rules of the road, … accountability, transparency, and enforcement mechanisms to make sure people are adhering to what they're supposed to do." He called for greater accountability of social media and other tech companies to bear responsibility for where their platforms are misused.

WAR AND ARTIFICIAL INTELLIGENCE

Kerry also spoke about some military applications of AI, such as missile defense. Referencing the time, energy, and resources put toward nuclear proliferation during the Cold War, he noted a similar arms race could arise from the weaponization of AI. This could lead to trillions of dollars globally being put toward new weaponry that would only yield conflicts "so much more brutally destructive than anything we've experienced before." Unlike many speakers on this issue, however, Kerry has enormous faith in the strength of the diplomatic process and its ability to counter threats posed by the weaponization of new technology. Having seen diplomacy work effectively in relation to the nuclear threat, and (temporarily) again with the Iran Deal, he knows firsthand how successful it can be in staving off some of the greatest threats to peace.

"I'm optimistic because we are doing a lot of great things you're not hearing about. We're seeing the first generation of children in Africa being born without AIDS—that's us, we did that… We are capable of great things in this country. If we get involved in and take ownership of the process again, I have no reason to believe we cannot continue to do so."

— **John Kerry**, *U.S. Secretary of State (2013–2017)*
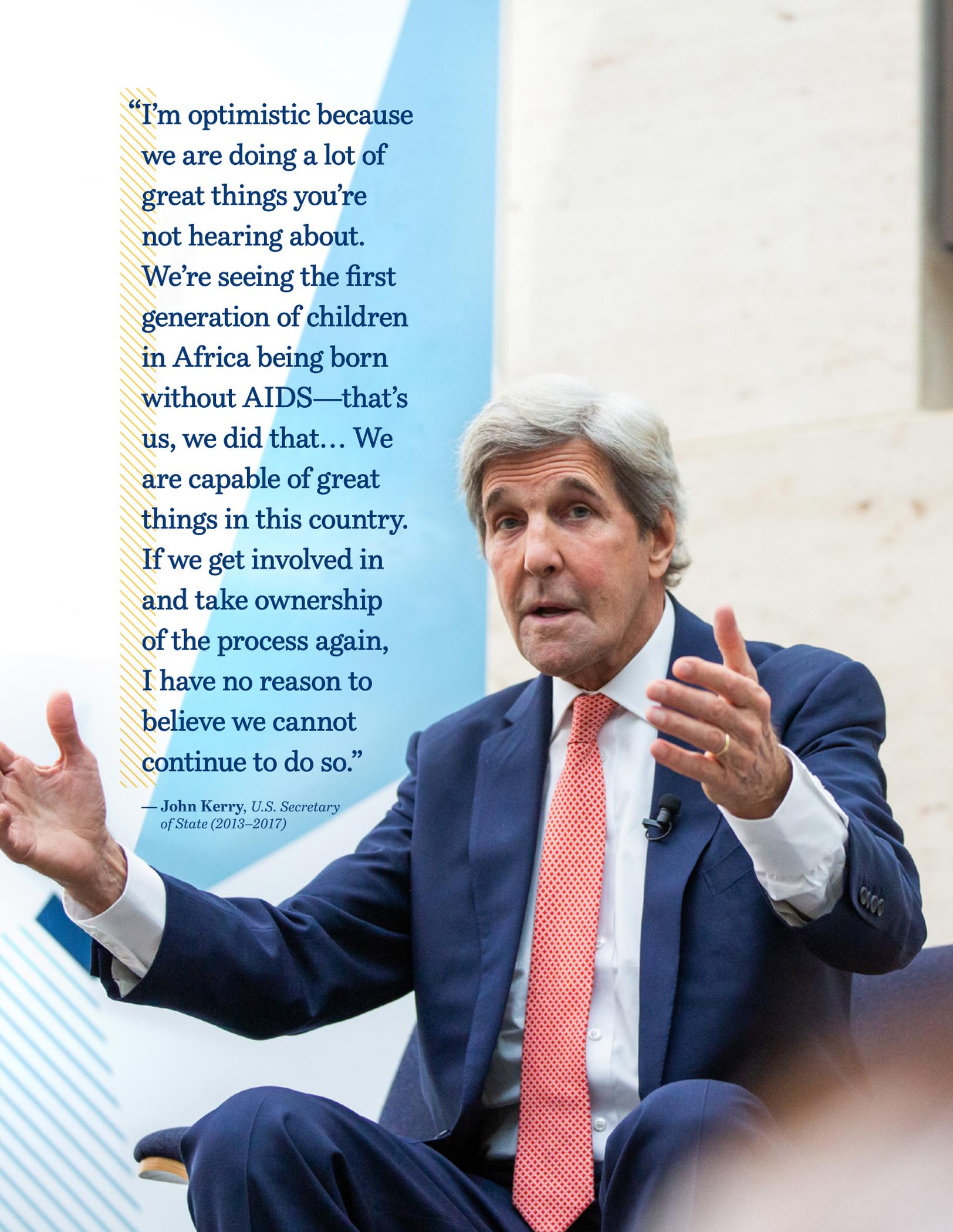
› CONVERSATION ON THE FUTURE OF TECHNOLOGY WITH RAQUEL URTASUN, CHIEF SCIENTIST, UBER ADVANCED TECHNOLOGIES GROUP

VISIONS FOR A SELF-DRIVING FUTURE

Raquel Urtasun began the conversation with a dynamic presentation about the AI applications in transportation that are making Uber a competitive leader in self-driving technology. Uber's recent AI innovations in self-driving vehicles hold the potential to improve transportation safety and access. Given estimates that over 90 percent of the 1.3 million annual road-traffic deaths worldwide are due to human error, Uber Advanced Technologies Group (ATG) investments rely on the bet that self-driving technology can reduce the chances for traffic accidents, increase mobility for populations with little access to transportation, and improve cities by decreasing pollution and road congestion.[7] Uber envisions self-driving vehicles and, perhaps one day, self-navigating air taxis becoming key pillars its business model and social value-added. When fully developed, these technologies could help transform transportation systems from ones that rely on personal vehicles to ones in which ride-sharing and public transport become the norm.

THE TECHNOLOGY UNDER THE HOOD

Three basic components make up the self-driving systems Uber is developing: (1) sensor-laden vehicles, (2) self-driving software systems built on AI, and (3) network connectivity. The "vehicles at scale" that Uber's system relies on are, in essence, robots with a complementary suite of camera and radar sensors. The 360-degree vision is a capability that adds something to a vehicle that goes beyond what a human can do, a possible opportunity to make autonomous vehicles a safer alternative to human-driven cars. Software systems represent the AI backbone of self-driving cars. Urtasun described how traditional systems operate by overlaying sensor data on two- and three-dimensional maps of an environment in a way that can track changes, use AI to predict how other cars and pedestrians will behave, and make plans that are adjusted as the process iterates repeatedly as a car moves along a trajectory. Yet, inherent inefficiencies in such systems

present safety liabilities that developers are working to overcome. Uber ATG relies on novel software that integrates data-processing steps into a single system and makes AI output more interpretable, thus making the self-driving system's predictions more speedily correctable when necessary. Finally, connectivity to networks link self-driving systems to the maps on which they rely. Generating comprehensive maps of roadways is a very costly enterprise: A onetime snapshot of every road in the United States would cost around $1 billion. Uber ATG is innovating more cost-efficient ways of building maps that could benefit other sectors, from the military and intelligence sector to geological surveying.

NAVIGATING AROUND THE ROADBLOCKS

In response to the technological progress Uber ATG has made, *Foreign Policy* magazine's Lara Seligman and colloquium participants engaged Urtasun with a provoking set of questions—and some skeptical critiques. Are fully self-driving cars as close to reality as we think? How safe can they really be? Is Uber doing enough to be transparent and facilitate effective government regulation of novel AI-based automobiles? Obstacles will surely litter any pathway to a future in which robotic cars dominate roadways—a future Urtasun believes is closer than someone like Missy Cummings, whose views were discussed earlier, does. Reasons for optimism over the eventual arrival of safe self-driving tech abound. For one, although errors will arise in any human-machine system, developers show signs they can learn from mistakes. The death of a pedestrian in Arizona who was struck by a self-driving Uber vehicle led to a sober review of the entirety of the corporation. The assessment identified ways Uber needed to better integrate safety concerns into its organization and its technical systems. As Urtasun argued, AI systems "need to adapt to people, not assume people will adapt to AI."

---

[7] World Health Organization, *Global Status Report on Road Safety 2015*, https://www.who.int/violence_injury_prevention/road_safety_status/2015/en/ (accessed October 8, 2019).

"Government needs to understand the technology before regulating it. But we in industry also need to play a role in education— the technology is evolving rapidly, so it can be hard to understand."

— *Raquel Urtasun, Chief Scientist, Uber ATG*

Beyond the technical solutions to the challenges of applying AI to human-dense environments, policy tools must also balance the dual goals of safety and risk-embracing innovation. Regulation will undoubtedly be part of ensuring innovators like Uber deliver the safer, more efficient future of transport that AI seemingly promises. Echoing observations made by the colloquium's other expert panels about the government's lack of AI understanding, Urtasun called on governments to make efforts to increase AI expertise within salient agencies to make regulation effective and efficient. In support of this goal and to build a broader base of public trust in its technology, Uber ATG publishes its safety-relevant research and makes its code publicly available to encourage dialogue among government, academia, and industry.

Urtasun spoke passionately on industry's role in explaining AI to the public and policymakers.

› CONVERSATION ON TECHNOLOGY & HUMAN RIGHTS WITH ZEID RA'AD AL HUSSEIN, PERRY WORLD HOUSE PROFESSOR OF PRACTICE OF LAW & HUMAN RIGHTS, AND KENNETH ROTH, EXECUTIVE DIRECTOR OF HUMAN RIGHTS WATCH

Kicking off the afternoon sessions, *NPR*'s Deborah Amos moderated a conversation between Zeid Ra'ad Al Hussein, the former UN high commissioner for human rights, a Perry World House visiting fellow, and the Perry World House professor of practice of law and human rights, and Kenneth Roth, the executive director of Human Rights Watch.

TECHNOLOGY AND PRIVACY

Privacy, Roth explained, comes up in different kinds of situations (from government surveillance to corporate data collection), and while the response to these situations may vary, it is in response to the same central questions: Who owns the data, and how should they be free to use it? The introduction of modern technology into this question adds another component: How much data is being taken in? Roth explained that cell phones and the data they constantly collect about their users' daily lives have changed expectations of privacy: "When you walk down the street, … theoretically, the government can follow you, but this requires teams of agents and is incredibly expensive. This has begun to change now that we carry around these tracking devices in our phones."

Technology may exacerbate existing divides in access to human rights protections. Nations that have strong track records in championing human rights and abiding by democratic and rights-based mechanisms will continue to do so and will expand those mechanisms to encompass threats to human rights posed by new technology. Both Al Hussein and Roth agreed that the next step must be to develop standards—but not in the form of a treaty. As Roth explained, with treaty-drafting, the result is "the lowest common denominator of standards—you have to compromise so much just to keep everyone at the table." As an alternative, he argued in favor of finding a forum where leaders can develop strong standards (such as the General Data Protection Regulation in Europe), create a doctrine that has moral authority, and enable those standards to gradually spread. By the end, even countries that would have never initially signed on to a treaty with those standards would end up complying with them just to stay in the market.

> **"This is an area where we have intuitions, but we have no laws. No rules whatsoever."**
>
> — *Zeid Ra'ad Al Hussein*

TECHNOLOGY AND FREE SPEECH

Discussing the challenges related to striking a balance between protecting free speech and policing hate speech and incitement to violence, Al Hussein emphasized the oversight problem when an algorithm or a single computer scientist makes these decisions, not a court. He said, "It is clear that Facebook is not dealing with this issue the way a court would be dealing with this issue." The process to take down posts on social media platforms, Al Hussein explained, differs from processes for traditional publications: "That decision is being made by an algorithm, about which we know very little, flagging content as hate speech, and then a young engineer perhaps is ultimately making the final decision of what stays available and what is removed. This is not a court or a judge deciding, and it can go horribly wrong."

Figure 6. On balance, do you think uses of AI are more likely to promote freedom and individual expression or aid in enabling human rights abuses?

FREEDOM & EXPRESSION

HUMAN RIGHTS ABUSE

■ 18%
■ 82%
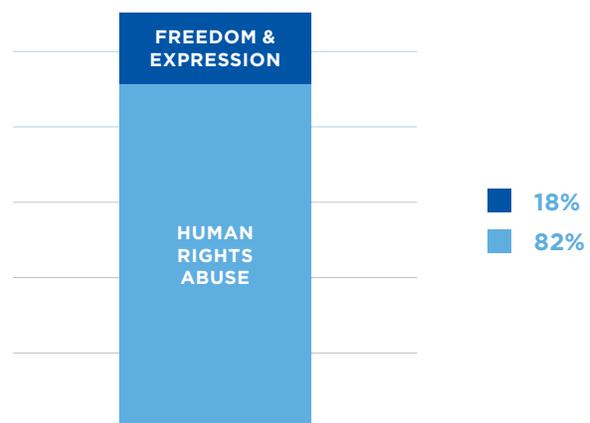
"Without that moral force, you have nothing. With that force, you can compel changes in government action," Kenneth Roth, Executive Director of Human Rights Watch, in response to a question regarding the influence of human rights on state behavior.

TECHNOLOGY AS A TOOL FOR HUMAN RIGHTS

The pre-colloquium survey revealed experts were pessimistic about AI's impacts on global governance and human rights. Only 18 percent of respondents predicted that AI would more strongly promote freedom and individual expression rather than limit them. Respondents' opinion was evenly split about whether threats to the social benefits of AI originate in the private sector or from the state. [8] Social media companies' roles in enabling or preventing state-led disinformation campaigns, surveillance, and nonconsensual use of private citizens' meta-data remain unclear.

Roth argued social media also can actively serve as a useful tool for advocacy, whistle-blowers, and activism. "We must not lose sight of that," he urged. As an example, Roth explained that social media and other new technologies have opened the door to numerous possibilities for remote open-source investigations with hundreds or even thousands of sources. Human rights investigators could use AI investigation tools and satellite imagery to find abuses, capture evidence, and thereby minimize the response time.

To capitalize on these potential benefits, however, Al Hussein was clear that governments must overcome other rising trends accompanying the spread of social media. "We are seeing fewer governments willing to stand up for human rights, even in the West," he said. While social media and other tech companies have the capability to aid in human rights causes, they have elected not to. Al Hussein was adamant that the people both can and *ought* to demand more of social media platforms. "They cannot continue to be agnostic about the ends for which their platforms are used," Al Hussein said, "and it is up to consumers to push these platforms toward being a force for good."

---

[8] Specifically, 27 percent of respondents agreed and 27 percent disagreed with the statement, "Governments are more likely than private companies to use AI for malicious ends, such as deep fakes, disinformation campaigns, or non-consensual surveillance." The remainder neither agreed nor disagreed.

› CONVERSATION WITH ROZA OTUNBAYEVA, FORMER PRESIDENT OF KYRGYZSTAN (2010–2011)

In an interview conducted by Trudy Rubin, a Perry World House visiting fellow and *The Philadelphia Inquirer's* worldview columnist, former President of Kyrgyzstan Roza Otunbayeva discussed her experience as the first female head of state in Central Asia.

GOVERNING LIKE A WOMAN

Trudy Rubin opened by asking Roza Otunbayeva what it was like to work as a female politician in a region that rarely sees one, and where male politics can be deadly. Otunbayeva described her presidency as a very purposeful, thoughtful undertaking, saying:

> *"We went against corruption, against family ruling—that was intentional. … We are learning from the globe how to fix the situation. … In three months, we put forward an agenda for a new constitution with a parliamentary government, we embedded such a group, who sit together, that is representative of our society, we promised to have parliamentary elections, and we promised that we will elect a new president, and that president will have a term of only six years."*

This agenda was strongly informed by the context in which Otunbayeva assumed power. There had been two very tumultuous presidencies preceding hers, and at the time she took office, Kyrgyzstan had just seen mass violence between two ethnic groups. Whoever assumed the role as the next head of state, Otunbayeva explained, could not be there for self-interested ambitions for power. Otunbayeva describes the rationale for electing her to the presidency as one of strategic foresight and a deeply felt need. She said: "The situation was absolutely unbearable, so these gentlemen had chosen me. … Because I was older, I was known in the international arena, and I was a leader of the opposition, a voice of opposition demanding what should be right and what is wrong. … So, logically, they chose me to lead us through this transition."

Otunbayeva discussed the challenges and opportunities of her presidency.

FUTURE DEVELOPMENT

What challenges does Otunbayeva hope to see her country tackle in the near future? Infrastructure. Over 90 percent of Kyrgyzstan's terrain is mountainous, with the population settling in the valleys, so the development of transportation systems like railways would transform the opportunities and resources available to Kyrgyzstan's citizens. Otunbayeva is hopeful this will be a priority for those in power. As an example of where she envisions Kyrgyzstan going, she hopes to use new technology to enhance communication and access to education across the country.

› CONVERSATION WITH ASH CARTER, U.S. SECRETARY OF DEFENSE (2015–2017)

MILITARY TECHNOLOGY: FROM SILICON VALLEY TO THE PENTAGON?

Since the end of the Cold War, the U.S. military has remained the largest and most capable in the world, but sources of international insecurity have also diversified. In his conversation with Jen Griffin of the *FOX News Channel*, former Secretary of Defense Ash Carter discussed how to keep up with contemporary threats, from transnational cyber threats to great-power military competition. "Having the best people and the best technology" is the maxim of effective stewardship over the massive pool of resources of the U.S. Department of Defense. Over the course of a long career at the Pentagon leading technology-centered projects, from preventing WMD proliferation out of the former Soviet Union to developing strategy for military AI, Carter sees clear changes in patterns of military technology research, development, and procurement. Major technological innovations now often flow from the private to public sector rather than in reverse, as during the Cold War. Now more than ever, having the "best technology" requires building bridges between Washington and Silicon Valley, where many actors are reluctant to work with the U.S. national security sector.

Cultural differences and a lack of dialogue between the U.S. military establishment and the tech industry have weighed on the pace and scope of collaboration. Carter is optimistic that overcoming these challenges is possible. During his tenure, intelligence surveillance programs revealed by Edward Snowden undermined the trust of many private actors considering collaborating with the U.S. government. That mistrust, he argues, hardly compares with the rancor young people felt for the military during the Vietnam War era, though. Still, governments need to work to interest young technologists in the mission of national security and public service. Carter described his work to convince engineers and business leaders in Silicon Valley to recognize the responsibilities they bear as experts in emerging technology with major implications for domestic and international society. Echoing the themes of his open letter to Google employees who protested participation in the Defense Department's AI research program Project Maven, Carter extended an invitation to public service to those who harbor mistrust or reluctance: "If you think we'll make mistakes, get in the game."

## LEADING WITH AN EDGE IN THE 21ST CENTURY

Looking to the future of technological change, Carter stated that it is not only necessary for the U.S. military to retain its technological edge, but also that it continues developing new, AI-based technologies consistent with American values and established norms governing the appropriate use of force. The public, too, must play an important role in demanding accountability from government to manage the effects—both good and bad—of emerging technologies. Instituting norms-based governance with respect to AI and data protection and privacy issues will have positive spillover effects, helping to mitigate the risks of other technological disruptions, as in biotechnology, that will mature in the coming decades. The only way to have sustainable commerce at scale for AI or biotechnology is to have trust alongside enforcement between individuals and businesses.

FRIENDS AND FOES IN FOREIGN POLICY

A high-stakes chessboard of international politics serves as the backdrop to U.S. efforts to enhance its national power with military innovation. Deterrence and political influence rely on more than capabilities, though. Carter emphasized the importance of maintaining a reputation and demonstrating values that attract and keep allies. Being able to gather a global coalition—including partners from long-time nation-state allies in Europe to Kurdish fighters in northern Syria and Iraq—was critical to the success of the U.S. mission to both defeat ISIS militarily and discredit the legitimacy of its ideology. Carter fears U.S. credibility, both among such allies and among adversaries, is eroding. Griffin asked if in response to Russian interference in U.S. elections, the United States has done enough. "No," was Carter's prompt reply. "Look at Putin. This is not a man who fears what would happen if he were to do it again. I'm a 'push-back' sort of guy. If you don't do that, … it gets worse." Ensuring adversaries remain certain of U.S. resolve to respond to acts of aggression becomes even more important as emerging technologies, from hypersonic missiles to cyber warfare, make international security relations increasingly uncertain and potentially unstable.

> "We keep our word. Being firm in that sense, keeping the values you and your allies have—that is what is important. These aren't signs of sentimentality, but sources of strength."
>
> — *Ash Carter, 25th U.S. Secretary of Defense*
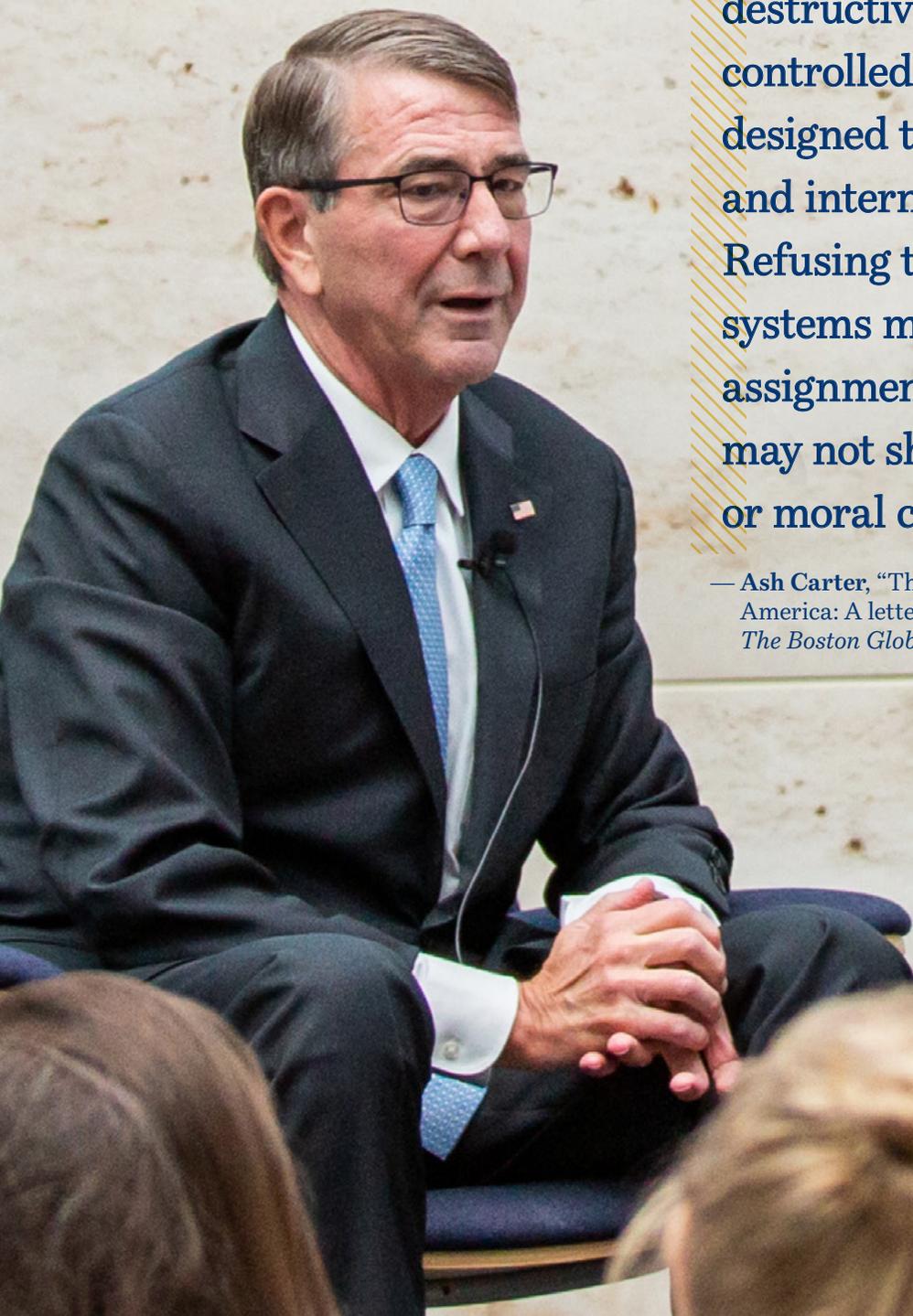
"Shouldn't people like you, who combine expertise with commitment to moral values, shape this tough arena? AI is an increasingly important military tool. Will it be crude, indiscriminate, and needlessly destructive? Or will it be controlled, precise, and designed to follow US laws and international norms? Refusing to help design these systems means ceding the assignment to others who may not share your skill or moral center."

— **Ash Carter**, "The morality of defending America: A letter to a young Googler," *The Boston Globe*, June 6, 2019.

› NEXT STEPS

The 2019 Global Order Colloquium advanced the conversation about emerging technologies and provided a call to refocus the attention of scholars and policymakers on the importance of both technological innovation and good governance of new technologies. Using the debates of the colloquium, participants' written analyses, and colloquium survey data, this section explains the key take-aways of the colloquium and makes three concrete recommendations for policymakers and scholars:

1. Prioritize international cooperation around emerging technologies by focusing on standard-setting and interoperability in both military and non-military realms.

2. Foster interdisciplinary research and education on AI applications, and embed AI experts in regulatory bodies.

3. Invest in the beneficial applications of new enabling technologies and focus on building digital public goods.

› POLICY AND RESEARCH RECOMMENDATIONS

PRIORITIZE INTEROPERABILITY IN EMERGING TECHNOLOGIES

To harness the competitiveness of global AI development, and to tap into the benefits of international cooperation, policymakers should prioritize interoperability and international standards for AI and data in both military and non-military realms. The pre-colloquium survey underscored the need for common standards. Among the experts surveyed, 86 percent said that AI generates more potential for competition than international cooperation. However, panelists recognized that advocates for international cooperation in AI governance will face an uphill battle. Pre-colloquium survey respondents said that AI will generate more potential for competition than cooperation.
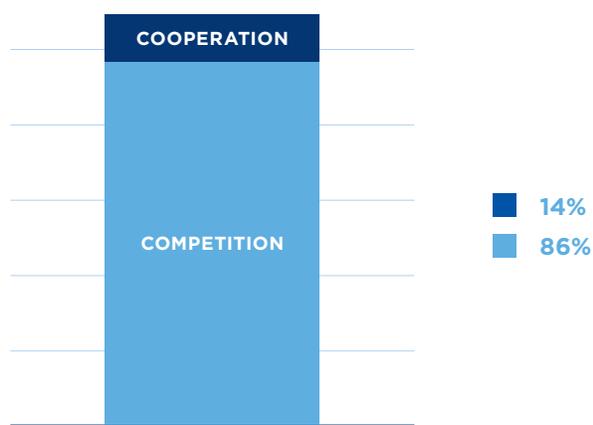

COOPERATION

COMPETITION

■ 14%
■ 86%

Figure 7. Do you view AI as generating more potential for international cooperation or competition?

Nonetheless, there are many examples where fierce competition occurs within a broader framework. As Ambassador Amandeep Singh Gill pointed out in his written analysis, civil aviation provides a model for success: Companies and states engage in cutthroat competition but still agree to a well-functioning distributed governance framework that allows travelers to jet around the world safely and easily (if not always in comfort). States could create international standard-setting bodies and working groups to agree on common terms and formatting guidelines.

The key to finding similar success in AI may lie in regional and international standard-setting. Colloquium participant Shawn Steene pointed to military aviation as another example where allies cooperate on complex technologies in his written analysis. In AI, the benefits of cooperation may be even greater; in addition to the usual win-win scenarios, cooperation may lead to greater access to pooled datasets, which may lead to better-functioning algorithms. To facilitate such cooperation among allies, Erik Lin-Greenberg noted in his written analysis that government needs to focus on interoperability on the format, storage, and labeling of data for machine learning. In national security, such cooperation may be limited to "trust networks," as Steene pointed out in his written analysis, but on less sensitive topics, standard-setting could be a major step toward more international cooperation and faster advances in technology. One possible implication is that NATO and other alliances prioritize interoperability at the regional level, while specific militaries and intelligence agencies seek to do the same for more sensitive data within "trust networks."

FOSTER INTERDISCIPLINARITY

Increased interdisciplinary research would help to uncover the most promising applications of AI and find creative solutions to AI's weaknesses. The 2019 Global Order Colloquium bridged divides among disciplines and between policymakers and the expert community. Many participants argued that such dialogue needs to be more common. Since AI is an enabling technology, its promise depends on its applications, which are best uncovered through interdisciplinary research. One difficulty of assessing "emerging technologies" is that their value depends on their proper application—drones can deliver both bombs and medical supplies, for example. An increased focus on interdisciplinarity—through research and policy fellowships and better public education, for instance—would help to reveal creative and productive applications and avoid the pitfalls of AI's weaknesses.

Within universities and machine-learning research centers, such interdisciplinarity could create better AI with baked-in values. However, interdisciplinarity is not just a goal for academia. In the policy world, the presence of technical experts in regulatory bodies would help to avoid panic and over-regulation, and would allow policymakers to understand the true issues of emerging technologies. In industry, embedding non-technical experts in the R&D process would help to embed good values and align software with societal goals.

Interdisciplinarity would also help to address the issues of bias and to foreground values debates. The future problem of global geo-technical divides and the gaps between the technological "haves" and "have nots" occupied much debate during the colloquium. The study and mitigation of these inequalities, however, will require the joint work of engineers, sociologists, policymakers, and many others.

INVEST IN DIGITAL PUBLIC GOODS AND STUDY THE BENEFITS OF AI

Several colloquium participants expressed dismay that most public conversations of AI focus only on the technology's dangers. Emerging technologies, after all, could improve human welfare dramatically, and future research and investment should look toward those benefits and the importance of preserving the public goods of technology.

To harness these potential benefits, states and other stakeholders will need to invest in "digital public goods." As Ambassador Amandeep Singh Gill explained in his written analysis, such goods could include digital identities to facilitate global access to financial systems or digital medical records to streamline and improve health care across the world. Opinion surveys suggest that the public is well-aware of these potential benefits—one 2018 U.S. opinion poll found that 76 percent of Americans believe AI will result in fundamental changes to society, and 77 percent of those believe the changes would be "mostly" or "very" positive.[9]

It is important that these three policy recommendations are deeply intertwined. Digital public goods, to be truly global, require greater international cooperation, as well as widely accepted standards on the formats of these goods. Well-functioning digital public goods, moreover, require algorithms that avoid discrimination and draw on clean and truly global datasets.

---

[9] RJ Reinhardt, "Americans Upbeat on Artificial Intelligence, But Still Wary," *Gallup.com*, last modified January 31, 2018, https://news.gallup.com/poll/226502/americans-upbeat-artificial-intelligence-wary.aspx (accessed November 12, 2019).

CONCLUSION

Although colloquium participants disagreed widely on the nature of AI's impact on the global order, all agreed that emerging technologies are likely to lead to significant disruption and require international cooperation. There is evidence that such cooperation is already underway on some issues related to emerging technologies. In the cyber domain, the Global Commission on the Stability of Cyberspace, the UN Group of Governmental Experts, and the Paris Call for Trust and Security in Cyberspace have all called for the protection of human well-being and digital public goods through norms to protect critical infrastructure and the "public core" of the internet, including the Domain Name System, from offensive cyber operations.[10,11]

It is important to remember, as Elina Noor's written analysis noted, that disruption from emerging technologies "may not necessarily be a bad thing." When guided in the right way, and with a human-centric approach that emphasizes cooperation, the rewiring of the global order could lead to a safer, more just, and more connected world. For further reading, the experts compiled the following list of resources.

WHAT THE EXPERTS ARE READING

Perry World House also asked participants to name one book or article that scholars and policymakers, in the United States or elsewhere, should read on AI or other emerging technologies. Here's what they recommended.

BOOKS

• Vannevar Bush, 2010 [1945], *Science, the Endless Frontier: A Report to the President on a Program for Postwar Scientific Research*, New York: American Council of Learned Societies.

• Ernest Davis and Gary Marcus, 2019, *Rebooting AI: Building Artificial Intelligence We Can Trust*, New York: Pantheon Books.

• Virginia Dignum, 2019, *Responsible Artificial Intelligence*, New York: Springer.

• David Edgerton, 2019, *The Shock of the Old: Technology and Global History Since 1900*, London: Profile Books.

• Kai-Fu Lee, 2019, *AI Superpowers: China, Silicon Valley, and the New World Order*, New York: Mariner Books.

• Paul Scharre, 2018, *Army of None: Autonomous Weapons and the Future of War*, New York: W.W. Norton & Company.

• Max Tegmark, 2018, *Life 3.0: Being Human in the Age of Artificial Intelligence*, London: Penguin Books.

• Shoshana Zuboff, 2018, *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power*, London: Profile Books.

ARTICLES

• Kate Crawford, Roel Dobbe, Theodora Dryer, et al., 2019, "AI Now 2019 Report," *AI Institute*, https://ainowinstitute.org/AI_Now_2019_Report.pdf.

• Richard Fontaine and Kara Frederick, "The Autocrat's New Tool Kit," *The Wall Street Journal*, March 15, 2019, https://www.wsj.com/articles/the-autocrats-new-tool-kit-11552662637.

• Michael C. Horowitz, 2018, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* 1(3), https://doi.org/10.15781/T2639KP49.

• Joshua A. Kroll, Joanna Huey, Solon Barocas, et al., 2017, "Accountable Algorithms," *University of Pennsylvania Law Review*, 165(633), https://scholarship.law.upenn.edu/penn_law_review/vol165/iss3/3.

• Heather Roff, 2019, "Artificial Intelligence: Power to the People," *Ethics and International Affairs*, 33(2): 127-140, https://doi.org/10.1017/S0892679419000121.

• Langdon Winner, 1980, "Do Artifacts Have Politics?" *Daedalus*, 109(1): 121-136, https://www.jstor.org/stable/20024652.

---

[10] Global Commission on the Stability of Cyberspace, "Final Report," *cyberstability.org*, last modified November 2019, https://cyberstability.org/report/ (accessed December 5, 2019).

[11] Paris Call, "Paris Call for Trust and Security in Cyberspace," *pariscall.international*, https://pariscall.international/en/ (accessed December 5, 2019).

PERRY
WORLD
HOUSE
UNIVERSITY of PENNSYLVANIA