Atlantic Council
SCOWCROFT CENTER
FOR STRATEGY AND SECURITY

POLICY PRIMER

AI, Society, and Governance: An Introduction

MARCH 2020

PETER ENGELKE

INTRODUCTION

After decades of largely unfulfilled promises, artificial intelligence (AI) has finally—and only recently—begun to demonstrate its massive power to reshape the marketplace, the public sector, the national security arena, and society more broadly. The technical developments that have occurred over the past decade, including new machine learning breakthroughs based on vast improvements in computational power and enormous increases in the quantity of data that can be used to "train" AI systems, have enabled the application of AI to an ever-wider range of sectors and activities. AI's increasing range of applications are having real-world consequences, both positive and negative. Those consequences, in turn, have animated spirited and at times emotional debates about how governments can craft policies to come to grips with a world increasingly shaped by AI.

With accelerating frequency over the past several years, public authorities, private sector firms, universities, and other organizations around the world have begun to address numerous AI-related policy questions. Within the public sector, AI policymaking is not limited to national governments. Rather, governments at every level—local, state/regional, and national governments in addition to multilateral institutions—have been grappling with AI-related challenges and attempting to craft policies to deal with them.

This newfound level of policymaking activity reflects how AI-driven applications have begun to affect nearly every dimension of human existence. In this arena at least, policy definitely has lagged behind technological advancement. Policymakers are now rushing to catch up.

As AI policymaking is so new, there is no accepted set of best practices. Experimentation is the order of the day. To illustrate, the first national AI strategy—developed by the Canadian government—did not appear until 2016. Since then, dozens of countries have raced to create their own AI strategies and policies, albeit with widely varying content, goals, mechanisms, and levels of funding.

AI policy, as one commentator puts it, can be defined as "public policies that maximize the benefits of AI, while minimizing its potential costs and risks."[1] This primer is intended to introduce and clarify AI policy across a wide range of policy domains. Although it is not exhaustive, it is intended to shed some light on the debates that have sprung up across these policy domains. It is intended for the layperson who may not be an AI expert, but who wants to better understand the central questions involved in AI policy debates.

1 Tim Dutton, "AI Policy 101: An Introduction to the 10 Key Aspects of AI Policy," *Medium*, July 5, 2018, https://medium.com/politics-ai/ai-policy-101-what-you-need-to-know-about-ai-policy-163a2bd68d65.

CATEGORIES OF AI POLICYMAKING

1: ETHICS AND NORMS

Many organizations have produced AI ethics and norms guidelines to place boundaries around the design of AI programs and their application to real-world phenomena.[2] A majority of such efforts have been created over the past few years.

The fact that so many organizations have endeavored to define the ethical uses of AI is itself testament to a fear that has long animated thinking about AI, involving its incredible power and how it might be used for harmful and unethical purposes. That fear has fired the collective imagination, inspiring countless science fiction books and films. Across a variety of other fields—scientific, technical, political, and humanistic—thinkers have expressed serious reservations about AI's potential for harmful application in the world we inhabit. Luminaries ranging from Henry Kissinger to Stephen Hawking have warned that AI someday will pose nothing less than an existential threat to humankind.[3]

Such fears of a machine takeover revolve around the development of "general AI," which is a term referring to AI systems possessing a kind of superhuman intelligence. General AI does not yet exist. What does exist is "narrow AI," which is AI applied to specific tasks. To clarify the distinction:

> "Artificial intelligence today is properly known as narrow AI (or weak AI), in that it is designed to perform a narrow task (e.g., only facial recognition or only internet searches or only driving a car). However, the long-term goal of many researchers is to create general AI (AGI or strong AI). While narrow AI may outperform humans at whatever its specific task is, like playing chess or solving equations, AGI would outperform humans at nearly every cognitive task."[4]

Research into general AI remains in its infancy. Researchers in the field estimate that general AI could be developed by the middle of this century or take far longer. However, given the stakes involved if general AI is ever developed, some experts advocate a prudent course starting now. They argue that governments should place safeguards around general AI's development today, long before it is perfected, in order to stave off a Terminator-esque scenario.[5]

Yet in the here and now, narrow AI is already demonstrating its enormous power. One response has been a proliferation of ethics statements regarding how narrow AI ought to be used. These statements tend to converge around a basic set of ethical propositions, for example, around the desirability of protecting human privacy, ensuring human oversight and control, and transparency. In 2017, as an example, the Future of Life Institute issued its Asilomar AI Principles, a list of 23 items focusing on how AI research ought to be conducted, what the objectives of AI systems should be, and how such systems can be made accountable.[6] The Institute is a high-profile body of scientists and ethicists dedicated to asking existential questions about emerging technologies.

In part to get ahead of the AI narrative, the world's biggest technology firms, including Google and Microsoft, have been among the most visible in the creation of such "ethical AI" guidelines. Google's AI principles, for example, list items such as accountability, safety, and a commitment to ensure that AI systems are unbiased. Such corporate efforts to create ethical

2 Darrell M. West, *The role of corporations in addressing AI's ethical dilemmas*, Brookings Institution, September 13, 2018, https://www.brookings.edu/research/how-to-address-ai-ethical-dilemmas/.

3 Henry A. Kissinger, "How the Enlightenment Ends," *Atlantic*, June 2018, https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/; Rory Cellan-Jones, "Stephen Hawking warns artificial intelligence could end mankind," *BBC News*, December 2, 2014, https://www.bbc.com/news/technology-30290540.

4 "Benefits & Risks of Artificial Intelligence," Future of Life Institute, https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/.

5 Future of Life Institute, "Benefits."

6 See "Asilomar AI Principles," *Future of Life Institute*, https://futureoflife.org/ai-principles/. "Asilomar" refers to a beach and conference grounds on California's Monterey Peninsula.

AI guidelines also have become among the most scrutinized, primarily because the world's largest tech firms are the global leaders in producing AI for their own for-profit purposes.[7] Google's efforts have fallen victim to such scrutiny: in 2019, the company scrapped an AI ethics board just one week into its existence after critics pounced on the board's composition.[8]

For governments, a common approach has been to create AI commissions. The activities of the European Union (EU) are especially noteworthy. In April 2019, the EU released its Ethics Guidelines for Trustworthy Artificial Intelligence, put together by a high-level expert group on AI.[9] The document spelled out a set of "fundamental rights" for "trustworthy AI," reflecting broadly shared norms that underpin European institutions. The rights include individual freedom, human dignity, democracy and the rule of law, and equality.

From these norms, the EU's expert group derived seven AI guidelines. AI systems should:

- be subject to human agency and oversight;
- be technically robust and safe;
- ensure privacy;
- be transparent (AI systems ought to inform people that they are interacting with an artificial system);
- enable diversity, non-discrimination and fairness;
- work in the service of societal and environmental well-being; and
- be accountable, including to external parties.

In June 2019, the same high-level expert group produced a second document recommending 65 policy and investment strategies for Europe. The recommendations range from upgrading workers' skills to placing limits on AI for surveillance purposes to using AI for improved public services to creating European AI-centric innovation ecosystems.[10]

One question is whether the EU will attempt to transition such principles into regulatory action, as it did regarding the digital economy with passage of the General Data Protection Regulation (GDPR). Having gone into effect in 2018, the GDPR has forced companies to prove compliance with the act in order to do business within Europe. Europe's AI ethics guidelines might be an important first step toward the issuance of AI regulations, with global effects similar to GDPR.

The act has had global effects. Other governments have modeled legislation or regulation after the GDPR, including Japan, which harmonized its data privacy regulations with European standards, and the state of California, which in 2018 passed the California Consumer Privacy Act (CCPA) and is now contemplating whether to align the CCPA more closely with the GDPR.[11] In both cases, these governments have been motivated in part by gaining an "adequacy determination" under the GDPR. An adequacy determination means that the EU would allow that country's firms to transfer their data from Europe to the home country (or state in California's case).

Critics argue that the GDPR's data privacy and transparency provisions are having an impact on European tech firms, specifically AI-related investment. For example, the GDPR requires firms to give individuals the right to a human review of a decision made by an automated (AI) system, which raises costs. Critics argue that, without reform, the GDPR will depress AI-related investment within Europe and shift even more of it to China and the United States.[12]

---

7 Sundar Pichai, "AI at Google: our principles," *Google*, June 7, 2018, https://www.blog.google/technology/ai/ai-principles/; "Microsoft AI principles," *Microsoft*, https://www.microsoft.com/en-us/ai/our-approach-to-ai; James Vincent, "The problem with AI ethics," *Verge*, April 3, 2019, https://www.theverge.com/2019/4/3/18293410/ai-artificial-intelligence-ethics-boards-charters-problem-big-tech.

8 Ed Adamczyk, "Google scraps new A.I. ethics board after member controversies," UPI, April 5, 2019, https://www.upi.com/Top_News/US/2019/04/05/Google-scraps-new-AI-ethics-board-after-member-controversies/2951554478709/.

9 European Commission, *Ethics Guidelines for Trustworthy AI*, 2019, https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top.

10 European Commission, *Policy and investment recommendations for trustworthy Artificial Intelligence*, 2019, https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

11 Andrei Gribakov, "Road to Adequacy: Can California Apply Under the GDPR?," *Lawfare*, April 22, 2019, https://www.lawfareblog.com/road-adequacy-can-california-apply-under-gdpr.

12 The Center for Data Innovation is a proponent of GDPR reform. See, e.g., Eline Chivot and Daniel Castro, *The EU Needs to Reform the GDPR to Remain Competitive in the Algorithmic Economy*, Center for Data Innovation, May 13, 2019, https://www.datainnovation.org/2019/05/the-eu-needs-to-reform-the-gdpr-to-remain-competitive-in-the-algorithmic-economy/.

Panelists debate AI ethics at an OECD conference. Paris, France, May 2019. Source OECD / Maud Bernos via Flickr

Pro-GDPR voices within Europe push back strongly against this criticism, arguing that the regulation will spur rather than inhibit innovation. Moreover, they observe, GDPR's primary purpose is not to enhance the AI market but to protect consumer rights.[13]

Other governments and multilateral institutions have crafted AI ethics guidelines that are similar to the EU's. The OECD's are among the most notable and recent. In May 2019, the OECD released five "complementary values-based principles" for responsible AI.[14] In June 2019, the G20 adopted its own set of principles that were drawn entirely from the OECD's principles.[15]

2: JUSTICE AND EQUITY

This policy domain is among the most visible and contested in the AI space. The dominant concerns revolve around whether AI systems reflect, reproduce, and even amplify society's problems. There is a running and often emotional debate about how to understand this challenge and, therefore, to build AI systems that would counter, prevent, or minimize such problems. This debate is especially fraught when it comes to AI systems' decision-making involving the human characteristics of gender, socioeconomic status, sexual orientation, and ethnic, racial, and religious status.

AI-infused systems produce decisions that, for a great many people, can be very real and non-trivial. Such systems, for example, can determine who has access to public and private resources and services, who is surveilled by state authorities, who is screened during hiring processes, what credit scores are assigned to consumers, and can shape how the police and courts interpret and enforce the law. Every one of these decision-making processes screens people in or out, or ranks them up or down, with positive and negative effects depending on selection status.

The justice and equity concerns involve, on the one hand, how AI tools are constructed and, on the other, how such tools are used by humans. In April 2019, the Partnership on AI (a nonprofit originally created by six large tech companies) released a report

13 See, e.g., Elizabeth Denham, "Letter: GDPR is showing clear promise as a modern law fit for the digital age," *Financial Times*, July 2, 2019, https://www.ft.com/content/87305816-9cb6-11e9-b8ce-8b459ed04726.

14 OECD, *Recommendation of the Council on Artificial Intelligence* (Paris: OECD Legal Instruments, May 21, 2019), https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

15 *G20 Ministerial Statement on Trade and Digital Economy* (Tsukuba City, Japan: G20 Trade Ministers and Digital Economy Ministers, June 8-9, 2019), https://trade.ec.europa.eu/doclib/docs/2019/june/tradoc_157920.pdf.

on the uses of AI in the criminal justice system in the United States. In a searing critique, the report's authors claimed:

> "There remain serious and unresolved problems with accuracy, validity, and bias in both the datasets and statistical models that drive these [AI] tools. Moreover, these tools are also often built to answer the wrong questions, used in poorly conceived settings, or are not subject to sufficient review, auditing, and scrutiny."[16]

Lack of validity signifies that an AI tool does not have "fidelity to the real world," which means the tool is applied out of context.[17] Bias refers to how AI tools can systematically err in their predictions, in particular for certain categories of people. Such AI bias can occur at the framing, data collection, and data preparation stages. During these stages, the AI system's designers can introduce biases into the algorithm and/or the training data, either deliberately or (most often) via blinders that prevent the designers from seeing how their efforts will bias the resulting analyses.[18]

There have been several high-profile cases involving what is sometimes referred to as "algorithmic bias" or "machine learning bias." Both Google and Amazon, for example, have suffered embarrassing revelations involving biases contained in their image search and hiring algorithms, respectively. Google's image search system was unable to accurately identify ethnic minorities. For Amazon, its hiring algorithm systematically assigned higher scores to males rather than females. Amazon's designers did not intend for this outcome to occur, but the AI system they built "learned" to select males over females anyway based on the algorithm's gender-imbalanced design parameters.[19]

Governments have begun crafting policies to combat algorithmic bias. In April 2019, two US senators crafted a bill, the Algorithmic Accountability Act, under which the Federal Trade Commission would require that firms screen their AI algorithms and training data for flaws leading to biased or discriminatory decisions.[20] Among other things, the Act would flag "high-risk" AI systems as those that include sensitive personal data, for example, data on a person's race, gender, sexual orientation, religion, genetic and biometric characteristics, and criminal background.[21] Its passage remains far from certain.

Taking a different approach, in March 2019 the British government announced that its Centre for Data Ethics and Innovation (created in 2017) and its Race Disparity Unit would cooperate on a research program to explore how AI systems can unfairly use ethnicity for decision-making within the United Kingdom's justice system.[22]

3: PRIVACY, CONSUMER PROTECTION, AND DATA AVAILABILITY

As with justice and equity, privacy and consumer protection issues are high-visibility and often fraught policy domains. AI systems glean insights about individuals from data collected from and about those individuals. Governments and firms, therefore, have access to an enormous amount of behavioral and biographical data, which describes a person's past, that is then is used by AI-based models to predict that person's behavior in the future.

AI adds a new dimension to the digital privacy policy debate, which involves questions regarding the type and volume of information actors have a right to collect from individuals and what can be done with that data. The EU's Ethics Guidelines for Trustworthy Artificial Intelligence, referenced above, includes a guideline devoted to privacy. Calling privacy a "fundamental right," the guidelines state that "adequate data governance [should cover] the quality and integrity of the data used, its relevance in light of the domain in which the AI systems will be deployed, its access protocols and the capability to process data in a manner that protects privacy."[23]

---

16 *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System, Partnership on AI*, April 2019, 10, https://www.partnershiponai.org/report-on-machine-learning-in-risk-assessment-tools-in-the-u-s-criminal-justice-system/.

17 Ibid, 14.

18 Karen Hao, "This is how AI bias really happens—and why it's so hard to fix," *MIT Technology Review*, February 4, 2019, https://www.technologyreview.com/s/612876/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/.

19 Tom Simonite, "When it Comes to Gorillas, Google Photos Remains Blind," *Wired*, January 11, 2018, https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/; Jeffrey Dastin, "Amazon scraps secret AI recruiting tool that showed bias against women," *Reuters*, October 9, 2018, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G.

20 Jack Corrigan, "Lawmakers Introduce Bill to Curb Algorithmic Bias," *Nextgov*, April 11, 2019, https://www.nextgov.com/emerging-tech/2019/04/lawmakers-introduce-bill-curb-algorithmic-bias/156237/.

21 Jerry Barbanel, "A look at the proposed Algorithmic Accountability Act of 2019," *IAPP*, https://iapp.org/news/a/a-look-at-the-proposed-algorithmic-accountability-act-of-2019/.

22 Cabinet Office, Department for Digital, Culture, Media & Sport, Home Office, Race Disparity Unit, and The Rt. Hon. Jeremy Wright MP, Investigation launched into potential for bias in algorithmic decision-making in society, press release, March 20, 2019, https://www.gov.uk/government/news/investigation-launched-into-potential-for-bias-in-algorithmic-decision-making-in-society.

23 European Commission, *Ethics Guidelines*.

AI-enabled facial recognition technologies exemplify the public debate concerning privacy rights. As has been widely reported, governments, firms, and other organizations increasingly are adopting AI-enabled facial recognition technologies. When married to video surveillance capabilities, facial recognition technologies allow people to be identified and tracked in real time, frequently without their consent or even knowledge. Facial recognition is being used in an increasingly wide variety of commercial and government applications, for example, to screen passengers at airports or identify and track people in city squares.[24]

Governments have begun to regulate their own adoption of this technology. Privacy concerns have been at the top of several local governments' ordinances in the United States that have banned facial recognition systems. Oakland, California, is the most recent to do so. Among other concerns, the advocates behind Oakland's ban argued that facial recognition technologies pose the risk of losing "the right to be anonymous in public, to freely associate."[25] Indeed, fears of "oppressive and continual mass surveillance" by governments and firms, as one report on facial recognition put it, provides much of the motivational force behind this slice of the AI privacy debate.[26]

Those fears are not unfounded, as China's use of facial recognition technologies already exemplifies. China famously is in the process of building and testing its social credit system, which aspires to monitor the behaviors of its citizens and reward or punish them accordingly. Facial recognition technologies, matched to near-ubiquitous video surveillance in public spaces, are key capabilities within this national social credit system.[27] Facial recognition's surveillance potential is being put to an even more serious test in the western province of Xinjiang, where the government is using the technology to monitor its Uighur population. There, facial recognition technologies have been integrated into a much larger and comprehensive effort to control Uighurs' entire lives.[28]

4: NATIONAL AI STRATEGY AND INDUSTRIAL POLICY

Governments are now developing national AI strategies and/or industrial policies that incorporate the commercial development of AI as a central objective.

National AI strategies represent a government's attempt to organize its thinking about AI and, therefore, to align its policy objectives and its stakeholders around a set of strategic objectives.[29] Industrial policy is not exactly the same thing. It is a malleable concept, but generally refers to direct and indirect state support for industrial development, particularly state support for targeted industrial sectors that are considered to be strategically vital for national economic prosperity and/or national security. A list of industrial policy tools might include creation of technology development funds, state-directed cluster investments (which are designed to foster geographic "clustering" effects), foreign tech transfer requirements (where foreign firms are required to transfer their technologies under certain conditions), targeted support for worker skills training, and more.[30]

While there is little to no controversy surrounding whether a government should create a national strategy focusing on a particular technology, ideological debates have surrounded the industrial policy concept for decades, particularly in the United States. Although the term "industrial policy" is, therefore, a loaded one, in practice governments implement a range of industrial policy-like policies across a range of sectors, the tech sector included.[31]

Regardless of nomenclature, national governments have been attempting to organize strategic approaches to AI. One study asserted that, as of July 2019, 41 countries have either produced national AI strategies or have demonstrated an interest in producing such strategies. Of the 19 states that have released framework documents, most (16) prioritized research and development into AI, followed by the development of AI

24 See, e.g., Madhumita Murgia, "London's King's Cross uses facial recognition in security cameras," *Financial Times,* August 12, 2019, https://www.ft.com/content/8cbcb3ae-babd-11e9-8a88-aa6628ac896c.

25 Sarah Ravani, "Oakland bans use of facial recognition technology, citing bias concerns," *San Francisco Chronicle*, July 17, 2019, https://www.sfchronicle.com/bayarea/article/Oakland-bans-use-of-facial-recognition-14101253.php.

26 Meredith Whittaker et al., AI Now Report 2018 (New York: New York University, AI Now Institute, December 2018), 4, https://ainowinstitute.org/AI_Now_2018_Report.pdf.

27 Karen Leigh and Dandan Li, "How China is Planning to Rank 1.3 Billion People," *Washington Post*, June 4, 2019, https://www.washingtonpost.com/business/how-china-is-planning-to-rank13-billion-people/2019/06/04/1cbdb2fe-86a3-11e9-9d73-e2ba6bbf1b9b_story.html?utm_term=.f031d338f802.

28 Isobell Cockerell, "Inside China's Massive Surveillance Operation," *Wired*, May 9, 2019, https://www.wired.com/story/inside-chinas-massive-surveillance-operation/.

29 This definition is adapted from Thomas A. Campbell, *Artificial Intelligence: An Overview of State Initiatives* (Evergreen, Colorado: FutureGrasp, 2019), 13.

30 *Perspectives on Global Development 2013. Industrial Policies in a Changing World: Shifting up a Gear* (Paris: OECD Development Centre, 2013), https://www.oecd.org/dev/pgd/COMPLETE-%20Pocket%20EditionPGD2013.pdf.

31 For an overview of industrial policy and its controversies, appeal, and limits, see Uri Dadush, *Industrial policy: a guide for the perplexed, Policy Center for the New South*, February 1, 2016, https://www.policycenter.ma/publications/industrial-policy-guide-perplexed#.VrJB1berTct.

talent (11), application of AI as part of an industrial strategy (9), development of ethical and legal guidelines (8), investment in AI-related infrastructure (6), and advancement of AI in government.[32]

Japan provides one example of such strategic thinking as applied to the AI domain. Its Industrialization Roadmap, released in 2017, clarifies how Japan should organize around AI and what kinds of investments the country ought to make.[33] The roadmap built on Japan's 2016 creation of a Strategic Council for AI Technology, designed to guide the government on AI questions. Japan aims to utilize AI in four priority areas: health, mobility, productivity, and information security.[34] The government appears to want to strengthen its AI hand given Japan's existing dominance in the robotics field. Japan both produces and consumes a large share of the world's most advanced robots.

Also in 2017, China's State Council released its own national AI strategy.[35] Calling AI a "strategic technology," the document articulated the need for China to possess world-class capabilities in AI research and industrial competitiveness by 2020 and to become "the world's primary AI innovation center" by 2030. China's strategy articulated the need to invest in AI research and development (R&D), "forcefully develop" new AI industries, and otherwise prepare society for swift AI adoption through the creation of "smart" sectors of every kind (health care, transport, cities, education, and so on.)

China's AI strategy thus focuses on becoming a world leader in AI through two key mechanisms: in the production of AI and related technologies on the one hand, and in society's ability to swiftly adopt and utilize AI on the other.

A European example is the UK's AI Sector Deal, which includes talent acquisition and talent development, scientific research and development, the creation of public-private partnerships focusing on AI, upgrading of the nation's digital infrastructure, and more.[36] The Sector Deal was part of a comprehensive industrial strategy produced by the UK government in 2017.[37] In the industrial strategy, the government named global leadership in AI as one of four priority "Grand Challenges" for the UK, alongside mobility, clean growth, and an aging society.

Although it is always difficult to compare the EU with national governments, given that the EU is a supranational entity, the bloc also has announced its strategic intentions regarding AI. In 2018, the EU released a pair of documents on its desire to lead Europe into an AI-defined future.[38] Together, these documents outline how the EU might lead a coordinated pan-European strategy centered on the creation of a common goal and vision, maximizing investment in AI, encouraging synergies across national AI platforms, exchanging best practices, and otherwise advocating for the treatment of AI in strategic terms.

In 2019, the US government released its own AI strategy via Executive Order, which emphasized R&D, standards, and workforce development. This document followed upon the Obama administration's 2016 release of a strategic R&D plan and a number of reports focusing on AI, society, and economy.[39] Also in 2019, three US senators proposed legislation that would create and, even more importantly, fund a national AI strategy. If passed, the Artificial Intelligence Initiative Act would allocate more than $2 billion over five years to fund a range of activities, including the establishment of an AI coordination office, several research centers, and an effort to create robust AI evaluation standards.[40]

---

32 Campbell, *Overview*, 13-14, 27-30.

33 *Artificial Intelligence Technology Strategy (Report of Strategic Council for AI Technology), Strategic Council for AI Policy*, March 31, 2017, https://www.nedo.go.jp/content/100865202.pdf.

34 "Japan: Fusing digital and physical," *MIT Technology Review Insights*, April 9, 2019, https://mit-insights.ai/japan-fusing-digital-and-physical/.

35 Graham Webster, Roger Creemers, Paul Triolo, and Elsa Kania, "Full Translation: China's 'New Generation Artificial Intelligence Development Plan' (2017)," *New America*, August 1, 2017, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

36 United Kingdom Department for Business, Energy & Industrial Strategy and Department for Digital, Culture, Media & Sport, "Policy Paper: AI Sector Deal," May 21, 2019, https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal.

37 United Kingdom Secretary of State for Business, Energy and Industrial Strategy, "Industrial Strategy: Building a Britain fit for the future," November 2017, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/730048/industrial-strategy-white-paper-web-ready-a4-version.pdf.

38 European Commission, Communication Artificial Intelligence for Europe, "Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions," April 25, 2018, https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe; European Commission, Coordinated Plan on Artificial Intelligence, "Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Coordinated Plan on Artificial Intelligence," December 7, 2018, https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence.

39 White House, "Executive Order on Maintaining American Leadership in Artificial Intelligence," February 11, 2019, https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/. For an overview of US AI policy, see Campbell, *Overview*, 26-27.

40 Khari Johnson, "US Senators propose legislation to fund national AI strategy," *Venture Beat*, May 21, 2019, https://venturebeat.com/2019/05/21/u-s-senators-propose-legislation-to-fund-national-ai-strategy/.

5: TECHNICAL STANDARDS

According to the International Organization for Standardization (ISO), a technical standard sets out "requirements, specifications, guidelines or characteristics that can be used consistently to ensure that materials, products, processes and services are fit for their purpose."[41]

By introducing common terminology and specifications, technical standards provide the foundations for markets to flourish. Specifically, technical standards create a floor upon which all manner of commercial activities can occur—allowing interoperability between technologies, enabling more rapid innovation, encouraging product differentiation, simplifying business contracts, and lowering trading costs.[42]

Broadly speaking, governments have a range of motives for creating (or encouraging adoption of) robust technical standards. On the domestic side, governments are interested in helping to create and scale promising domestic markets in new technologies. Technical standards help promote such markets. But governments also have an obvious interest in protecting the public good, to which technical standards contribute. Public health and safety considerations, for example, ranging from food to surgery to transport, in addition to many other areas of public concern, are paramount considerations in the formulation of technical standards.

Nowhere is the need for AI-related technical standards greater than in autonomous transport systems. Although there is now much activity, no country has created a full set of standards for autonomous systems. The US government, in keeping with its approach to technical standards in general, only supports the development of "stakeholder-driven voluntary technical standards" for autonomous vehicles, preferring to defer to nongovernmental organizations such as SAE International (formerly the Society of Automotive Engineers) to develop them.[43] Dozens of US states have considered or enacted initial

A fleet of Waymo's autonomous vehicles await further on-street testing. Tempe, Arizona, May 2018. Source zombiete via Flickr

legislation related to autonomous vehicles, several with the goal of creating their own (state-level) technical standards to enable autonomous vehicle testing.[44]

Germany is another matter altogether. In 2016, Germany's government established, under the federal transport ministry, the Ethics Commission on Automated and Connected Driving, charged with "develop[ing] the necessary ethical guidelines for automated and connected driving." The commission drew members from diverse fields, including the humanities, social sciences, law, business, the automotive industry and the tech sector. The Commission's report, released in 2017, contained 20 "propositions" about the use of autonomous vehicles. For example, that protection of human life must always have highest priority (over, e.g., damage to property).[45] The report appeared not long after the German government passed legislation that would allow automakers to test autonomous vehicles on roadways under specific conditions.[46]

Given the youthfulness of commercially viable AI technologies, the regulatory environment lags behind. Some industries have

41 "Standards," n.d., *International Organization for Standardization*, https://www.iso.org/standards.html.

42 US Department of Commerce, National Institute of Standards and Technology, *U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools*, 2019, 5; "Benefits of standards," European Commission, https://ec.europa.eu/growth/single-market/european-standards/policy/benefits_en.

43 US Department of Transportation, "Preparing for the Future of Transportation: Automated Vehicles 3.0," October 2018, 49, https://www.transportation.gov/av/3/preparing-future-transportation-automated-vehicles-3. SAE International's page on autonomous and unmanned vehicles is https://www.sae.org/automated-unmanned-vehicles/.

44 "Autonomous Vehicles: Self-Driving Vehicles Enacted Legislation," *National Conference of State Legislatures*, March 19, 2019, http://www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx#additional.

45 *Ethics Commission on Automated and Connected Driving* (Berlin: Federal Minister of Transport and Digital Infrastructure, June 2017), https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile.

46 "Germany adopts self-driving vehicles law," *Reuters*, May 12, 2017, https://www.reuters.com/article/us-germany-autos-self-driving-idUSKBN1881HY.

begun working on their own standards in order to shape the public sector's development of regulations.[47] Germany's auto industry, which has made heavy investments in autonomous capabilities, strongly favored the above legislation (passed in 2017) as it provided a regulatory go-ahead for the testing of vehicles in real-world conditions.

Governments also have global geo-economic motives for the creation and adoption of technical standards. They want to ensure that their firms' products are competitive in global markets (a firm can sell goods in a foreign market only when those products conform to that market's technical standards). Moreover, if a country's preferences are adopted in globally or regionally binding technical standards, that country's firms should have an enormous advantage. This is why the largest global powers—China, the United States, and the EU, in particular—increasingly see technical standards as part of a geo-economic and arguably even a geopolitical game. They view the international adoption of technical standards based on their own preferences as a key to global market power.[48]

Indeed, Europe, China, and the United States are pushing hard on this front. As a recent report issued by the Swedish Institute of International Affairs put it, China has identified technical standard setting as "an important angle for promoting and projecting its growing international power," by increasing its presence within international standard-setting institutions such as the ISO and through leveraging its own infrastructural investments via the Belt and Road Initiative.[49] In a December 2018 policy paper, the Chinese government expressed its desire to increase "exchanges and cooperation" with the EU regarding standards and to discuss "standardization issues of common interest to provide Chinese and European enterprises with timely, effective and authoritative information on standards."[50]

For its part, in expressing its strong interest in developing common standards with the United States, the EU has pointed out that although the United States and the EU have been the global "rule-makers" in common standard setting, both are now faced with the threat from "emerging, often heavily state-controlled economies," meaning China.[51] Regarding China, the European Commission has been uncharacteristically blunt:

> "The EU and the US should engage in a joint reflection on how to reinforce and deepen their cooperation in global standards setting given the increasingly visible ambition of certain third countries to influence this process to their own advantage. A good example is provided by China's ambitions in the 'Made in China 2025' sectors."[52]

6: INNOVATION ECOSYSTEMS

National and subnational governments alike have a strong interest in the creation of innovation ecosystems. The world's leading tech innovation ecosystems—cities and regions such as California's Bay Area—not only produce a large share of the world's applied technologies, but also create vast amounts of wealth for their residents and, by extension, the countries in which they sit. Given the money that will be made from AI-related technologies, policymakers are focused on ensuring that they can capture a share of this wealth creation within their boundaries.

Numerous public sector policies and investments, at national and subnational levels, support (or hinder) creation and growth of innovation ecosystems. Such policies and investments can overlap with the kind of policies and investments underpinning

47 "Daimler, BMW Partner to Develop Industry Standards for Autonomous Driving Technology," *Insurance Journal*, March 8, 2019, https://www.insurancejournal.com/news/international/2019/03/08/519987.htm.

48 This argument is advanced in, e.g., Alan Beattie, "Technology: how the US, EU and China compete to set industry standards," *Financial Times*, July 24, 2019, https://www.ft.com/content/0c91b884-92bb-11e9-aea1-2b1d33ac3271.

49 Björn Fägersten and Tim Rühlig, *China's standard power and its geopolitical implications for Europe, Swedish Institute of International Affairs*, February 2019, 3, https://www.ui.se/globalassets/ui.se-eng/publications/ui-publications/2019/ui-brief-no.-2-2019.pdf.

50 "Full text of China's Policy Paper on the European Union," *Xinhuanet*, December 18, 2018, http://www.xinhuanet.com/english/2018-12/18/c_137681829.htm.

51 European Commission, *Greater together: boosting transatlantic trade and addressing global challenges. Progress Report on the implementation of the EU-U.S. Joint Statement of 25 July 2018*, July 25, 2019, 6, http://trade.ec.europa.eu/doclib/docs/2019/july/tradoc_158272.pdf.

52 European Commission, *Greater together*, 11.

industrial policy, although public involvement in building innovation ecosystems typically is not classified as industrial policy.

The range of policies that are critical for developing AI innovation ecosystems include:

i Research and development support

Support for basic science—R&D—is a longstanding and widely utilized form of state involvement in the tech sector and forms a bedrock component of most innovation ecosystems. Those countries that make significant public investments in R&D also tend to be the world's leading innovators.[53]

State R&D support often takes the form of direct funding for research, frequently through universities and national labs, and sometimes via the creation of new research institutions. In the AI space, a recent example is the UK's creation of the Alan Turing Institute, which is the UK's "national institute for data science and artificial intelligence."[54] The institute has three goals: to "advance world-class research and apply it to real-world problems," including research leading to new businesses and employment; "train the leaders of the future" in data science and AI; and "lead the public conversation" around AI.

In the United States, in 2016, the Obama administration released a National Artificial Intelligence R&D Strategic Plan containing seven distinct research strategies.[55] The first strategy recommended that the federal government make a series of long-term investments in AI research, for example, in advancing "general AI," in creating more robust hardware to support AI, and developing more advanced robotics. Bloomberg has reported that in FY 2020 the US government will budget just shy of $5 billion for unclassified AI research and development.[56]

China has invested significant and rapidly increasing resources into AI R&D. Although it is impossible to determine precisely how much the United States and China are spending on AI R&D from all sources, the consensus within both countries is that China's spending is on pace to outstrip spending by the United States, and soon. A 2018 report by the US House of Representatives' IT subcommittee affirmed an earlier National Science Board/National Science Foundation assessment that China likely would surpass the US government in AI R&D spending by the end of that year.[57]

Given their rising R&D investments, the Chinese also are bullish about their capabilities. A 2018 report by the China Institute for Science and Technology Policy at Tsinghua University claimed that China has surpassed the United States in several AI research metrics, for example, in the number of first-tier AI research papers produced by Chinese scientists and in the number of AI-related patents produced by Chinese individuals, universities, and firms.[58]

ii Startup formation and scaling

Cultivating a robust startup economy is at the heart of any strong innovation ecosystem. Startup formation and scaling are core features of all the world's most innovative cities, regions, and countries.

Although the data on AI startups are incomplete, AI startups are formed most frequently in the same places as other types of tech startups. A 2018 study by Asgard Capital and Roland Berger, a German consultancy, found that North America, China, Europe, and Israel accounted for the bulk (83 percent) of the world's AI startups, a map that overlaps with the global tech startup map generally. Their data on AI startup cities revealed much of the same, with California's Bay Area, London, Tel Aviv,

53 This point is made at length in: Peter Engelke and Robert A. Manning, *Keeping America's Innovative Edge: A Strategic Framework, Atlantic Council*, April 2017, https://www.atlanticcouncil.org/publications/reports/keeping-america-s-innovative-edge; and Robert A. Manning and Peter Engelke, *The Global Innovation Sweepstakes: A Quest to Win the Future, Atlantic Council*, June 2018, https://www.atlanticcouncil.org/publications/reports/the-global-innovation-sweepstakes-a-quest-to-win-the-future.

54 The Turing Institute, https://www.turing.ac.uk/.

55 Executive Office of the President, National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee, *National Artificial Intelligence R&D Strategic Plan*, October 2016, https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf.

56 Chris Cornillie, "Finding Artificial Intelligence Money in the Fiscal 2020 Budget," *Bloomberg Government*, March 28, 2019, https://about.bgov.com/news/finding-artificial-intelligence-money-fiscal-2020-budget/.

57 US House of Representatives, Committee on Oversight and Government Reform, Subcommittee on Information Technology, *"Rise of the Machines: Artificial Intelligence and its Growing Impact on U.S. Policy,"* September 2018, https://fas.org/irp/congress/2018_rpt/hogr-ai.pdf.

58 Tsinghua University, China Institute for Science and Technology Policy, *China Artificial Intelligence (AI) Development Report 2018*, 2018, http://www.sppm.tsinghua.edu.cn/eWebEditor/UploadFile/China_AI_development_report_2018.pdf.

AI as business opportunity: panelists discuss how to position Canada as a leader in AI-driven innovation. Fortune Global Forum, Toronto, Canada, October 2018. Source John Lehmann/Fortune via Flickr

New York, and Beijing emerging as the biggest AI startup centers of activity. Although ranking cities by innovative activity is an inexact science, no observer of the global tech scene would be surprised to see this list of cities at or near the top of global AI startup rankings.[59]

This geographic overlap demonstrates how the world's leading tech innovation ecosystems almost certainly will be the centers of AI startup formation well into the future. Those ecosystems possess a collection of attributes that attract entrepreneurs, researchers, scientists, talented workers, and venture capitalists. Once such ecosystems are established in the global firmament of innovation ecosystems, then inertia becomes a friend, allowing replication of success due to the critical mass of people, institutions, and capital found there.

However, with smart policies and enough time to mature, new entrants can become important tech-based innovation ecosystems. Many of today's leading tech innovation ecosystems, including Tel Aviv, Stockholm, Singapore, Hyderabad, Berlin, and more, had low profiles in the startup space until not long ago.[60]

iii Talent: domestic workforce

The world's most innovative countries have deep pools of talented people who possess the necessary skills and capabilities to perform basic research, create viable technologies, find ways to match those technologies to markets, and otherwise apply the technologies, including digital tools, in the workplace. Any society that seeks to become a cutting-edge home for innovation around AI thus needs to find ways to attract such people from abroad and keep them within national boundaries, and/or create skilled workers from the domestic population.

Workforce development programs are designed to upskill a country's domestic workforce to enable it to take advantage of AI-driven technologies. The logic is twofold: (1) doing so will contribute to a country's competitiveness within the global economy; (2) doing so will help prevent citizens' skillsets from becoming obsolete as technology progresses. The objective is to build a labor force that possesses the skills necessary to compete in a world increasingly shaped by AI technologies and systems.

59 Roland Berger and Asgard, *Artificial Intelligence—A strategy for European startups,* 2018, https://www.rolandberger.com/publications/publication_pdf/roland_berger_ai_strategy_for_european_startups.pdf.

60 Histories and analyses of several of these ecosystems can be found in Manning and Engelke, *Sweepstakes*.

Experts diverge in their assessments of AI's future impact on employment and wages. Some argue that AI will destroy more jobs than it will create, especially in fields ripe for automation (trucking, retail, warehouses, call centers, and more), with impacts happening within the next decade. This argument rests partly on historical grounds, with a focus on how machines have displaced human labor many times since the dawn of the Industrial Revolution. Although machines frequently have opened up space for new categories of work and employment, this reading of history emphasizes the significant and widespread pain that such shifts have caused, with workers being exposed to wrenching and often lengthy periods of adjustment.[61]

One of the most prominent critics, Andrew Yang, argues that the most disruptive (negative) impacts will be on those with the fewest resources to reskill themselves: the poor and least well educated members of the workforce.[62] Yang, who recently ended his bid for US president, is one of a growing number of people in the United States and elsewhere who have embraced the universal basic income (UBI) concept, an idea that has found a significant following in Silicon Valley. UBI traces its roots back to the eighteenth century, but was first proposed in its modern form in the 1960s. UBI provides a minimum income floor for all citizens of a country, no strings attached, regardless of a person's economic status.[63] Although there have been a few policy experiments with UBI implementation, perhaps most notably by the Finnish government, the evidence is mixed as to its effectiveness.[64]

Although UBI's supporters have different motives, those like Yang are motivated most by a belief that technology—AI in particular—is set to replace a wide range of human capabilities. This downbeat forecast anticipates that AI-powered machines soon will become the most capable "workers" in the world, thereby making obsolete entire categories of human labor and, along with those work categories, the humans themselves. UBI is seen as a necessary corrective.

On the other end of the spectrum are those experts who argue that while AI will destroy some forms of labor, it will create many new employment opportunities. This argument rests upon a belief that AI will augment rather than replace human labor, making workers more productive instead of obsolete.

In this reading, the challenge is to find ways to ensure that people who are displaced by technology can reskill themselves to take advantage of the new categories of work that the same technologies help create, and to do so quickly. This argument, too, is partly historical, resting upon a happier interpretation of economic history that emphasizes the virtuous role that technology has played in augmenting, rather than diminishing or replacing, human labor.

Accenture's Paul Daugherty and Jim Wilson, authors of Human + Machine, are prominent voices in making this case. As the title of their book implies, their argument is that machines—AI-based systems—will allow human workers to do their jobs better, faster, and with greater personal fulfillment.[65] By removing much of the drudgery that is embedded in most jobs, they argue, AI will help free workers to focus their time on higher-value tasks that only humans can or should do, as with those tasks requiring human judgment or creative thinking.

Daugherty and Wilson, therefore, believe that AI will enable workers to become more productive, leading to a virtuous circle that includes more employment, greater wealth, and more happiness.

iv Talent: immigration

Tech-centric immigration policies focus on identifying, recruiting, and retaining skilled AI talent from abroad. Many governments believe that attracting foreign talent should form a critical piece of their overall strategies for global leadership in AI research and, therefore, in the commercial applications of AI. Immigration policy in this context largely focuses on attracting and retaining scientists and engineers who conduct

61 See, e.g., Carl Benedikt Frey, *The Technology Trap: Capital, Labor, and Power in the Age of Automation* (Princeton: Princeton University Press, 2019).

62 Kevin Roose, "A 2020 candidate sounds the alarm about robots and your job," Seattle Times, February 20, 2018, https://www.seattletimes.com/business/a-2020-candidate-sounds-the-alarm-about-robots-and-your-job/. Andrew Yang's book is Andrew Yang, *The War on Normal People: The Truth About America's Disappearing Jobs and Why Universal Basic Income Is Our Future* (New York: Hachette Books, 2018).

63 Stephen Mihm, "Why Legendary Economists Liked Universal Basic Income," *Bloomberg*, February 19, 2019, https://www.bloomberg.com/opinion/articles/2019-02-19/universal-basic-income-wasn-t-invented-by-today-s-democrats.

64 Emma Charlton, "The results of Finland's basic income experiment are in. Is it working?," *World Economic Forum*, February 12, 2019, https://www.weforum.org/agenda/2019/02/the-results-finlands-universal-basic-income-experiment-are-in-is-it-working/.

65 "Human + Machine: Reimagining work in the age of AI," *Accenture*, https://www.accenture.com/us-en/insight-human-machine-ai.

the basic and applied R&D leading to technical application in defense, health care, transport, and numerous other sectors.

For decades, the United States has successfully attracted high-skilled talent from abroad, forming a core reason why its tech innovation ecosystem has been the best in the world since at least 1945, if not earlier.[66] Recently, however, the US government has begun restricting H-1B visas, which admit high-skilled workers to the United States and is the mechanism through which the tech sector has long has relied on to attract such talent.[67]

As the United States does not produce enough domestic high-skilled talent to fill the tech sector's demand, restricting high-skilled immigrant talent appears counterproductive for building the United States' nascent AI-based economy.

Other countries have been heading in the opposite direction, in some cases spurred by a belief that the United States' current policies surrounding immigration mean that foreign tech talent may now be more interested in non-US destinations.

Canada provides an apt example. In 2017, it launched the Global Skills Strategy, a streamlined and simplified work visa process designed to incentivize skilled workers to choose Canada as a work destination.[68] Recent tech sector hiring trends suggest that Canada's open-door policy has been successful (relative to US performance) after the United States implemented restrictions on H-1B visas.[69] Regarding AI specifically, Canada also has implemented policies designed to attract highest-skill AI talent from abroad. The government's CIFAR Chairs in AI Program aims to "attract and retain the best AI talent to Canada," with the majority of chairs named thus far given to foreign researchers.[70]

v Intellectual property

The world's leading innovation ecosystems tend to exist within strong intellectual property (IP) regimes, meaning innovation occurs most in places that protect new ideas and the inventions that flow from them.[71] AI is in the process of upending the IP world, challenging basic assumptions and forcing IP governing institutions to adapt. IP policymaking related to AI is a brand new arena, with little in the way of formal public sector guidance.

AI challenges accepted definitions of who produces IP and, therefore, owns it, what types of AI-related materials and outputs constitute IP, and how AI-related IP infringements occur. As an example, although AI-related patent applications are rising globally, patent law does not protect AI data sets or compilations (for instance, the AI system's training data sets), which are fundamental components of AI systems. Tech firms that are interested in protecting their IP are having to find ways to fit their AI system features within existing law, which involves alternatively defining AI-related IP under patent, copyright, or trade secret classifications.[72]

The "who" question in IP law appears to be the most salient because it speaks directly to who is rewarded for AI systems' creations. Since AI systems create products based on their own "learning," the IP questions involve who should be given credit for a result that the AI system has produced. Would IP rights accrue to the AI system's owner, its programmers, or someone else? For example, AI systems soon will begin to produce their own creative works (music, etc.). When that occurs, will the IP belong to the artist(s) whose content was originally fed into the AI system and upon which the system crafted its own work, will it belong to the AI system's designers and owners, or to both?[73] Making this situation even more complicated is the fact that firms such as

66 Engelke and Manning, *Innovative Edge*.

67 Rani Molla, "Visa approvals for tech workers are on the decline. That won't just hurt Silicon Valley," Vox, February 28, 2019, https://www.vox.com/2019/2/28/18241522/trump-h1b-tech-work-jobs-overseas.

68 Immigration, Refugees and Citizenship Canada, Government of Canada launches the Global Skills Strategy, news release, June 12, 2017, https://www.canada.ca/en/immigration-refugees-citizenship/news/2017/06/government_of_canadalaunchestheglobalskillsstrategy.html.

69 Rani Molla, "Foreign tech workers are turning to Canada as US immigration becomes more difficult," *Vox*, June 7, 2019, https://www.vox.com/recode/2019/6/7/18653790/foreign-tech-workers-canada-immigration-indeed-trump.

70 Krista Davidson, "CIFAR expands Canada CIFAR AI Chairs Program to 46," *CIFAR*, April 7, 2019, https://www.cifar.ca/cifarnews/2019/04/08/cifar-expands-canada-cifar-ai-chairs-program-to-46.

71 For overviews, see Engelke and Manning, Innovative Edge, and Manning and Engelke, *Sweepstakes*.

72 "Protecting Artificial Intelligence IP: Patents, Trade Secrets, or Copyrights?," *Jones Day*, January 2018, https://www.jonesday.com/en/insights/2018/01/protecting-artificial-intelligence-ip-patents-trad.

73 See, e.g., "Artificial intelligence and intellectual property: an interview with Francis Gurry," *WIPO Magazine*, September 2018, https://www.wipo.int/wipo_magazine/en/2018/05/article_0001.html; Jonathan Weinberger, "Effective intellectual property rights protections are essential to advancing AI technology," *Hill*, March 14, 2019, https://thehill.com/blogs/congress-blog/technology/433957-effective-intellectual-property-rights-protections-are; Emma Woollacott, "Should AI own their own IP?," *Raconteur*, March 21, 2019, https://www.raconteur.net/risk-management/ai-ip-rights.

Google are creating AI support infrastructure (e.g., TensorFlow) to enable third parties (individuals, researchers, smaller firms) to more easily create their own models using open-source data found online, as, for example, via Google image searches. When someone creates an app using TensorFlow or other third-party provided inference architecture, who owns the IP?

7: CYBERSECURITY

AI complicates cybersecurity because it dramatically speeds up the problems and opportunities found in cyberspace. The cybersecurity industry views AI as a tool for combatting cybercrime and hacking, for example, through automation of threat detection and response. Its belief is that AI can perform these tasks more swiftly and more efficiently than humans or software. Yet, at the same time, these advantages can be applied by attackers as well, who could use AI to more rapidly discover and exploit software vulnerabilities, generate far more malware in order to overwhelm cyber defenses, and create more sophisticated and adaptive email-based scams (phishing, etc.). One fear is that hackers might successfully trick AI systems into categorizing malware as clean code, resulting in the mis-categorization of AI data and, therefore, producing flawed AI-generated program outcomes.[74]

A growing concern in this vein is "adversarial ML" (or "adversarial AI"), which occurs when attackers seek to disrupt, fool, or steal from an ML/AI system. There are several variants of adversarial attacks, for example, "model inversion" (where attackers acquire the training data used to train the system) and "model stealing" (where attackers seek to steal the model's underlying algorithm).[75] "Poisoning attacks," which insert inputs into machine learning models to trick the system into making categorical decisions (yes/no, in/out, etc.) that differ with the system's design parameters, are another common type of adversarial ML.[76] In poisoning attacks, unethical actors—frauds, cheats, thieves, hostile foreign governments, and so forth—manipulate how ML models assess inputs in order to reconfigure the resulting outputs in ways favoring the attacker. The poisoning attack introduces subtle changes to input data at the model's training stage, changes designed to be small enough to evade researchers' detection yet in aggregate significant enough to produce systematic deviation in the program's outputs. This "addition of a small amount of carefully engineered noise" can have real consequences, for example, a "duped" AI system in the medical field might order unnecessary medical treatments, constituting health insurance fraud.[77]

Adversarial ML is so new that policymakers, legal experts, and others are just now beginning to understand how and even whether existing law applies, and how legislation and policy might be changed to address this problem going forward. One review of US law found that although some types of adversarial attacks might fall under the jurisdiction of some statutes under some circumstances, hence might give the damaged party (the AI system owner or researcher) legal recourse against an identified attacker, for many types of adversarial attacks the application of existing law is far less clear.[78]

8: CRIME, LAW ENFORCEMENT, AND FRAUD AND DECEPTION

AI is a powerful new tool for both combatting criminal activity and for enabling it. On the combatting side, it is true that AI has enormous potential to assist law enforcement and court systems in the prosecution and prevention of crime. But at the same time, its use within criminal justice systems brings with it significant risks. On the enabling side, AI also will enable criminals to commit crimes, including entirely new types of crime, and to better hide their illicit activities.

The United Nations Interregional Crime and Justice Research Institute (UNICRI) and the International Criminal Police Organization (INTERPOL) are the leading multilateral institutions at the global level that are looking into this nexus of issues. In 2019, they released an overview report arguing that while

74 Martin Giles, "AI for cybersecurity is a hot new thing—and a dangerous gamble," ***MIT Technology Review***, August 11, 2018, https://www.technologyreview.com/s/611860/ai-for-cybersecurity-is-a-hot-new-thing-and-a-dangerous-gamble/.

75 Ram Shankar, "Law and Adversarial Machine Learning," *Medium*, December 20, 2018, https://medium.com/berkman-klein-center/law-and-adversarial-machine-learning-5c3badccea0e.

76 "The new cyberattack surface: Artificial Intelligence," *Accenture*, April 1, 2019, https://www.accenture.com/us-en/insights/artificial-intelligence/adversarial-ai.

77 Jonathan Shaw, "AI and Adversarial Attacks: vulnerabilities to manipulation," *Harvard Magazine*, January-February 2019, https://harvardmagazine.com/2019/01/ai-and-adversarial-attacks.

78 Ram Shankar Siva Kumar, David R. O'Brien, Kendra Albert, and Salomé Viljoen, "Law and adversarial machine learning," paper submitted to the 32nd Conference on Neural Information Processing Systems, Montréal, Canada, 2018, https://arxiv.org/pdf/1810.10731.pdf.

"many countries" are applying AI to law enforcement, there is a generally poor understanding of its effects and low coordination among law enforcement agencies across international boundaries. The report found that law enforcement agencies are using AI across a range of applications, for example, through creation of virtual autopsy tools, autonomous patrol robots, computer vision software, tracking and tracing systems, forecasting tools (predictive policing, crime hot spot analytics), and more.[79] UNICRI and INTERPOL also are leading a global conversation on the issue, with the second annual Global Meeting on Artificial Intelligence (AI) for Law Enforcement occurring in July 2019.[80]

Although law enforcement agencies are enthusiastic about using AI tools to combat crime, the use of such tools to counter criminal activity also creates the same concerns about justice, equity, data privacy, individual rights, and algorithmic bias as discussed at length above.

For example, AI creates the possibility of a Minority Report-style pre-crime world, wherein police and law enforcement possess the ability to better forecast not only where and when crimes are likely to occur, but also who is likely to commit them. Local governments in the United States have been using algorithm-based predictive policing tools for years. Controversy has followed the application of these tools, with critics charging (among other things) that they reinforce police biases against certain types of people and subject targeted communities to constant surveillance.[81] The fear is that human decision-making biases are simply replicated by algorithmic ones, one buttressed by the fact that AI-based decisions often are not transparent to the outside observer.

Likewise, in the UK, local police are experimenting with an AI-based program called the National Data Analytics Solution (NDAS), which attempts to assess the likelihood of a person committing a crime or becoming a victim of one. Although the police departments involved insist that the police would take no action against any individuals identified by NDAS as at risk of committing a crime (the individuals would be contacted by social services for counseling), as in the United States, there are clear rights issues involved in the application of such tools.[82]

The application of AI to policing overlaps with the surveillance and privacy debate as discussed in the above paragraphs.

One of AI's more frightening dimensions is its malicious use by criminals (both individuals and organized criminal networks), hostile foreign governments, hackers, and others. The wide range of malicious uses of AI include not only the expected cyberattacks and cyberthefts, but also an entirely new set of criminal tools that AI systems create.

Among the latter are increasingly cheap and widely available AI-driven video and audio spoofing tools that enable malevolent actors to manipulate and mimic human voices and imagery. Although these tools do not yet perfectly mimic a person's image or voice, experts believe that such perfection will occur not long in the future. When that occurs, any person could have exact or near-exact replicas of their voice, face, and body posted online as if it were a real photo, video, or audio recording of themselves. Such tools might be used to target individuals directly, for example, a person might have their voice and image used against them, whether for blackmailing purposes or as part of a financial scam. A major problem will involve disentangling the real from spoofed imagery, as even experts in this sub-field anticipate that they will have difficulty distinguishing between authentic and fake.[83]

There is rising awareness about the threats posed by these tools, both as criminal threats against individual people and as political threats against public officials. Within the United States, policymakers have begun expressing their concern that "deep fake" videos and images, created by hostile foreign nations as

79 United Nations Interregional Crime and Justice Research Institute (UNICRI) and the International Criminal Police Organization (INTERPOL), *Artificial Intelligence and Robotics for Law Enforcement*, 2019, http://www.unicri.it/news/files/ARTIFICIAL_INTELLIGENCE_ROBOTICS_LAW%20ENFORCEMENT_WEB.pdf.

80 United Nations Interregional Crime and Justice Research Institute (UNICRI), "2nd INTERPOL—UNICRI Global Meeting on Artificial Intelligence for National law enforcement agencies, private sector and academia," July 3, 2019, http://www.unicri.it/news/article/UNICRI_INTERPOL_Artificial_Intelligence.

81 Caroline Haskins, "Dozens of Cities Have Secretly Experimented with Predictive Policing Software," *Vice Motherboard*, February 6, 2019, https://www.vice.com/en_us/article/d3m7jq/dozens-of-cities-have-secretly-experimented-with-predictive-policing-software; Jonathan Capehart, "How your data is used by police, and where it goes wrong," *Washington Post*, July 17, 2018, https://www.washingtonpost.com/blogs/post-partisan/wp/2018/07/17/how-your-data-is-used-by-police-and-where-it-goes-wrong/?utm_term=.12bc5db6831f; Mark Puente, "LAPD to scrap some data programs after criticism," *Los Angeles Times*, April 5, 2019, https://www.latimes.com/local/lanow/la-me-lapd-predictive-policing-big-data-20190405-story.html.

82 Chris Baraniuk, "Exclusive: UK police wants AI to stop violent crime before it happens," *New Scientist*, November 26, 2018, https://www.newscientist.com/article/2186512-exclusive-uk-police-wants-ai-to-stop-violent-crime-before-it-happens/.

83 John Markoff, "As Artificial Intelligence Evolves, So Does Its Criminal Potential," *New York Times*, October 23, 2016, https://www.nytimes.com/2016/10/24/technology/artificial-intelligence-evolves-with-its-criminal-potential.html.

well as other malevolent actors, will flood the digital landscape ahead of the 2020 election.[84]

Besides the potential impact that deep fakes could have on individual politicians and election cycles, the larger concern involves the integrity of the democratic process itself. Further erosion of the distinction between an objective, empirically valid reality on the one hand and fabricated information on the other—a distinction already seriously compromised by the existing digital landscape—will deliver yet another blow to the democratic ideal of a citizenry capable of making well-informed, hence rational, decisions at the ballot box.

9: PUBLIC SECTOR EFFICIENCY AND EFFECTIVENESS

A burgeoning field of policy involves AI and the public sector, specifically regarding how governments can benefit from the systematic application of AI to their own processes. There is high enthusiasm regarding AI's positive impacts on the sector. Government operations are ideal for AI application because they involve standardized procedures based on legal requirements that (most often) change infrequently. AI can churn through these kinds of processes faster and with fewer errors than can human workers, reducing case backlogs, increasing customer satisfaction, more readily identifying fraud and abuse, and reducing costs.[85]

In addition, AI can assist with other important government functions. Regulatory oversight is one of these. Factory and worksite inspections can be made more common, effective, and cheaper via the application of AI-integrated technologies (sensors plus remote video in addition to AI). Moreover, and just as critically, AI can generate predictive analytics to forecast when and where serious adverse incidents might occur, for example, a mechanical breakdown at a power plant, a pathogenic outbreak at a meat packing plant, or a chemical spill at a factory.

As with the labor augmentation logic, a common argument is that AI will reduce the drudgery involved in processing government operations, in effect freeing government workers to focus on cases where humans must make the difficult decisions.

Although governments at all levels have been experimenting with AI-based applications over the past several years, none has been more aggressive than the United Arab Emirates (UAE) in embracing AI for public sector reform purposes. In 2017, the UAE created a State Minister for Artificial Intelligence and gave the position to Omar Bin Sultan Al Olama, who was at the time all of 27 years old. Among other objectives, his ministry's goals include the swift adoption of AI technologies within the government in order to improve public services, increase public sector efficiency, and help drive the UAE's global economic competitiveness.[86]

10: COSTS: FINANCIAL AND ENERGY

A barely noticed but important question involves the high costs of AI research, which both limit who can participate in development of AI systems and how much in the way of resources—money and energy, specifically—are required to conduct the research. A June 2019 paper published by researchers at the University of Massachusetts found that training large AI models is costly in terms of both resources. Although there are many AI models and there are wide variations in the amount of money and energy required to run them (money is needed for cloud computing, primarily), the researchers found that such costs grew exponentially with model complexity.[87]

In a summary of the paper, the MIT Technology Review observed that the "computational and environmental costs of training [the models] grew proportionally to model size and then exploded when additional tuning steps were used to increase the model's final accuracy." The most complex model

84 Cristiano Lima, "'Nightmarish': Lawmakers brace for swarm of 2020 deepfakes," *Politico*, June 13, 2019, https://www.politico.com/story/2019/06/13/facebook-deep-fakes-2020-1527268.

85 *Artificial Intelligence Unleashed: How agencies can use AI to automate & augment operations to improve performance* (Arlington, Virginia: Accenture Federal Services, 2018), https://www.accenture.com/_acnmedia/PDF-86/Accenture-Essential-Insights-POV.pdf#zoom=50.

86 Cynthia Johnson, "How the UAE's New Minister of AI Views the Future of Tech in His Desert Nation," *Entrepreneur*, February 20, 2018, https://www.entrepreneur.com/article/308709.

87 Emma Strubell, Ananya Ganesh, and Andrew McCallum, "Energy and Policy Considerations for Deep Learning in NLP," (paper submitted to the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, June 2019), https://arxiv.org/abs/1906.02243v1.

examined by the researchers cost between $940,000 and $3.2 million to run (in cloud computing costs) and generated 284 tons of CO_2. The latter figure is roughly five times the amount of CO_2 generated by a car over its entire life cycle (including the energy required to build the car) and 17 times the annual amount of carbon generated by the average American.[88]

In its own review of this study, the Financial Times asserted that besides the massive carbon footprint produced by AI training models, "the resources and costs involved in conducting research into machine learning [threaten to] shut out many academic researchers. That leaves it to the big tech companies that provide cloud computing (i.e., Amazon, Microsoft, and Google) and therefore have access to vast computational resources."[89]

As of yet, there are no policy experiments, nor even serious policy conversations, regarding how to grapple with this set of issues.

11: AUTONOMOUS WEAPONS AND LETHAL FORCE

A final set of policy questions rests squarely within the national security arena. These involve AI and lethal force, specifically the development of "lethal autonomous weapons systems" (LAWS) and their use on the battlefield.

For the major world powers' national security apparatuses, the highest policy priority is simple: in the absence of binding global limits on autonomous weapons systems (arms control treaties), the most important priority is having access to sufficient resources to develop, test, and deploy AI-integrated systems so as to ensure superiority over one's geostrategic and military rivals. Regarding the uses of AI, their dominant concerns involve the military chain of command: who (or what) issues a kill order on the battlefield, or at the very least who in the chain of command makes "decisions about how, when, where, and why the weapon will be employed"?[90]

For everyone else in the world, there is a greater range of relevant policy questions, up to and including whether anyone anywhere ought to possess such weapons.

The same ethical and humanitarian concerns about AI development in general animate the global debate about autonomous weapons. Kissinger, as an example, has argued that it might be impossible to control AI-based weapons systems, including LAWS, given the lack of transparency surrounding their development (transparency is the basis of all arms control).[91] His remarks were intended as a dark warning about the grim consequences that might result if AI-based weapons systems are allowed to be developed without constraint.

The UN has taken the lead in organizing a global debate concerning such questions. In 2013, under the auspices of the Convention on Certain Conventional Weapons (CCW), the UN began convening an informal expert body to discuss emerging issues, including ethical and humanitarian challenges, presented by the specific threats that will be posed by LAWS whenever such systems are developed. In 2017, this group was replaced by a more formal structure, the Group of Governmental Experts (GGE), which convenes large meetings attended by dozens of state and non-state actors.[92]

Civil society is heavily involved within the GGE's deliberations, enjoying participatory status. The groups include nonprofits, universities, think tanks, and other organizations. This breadth of civil society representation reflects widespread concern about a future world in which autonomous weapons exist and are used. The name of the most prominent nonprofit involved, the Campaign to Stop Killer Robots, is itself a case in point.[93]

GGE's most recent informal meeting occurred in late June 2019 and focused on themes such as ethics and humanitarian considerations, the "human-machine interaction in the development, deployment and use" of LAWS, the military applications of LAWS, and more.[94]

---

88 The most expensive model was titled "Transformer (213M parameters) w/ neural architecture search." See Karen Hao, "Training a single AI model can emit as much carbon as five cars in their lifetimes," *MIT Technology Review*, June 6, 2019, https://www.technologyreview.com/s/613630/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/.

89 Jemima Kelly, "AI: not so benevolent after all," *Financial Times*, June 20, 2019, https://ftalphaville.ft.com/2019/06/20/1561003239000/AI--not-so-benevolent-after-all-/.

90 Kelley M. Sayler, *Defense Primer: U.S. Policy on Lethal Autonomous Weapon Systems* (Washington: Congressional Research Service, March 27, 2019), https://fas.org/sgp/crs/natsec/IF11150.pdf.

91 Anshula Gandhi, "Henry Kissinger speaks at College of Computing celebration," *Tech*, March 7, 2019, https://thetech.com/2019/03/07/kissinger-talk-college-of-computing-celebration.

92 "Background on lethal autonomous weapons systems in the CCW," the United Nations Office at Geneva, https://www.unog.ch/80256EE600585943/(httpPages)/8FA3C2562A60FF81C1257CE600393DF6?OpenDocument.

93 Campaign to Stop Killer Robots, https://www.stopkillerrobots.org/.

94 "Background," the United Nations Office at Geneva.

Notably, the world's greatest military powers (assuming these are the United States, China, and Russia) appear to be heading in the opposite direction as the UN system. Whereas the conversation within the UN system is about the threats posed by anyone possessing and using LAWS, the world's great military powers are singularly focused on developing these technologies, given their rivals' desires to do the same.[95]

All three powers have been investing huge sums of money into AI systems, including LAWS, the result of which is that we are living through a kind of global AI arms race with no limits on what can be developed, deployed, and used on the battlefield. Those investments in turn reflect confidence within each of the powers' national security communities that developing their military's AI capabilities will give them technical advantages over their strategic rivals. Conversely, all appear convinced that not doing so will amount to a form of unilateral disarmament vis-à-vis their strategic rivals.

In addition, the major powers believe that greater AI capabilities will have other military benefits. For example, autonomous systems might reduce the number of personnel who are exposed to dangerous conditions in war zones (as an example, autonomous vehicles running supplies through hazardous territory). AI also will introduce other efficiencies, for example, through dramatically increasing the power of real-time battlefield analytics.

The US government does not prohibit the development and use of autonomous weapons systems by its military, and there is no statutory guidance or restriction on semiautonomous or fully autonomous systems. A 2019 Congressional Research Service paper on AI-based warfighting systems, including LAWS, asserted that the US military is in control of the development and utilization of such technologies.[96] An initial governing mechanism appeared in 2012, when the US Department of Defense established policy via a departmental directive ("Autonomy in Weapon Systems").[97] Among other things, the directive requires that all autonomous and semiautonomous systems "allow commanders and operators to exercise appropriate levels of human judgment over the use of force," that any such system be tested and evaluated for functionality, that any operators using such systems be fully trained, and that a senior-level departmental review must occur before any such system becomes operational. In February 2020, the department issued a set of ethical principles governing the military uses of AI. These principles are consistent with the 2012 policy, stressing that the department's AI systems be "responsible, equitable, traceable, reliable, and governable."[98]

In 2018, the Department of Defense (DoD) issued an AI strategy that outlined "strategic focus areas," including the use of AI to improve analytics and decision-making, increase operational safety, improve logistics and maintenance, and improve business processes. The DoD also established a Joint Artificial Intelligence Center to accelerate, scale, and synchronize the department's AI efforts.[99]

Regarding LAWS, the Congressional Research Service paper states flatly that "the United States is not currently developing LAWS," a claim that stretches credibility.[100] The DoD's 2018 AI strategy makes no reference to its own development of AI-enabled weapons systems, aside from some initial passing references. The DoD has attempted to channel the public's ethical concerns about LAWS development through a public comment process under the aegis of the Defense Innovation Board, an advisory body.[101]

---

95 Kendrick Foster, "The Modern Pen and the AI Sword," *Harvard Political Review*, May 13, 2019, https://harvardpolitics.com/united-states/pen-ai-sword/.

96 Sayler, *Defense Primer.*

97 US Department of Defense, *Department of Defense Directive No. 3000.09: Autonomy in Weapon Systems*, November 21, 2012, https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf.

98 Ankit Panda, "US Department of Defense adopts artificial intelligence ethical principles," *The Diplomat*, February 25, 2020, https://thediplomat.com/2020/02/us-department-of-defense-adopts-artificial-intelligence-ethical-principles/.

99 US Department of Defense, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy,* February 12, 2019, https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF, 7-11.

100 Sayler, *Defense Primer*, 2.

101 Khari Johnson, "The US military wants your opinion on AI ethics," *Venture Beat,* April 26, 2019, https://venturebeat.com/2019/04/26/the-u-s-military-wants-your-opinion-on-ai-ethics/.

CONCLUSION AND RECOMMENDATIONS

We are now in a transitional period wherein AI has moved from potential to reality. As surveyed in this document, AI's enormous power already is upending human existence, for better and for worse. If governments wish to take advantage of AI's potential for positive change, and to avoid AI's equally significant negative impact, they will need to define societal priorities, develop actionable strategies, and then follow through with smart policies backed by real funding.

What follows is a non-exhaustive list of recommendations:

1: ETHICS AND NORMS

The creation of ethics and norms statements is now well-trodden territory. A cynic would argue that these statements are common precisely because they are nonbinding, cost little (in terms of financial costs and political capital), have no adverse impact on affected economic sectors, and signal virtuous but often vague institutional action.

But there is a compelling counterargument here, too. When defined through institutionalized, public settings, ethics and norms statements can force governments to clarify their most cherished values and thereby define their highest priorities. That process forces stakeholders to sharpen their arguments and negotiate their positions alongside others. Moreover, nonbinding processes help generate consensus around AI goals ahead of the more binding, hence more difficult, policymaking processes that follow.

Governments, therefore, should follow the EU's lead in establishing high-level commissions, led by well-recognized and -respected chairpersons and with membership broadly representative of society and tasked with going through a rigorous process leading to a well-publicized and -distributed ethics and norms statement.

2: STRATEGY

When it comes to AI, anticipatory governance is critical. Although the exact contours of the AI-driven future are not known, governments should expect that AI will drive significant and disruptive change. Governments, above all, will need to develop actionable and practicable strategies to shape the future in a positive direction while avoiding or minimizing the negative consequences that inevitably will follow in AI's train.

Comprehensive strategies should be based upon ethics and norms statements, in order to help define strategic ends. Ideally, an AI strategy not only would define what the government hopes to accomplish but also, and as importantly, would define the outcomes it is not willing to tolerate. Further, it is important to acknowledge at the outset that some strategic outcomes will contradict others. Governments will need to strike a balance between aiming on the one hand for the cutting edge of innovation—to be a global leader in the creation of AI-based startups and AI-driven industries—and on the other dealing with AI's societal consequences.

As an example, the chipmaker Intel recently published a prototype national strategy for the United States. It argued that the US government should adopt a four-pronged strategy focusing on innovation, employment and human welfare, data "liberation," and removal of barriers for AI development.[102] Inclusion of the employment and human welfare plank, wherein Intel implores the US government to invest more heavily and creatively in human capital, is an acknowledgment of AI's dual implications for working people. Intel is saying, in effect, that unless the United States revamps its labor policies, AI will have profoundly negative impacts on many workers even as other workers benefit.

Strategies must be backed by sufficient implementation resources, otherwise they are just paper statements. Besides providing enough financial resources, governments also should create or authorize implementing institutions. The United States' proposed Artificial Intelligence Initiative Act would create a National AI Coordination Office plus an interagency mechanism and a nongovernmental expert group. Whether these would be sufficiently powerful institutions that could have a real say in directing a national AI strategy is an open question. Regardless, implementing institutions ought to have enough power and resources to shift implementation tactics given AI's uncertain societal impacts and its high potential for contradictory outcomes.[103]

102 "White Paper: Intel's recommendations for the US National Strategy on Artificial Intelligence," *Intel*, March 2019, https://newsroom.intel.com/wp-content/uploads/sites/11/2019/03/intel-ai-white-paper.pdf.

103 An analogous recommendation is made in Joël Blit, Samantha St. Amand, and Joanna Wajda, *Automation and the Future of Work: Scenarios and Policy Options*, CIGI Papers No. 174, May 2018, 7, https://www.cigionline.org/publications/automation-and-future-work-scenarios-and-policy-options.

A protestor demonstrates against facial recognition and video surveillance technologies. Berlin, Germany, November 2017
Source Stefanie Loos via Flickr

3: PRIVACY AND DATA

As individuals' data often provide the raw material upon which AI systems work, privacy issues are at the core of most policy debates. For every privacy concern, often there is a powerful counterpoint regarding the public benefits that AI could generate in areas ranging from health care to transport to environmental protection. Striking this balance, between protection of individuals' rights on the one hand and maximizing the potential benefits that AI could generate for the broader public on the other, is at the core of AI policymaking.

This trade-off is never far from the surface of AI policy debates. The more personal the data, the greater the concerns and, often, the pushback. Perceptions of who gathers the data and under what circumstances, who owns and controls the data, and what rules exist regarding what can be done with the data, appear to have as much importance as any other factor. For example, in November 2018, Google announced that it would absorb DeepMind Health, a British AI health analytics company, bringing it closer to Google's core operations. Although Google had purchased DeepMind Health in 2014, it had maintained the company's independence until the 2018 announcement.[104] Within the UK, the news that the United States' tech giant, Google, would have access to a vast amount of National Health Service data without patient knowledge or control set off alarm bells among experts and the general public.[105]

AI-based facial recognition technology is one area where privacy concerns might emerge triumphant over other considerations. The European Commission reportedly is considering regulatory interventions to severely limit the uses of facial recognition technologies across Europe. Such a move logically would build

104 Christina Farr, "The new Google Health unit is absorbing health business from DeepMind, Alphabet's AI research group," *CNBC*, November 3, 2018, https://www.cnbc.com/2018/11/13/google-health-unit-absorbs-deepmind-health.html.

105 Chris Stokel-Walker, "Why Google consuming DeepMind Health is scaring privacy experts," *Wired*, November 14, 2018, https://www.wired.co.uk/article/google-deepmind-nhs-health-data.

upon the EU's GDPR and AI ethics guidelines, in the sense that it would retain the EU's interest in privileging the individual's right to privacy above the interests of firms, institutions, and governments.[106] This kind of outcome might occur more frequently as publics become more aware of the technology's power, intrusiveness, and increasing ubiquity.[107]

Governments will need to establish data ownership and usage rights and be prepared to accept the consequences. If large tech companies own data outright and can use it at will, then the positive outcomes might include building the data economies of scale considered necessary for constructing robust commercially viable AI systems. But negatives include the loss of individuals' privacy.

To those who brush privacy concerns aside, it should be clear by now that populations in the United States, Europe, and elsewhere have become, over a very short period of time, far less tolerant of the tech sector and its claims to benevolence. (In August 2019, Apple acknowledged that it had hired humans to listen to audio gathered by its voice assistant Siri, without the knowledge and approval of its customers; Apple did so in order to feed better training data into its voice recognition machine learning system.)[108]

If, on the other hand, individuals have the greatest control over data, then privacy is privileged over commerce. The EU has taken this road. In calling for GDPR reform, the Center for Data Innovation (CDI), a Washington think tank, argues that the regulation is skewed too heavily toward privacy, in effect stifling investment in and innovation around the "algorithmic economy."[109] The EU, for its part, insists privacy protection will benefit commerce. Results remain to be seen.

Proactive government policy can seek to find a balance. One strategy is to clarify and then regulate data usage protections, regardless of data ownership. The policy question is to ask who is allowed to do what, with what data, and under what conditions? In 2016, France passed the Digital Republic Act, which established a comprehensive national open data policy, focused on government provision of standardized, publicly available data sets to (among other things) facilitate formation of startups and enable AI-driven analytics.[110] In 2018, the French government articulated a national AI strategy embraced by President Emmanuel Macron.[111] The strategy asserts that AI systems should serve the public interest, as, for example, in generating health care solutions to combat diseases, while at the same time ensuring that the data upon which such analyses are conducted do not compromise individual privacy.

4: INNOVATION ECOSYSTEMS AND HUMAN CAPITAL

Governments the world over see AI development as an opportunity to build and strengthen their innovation ecosystems. As noted in the above section, the goal is the same as for other emerging technologies: to facilitate startup formation and new economic sectors within a country's borders so as to reap the economic benefits that follow.

As the author discusses at length in two previous Atlantic Council reports, there are multiple policy interconnections involved in creating strong tech innovation ecosystems.[112] These span research and development spending, IP protection, infrastructural investment, tax policy, housing policy, and much more. Such policy interconnections are as relevant to spurring AI-related innovation as they are to spurring innovation related to other emerging technologies. The reader is advised to consume these previous reports.

With respect to AI in particular, policymakers need to maximize human capital to build innovation ecosystems while reducing downside impacts on labor.

---

106 Mehreen Khan, "EU plans sweeping regulation of facial recognition," *Financial Times*, August 22, 2019, https://www.ft.com/content/90ce2dce-c413-11e9-a8e9-296ca66511c9.

107 The media is focusing on the degree to which facial recognition technology is being used without the knowledge or consent of individuals. See, e.g., Madhumita Murgia, "Who's using your face? The ugly truth about facial recognition," *FT Magazine*, April 19, 2019, https://www.ft.com/content/cf19b956-60a2-11e9-b285-3acd5d43599e.

108 Patrick McGee, "Apple apologises for listening to Siri conversations," *Financial Times*, August 28, 2019, https://www.ft.com/content/2563911e-c9a9-11e9-a1f4-3669401ba76f.

109 Chivot and Castro, *Reform*.

110 Anoush Darabi, "A digital republic? The unrivalled open data experiment transforming France," *apolitical*, November 26, 2018, https://apolitical.co/solution_article/a-digital-republic-the-unrivalled-open-data-experiment-transforming-france/.

111 *AI for Humanity: French Strategy for Artificial Intelligence* (Paris: French Digital Council, 2018), https://www.aiforhumanity.fr/en/.

112 Engelke and Manning, *Innovative Edge*; Manning and Engelke, *Sweepstakes*.

Regarding foreign labor, there are vanishingly few successful tech innovation ecosystems that depend primarily on domestic workforces. The rest compete—fiercely—for talented workers from all over the world. Given their massive populations, perhaps China or India can get away with relying primarily on domestic talent. No one else, including the United States, has this luxury.

The policy implication, therefore, is clear. Policymakers should encourage the immigration of highly skilled foreign talent through as many pathways as possible, including offering citizenship to those migrants who prove their interest in staying and contributing to the host country over the long run.

Regarding the domestic workforce, the policy equation becomes more complicated. The objective no longer involves attracting existing talent, but rather creating that talent domestically and otherwise ensuring that a country's citizens do not fall victim to AI-generated obsolescence.

There is no magic-bullet solution to ensuring that the domestic workforce possess the right skills and competencies necessary to survive in an AI-driven world. Policymakers will have to approach the problem through multiple pathways.

For several decades, lifelong education via degree-granting formal institutions has been sold as the universal solution. But the speed with which AI already is altering the workplace has forced an intense conversation about how education, including higher education, must change in order to prepare students for a very different future labor market. Getting more students into STEM fields ("science-technology-engineering-math"), training more people how to code, and otherwise investing in worker retraining programs have been key components of this message.[113] Some governments are revisiting their educational and training models.

But the formal education model has big limitations. Formal education imposes high costs—money and time—upon people who might be out of work, poorly compensated if they are working, or overwhelmed with other obligations such as family. The model assumes that skills are developed primarily in the classroom versus on the job, which economic history shows is only partially true. The model also does not address labor market shortcomings. Examples include occupational licensing, which requires people to certify themselves in order to enter a field (often through time- and money-consuming certification programs) and non-compete agreements, which prevent workers from taking their knowledge and skills to other firms.[114]

As AI is taking labor into uncharted territory, policymakers will have to be flexible and innovative in crafting solutions for a turbulent era. There is no single template from which policymakers can draw. If there is a consensus, it is that:

- high-quality and well-rounded education matters as much early in life as it ever has, albeit adapted to today's realities (e.g., ensuring youth become comfortable working alongside machines and AI-driven systems, plus ensuring they acquire an entrepreneurial mindset);

- workers will need to acquire new and transferable skills throughout their working lives and at a faster pace with fewer obstacles in the way (time and money);

- the social welfare system, which was built for an industrial economy with more stable employment, will need to be refashioned for a future when the workplace is less secure.

113 Joshua Kim, "If you read one higher ed book this year, make it 'Robot-Proof,'" *Inside Higher Ed*, April 18, 2018, https://www.insidehighered.com/digital-learning/blogs/technology-and-learning/read-robot-proof-if-you-only-read-one-higher-ed-book.

114 Thinkers in this vein include the economic historian James Bessen and the entrepreneur Nicolas Colin. See, e.g., James Bessen, *Learning by Doing: The Real Connection between Innovation, Wages, and Wealth* (New Haven: Yale University Press, 2015); Nicolas Colin, Hedge: *A Greater Safety Net for the Entrepreneurial Age* (London: Family Stories, 2018).

The UN gathers to debate autonomous weapons under the auspices of the Convention on Certain Conventional Weapons (CCW). Geneva, Switzerland, May 2014 Source UN Photo / Jean-Marc Ferré via Flickr.

5: AUTONOMOUS WEAPONS AND LETHAL FORCE

Gill Pratt, a roboticist formerly at DARPA (US Defense Advanced Research Projects Agency), argued in a widely cited 2015 paper that advances in machine learning, computing power, data and energy storage systems, and sensors stand to deliver a "Cambrian Explosion" in AI-enabled robotics, by which he meant an "explosion in the diversification and applicability of robotics."[115]

For the world's militaries, such advances in AI-enabled systems are irresistible. They promise advances in transport, supply, logistics, surveillance, and of course warfighting capabilities far beyond those possessed by any military today.

Therein lies the crux of the dilemma. In an anarchic system without any binding and enforceable treaties or other international constraints on autonomous weapons systems, states believe they have no choice but to develop such systems. As the authors of a 2017 Harvard Kennedy School study on the military uses of AI argued, states "will face increasing temptation to delegate greater levels of authority to a machine, or else face defeat" owing to machines' superior warfighting capabilities.[116]

States invest in technology to ensure their own security. As was true of the computer, global positioning system (GPS) technologies, satellites, and many other technologies, the world's militaries are funneling huge funds into AI development because they fear falling behind their rivals in this key technology. Indeed, the US military is one of the world's greatest funders of AI research: the DoD argues that the United States "must adopt AI to maintain its strategic position" vis-à-vis its geopolitical rivals.[117]

And although smaller powers often push for limits on arms, some conversely may see advantages in developing LAWS. The Harvard study's authors speculate, as an example, that in

115 Gill A. Pratt, "Is a Cambrian Explosion Coming for Robotics?," *Journal of Economic Perspectives* 29, 3 (2015), 51-60 (quotation p. 51), https://www.aeaweb.org/articles?id=10.1257/jep.29.3.51.

116 Greg Allen and Taniel Chan, *Artificial Intelligence and National Security* (Boston: Harvard Kennedy School, Belfer Center, July 2017), 21.

117 US Department of Defense, *Summary*, 5 (emphasis added).

the future rich countries having "small, elderly, and declining populations may be able to use robotics and autonomy to possess robotic 'manpower' far beyond their human population size…[so as to] field greater numbers of more capable robotic 'warfighters' than some more populous adversaries."[118]

In common parlance, then, we have an arms race on our hands. And there are few, if any, brakes to check it.

Most, if not all, observers expect that the pace of AI-driven weapons systems will result in fully autonomous capabilities within the next decade, if not sooner. And there is sound logic for deploying such systems, given that the "pace of war"—the speed at which battle occurs—is becoming so fast that human processing soon will be too slow to compete with machines.[119]

Within the UN's GGE process, the lineup of states opposing regulation or a ban includes countries having the most advanced militaries: besides the United States, the list includes Russia, the UK, Germany, France, South Korea, Australia, Israel, Sweden, Belgium, Spain, Turkey, and China. Those in support are generally poorer countries from the global South.[120] This split neatly divides the world into the military haves and have nots, in turn suggesting that fears of being outflanked by one's geopolitical rivals have the upper hand compared with fears of experiencing real, AI-enabled shooting wars.

As with the future of labor in an AI-enabled landscape, there is no simple and easy policy fix here. It is unrealistic to expect that all of the world's leading militaries will agree to a preemptive ban on LAWS and—critically—stick to an agreement. However, absent regulation if not the outright ban of LAWS, it is certain that at some point in the future these weapons systems will exist and be deployed, perhaps with devastating and unforeseen consequences.

It is with this scenario in mind that the United States and its key allies and partners, within the UN's GGE process, should drop their collective refusal to entertain such restrictions and at least begin exploring constructive ways in which the international community might put real and enforceable boundaries around this novel and frightening category of weapons.

Dr. Peter Engelke is a deputy director and senior fellow within the Atlantic Council's Scowcroft Center on Strategy and Security. His diverse work portfolio at the Center's Foresight, Strategy, and Risks Initiative spans global and regional futures, geopolitics and international affairs, innovation and technological disruption, climate change and natural resources, and global urbanization among other topics.

118 Allen and Chan, *National Security*, 23.

119 Zachary Fryer-Biggs, "Coming Soon to a Battlefield: Robots That Can Kill," *Atlantic*, September 3, 2019, https://www.theatlantic.com/technology/archive/2019/09/killer-robots-and-new-era-machine-driven-warfare/597130/.

120 This list is derived from Michael Moodie, *International Discussions Concerning Lethal Autonomous Weapon Systems* (Washington: Congressional Research Service, August 16, 2019), Table 1, https://assets.documentcloud.org/documents/6305453/International-Discussions-Concerning-Lethal.pdf.

Atlantic Council Board of Directors

The Atlantic Council is a nonpartisan organization that promotes constructive US leadership and engagement in international affairs based on the central role of the Atlantic community in meeting today's global challenges.

Atlantic Council

1030 15th Street, NW, 12th Floor,
Washington, DC 20005

(202) 463-7226, www.AtlanticCouncil.org