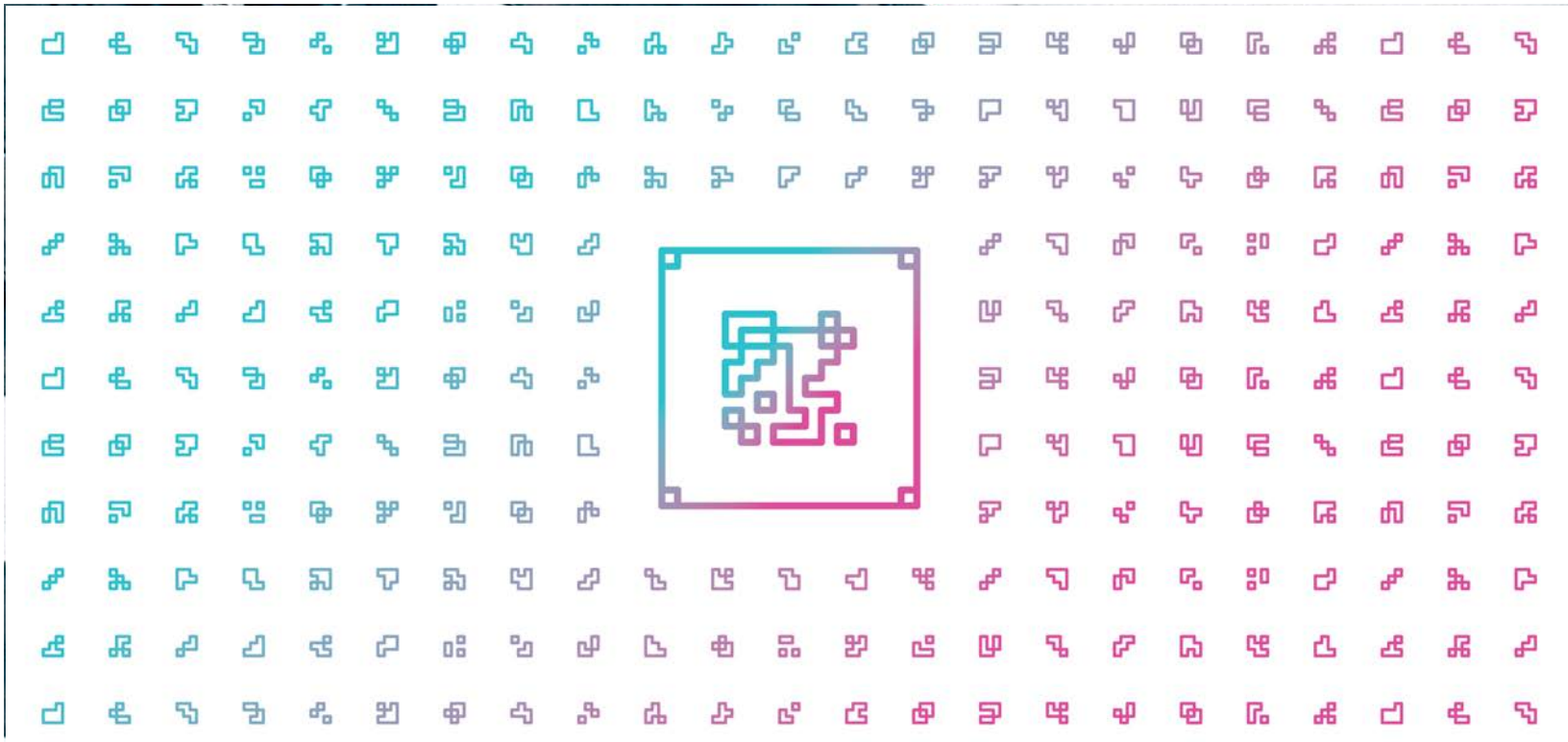


Artificial Intelligence and International Security

July 2018

By Michael C. Horowitz, Gregory C. Allen,
Edoardo Saravalle, Anthony Cho,
Kara Frederick, and Paul Scharre

Bold. Innovative. Bipartisan.

1152 15th Street NW
Suite 950
Washington, DC 20005

CNAS.ORG / @CNASDC
T: 202.457.9400
F: 202.457.9401

ABOUT THE AUTHORS

Gregory C. Allen is an Adjunct Fellow in the Technology and National Security Program at the Center for a New American Security (CNAS).

Anthony Cho is a former Researcher in the Technology and National Security Program at CNAS.

Kara Frederick is a Research Associate in the Technology and National Security Program at CNAS.

Michael C. Horowitz is an Adjunct Senior Fellow at CNAS and a professor of political science at the University of Pennsylvania.

Elsa Kania is an Adjunct Fellow in the Technology and National Security Program at CNAS.

Edoardo Saravalle is a Researcher for the Energy, Economics, and Security Program at CNAS.

Paul Scharre is a Senior Fellow and Director of the Technology and National Security Program at CNAS.

ABOUT THIS REPORT

This report is part of the Center for a New American Security's series on *Artificial Intelligence and International Security*. The series examines the potential consequences of advances in artificial intelligence for the national security community. Nearly every aspect of national security could be transformed by artificial intelligence. AI has applications for defense, intelligence, homeland security, diplomacy, surveillance, cybersecurity, information, and economic tools of statecraft. The United States must not only anticipate these developments, but act decisively to prepare for uses by competitors and take advantage of the opportunities AI presents.

ALSO IN THIS SERIES

The *Artificial Intelligence and International Security* series includes:

- **Artificial Intelligence: What Every Policymaker Needs to Know** by Paul Scharre and Michael C. Horowitz with Preface by Robert O. Work
- **Strategic Competition in an Era of Artificial Intelligence** by Michael C. Horowitz, Gregory C. Allen, Elsa B. Kania, and Paul Scharre (forthcoming)

This series is part of the Center for a New American Security's multi-year Artificial Intelligence and Global Security Initiative. Learn more at cnas.org/AI.

ACKNOWLEDGEMENTS

We would like to thank Loren Schulman for her helpful comments on an early draft of this report. We would also like to thank Maura McCarthy and Allene Bryant for their role in the production and design of this report. Any errors or omissions are the sole responsibility of the authors. CNAS does not take institutional positions.

TABLE OF CONTENTS

NATIONAL SECURITY-RELATED APPLICATIONS OF ARTIFICIAL INTELLIGENCE

Introduction

There are a number of direct applications of AI relevant for national security purposes, both in the United States and elsewhere. Kevin Kelly notes that in the private sector "the business plans of the next 10,000 startups are easy to forecast: *Take X and add AI*."[1] There is similarly a broad range of applications for AI in national security. Included below are some examples in cybersecurity, information security, economic and financial tools of statecraft, defense, intelligence, homeland security, diplomacy, and development. This is not intended as a comprehensive list of all possible uses of AI in these fields. Rather, these are merely intended as illustrative examples to help those in the national security community begin to think through some uses of this evolving technology. (The next section covers how broader AI-driven economic and societal changes could affect international security.)

Cybersecurity

The cyber domain represents a prominent potential usage arena for AI, something senior leaders have expressed in recent years. In October 2016, National Security Agency (NSA) Director Michael Rogers stated that the agency sees AI as "foundational to the future of cybersecurity." Rogers' remarks occurred only two months after DARPA held its first Cyber Grand Challenge, a head-to-head fight between autonomous machines in cyberspace. Each system was capable of automatically discovering and exploiting cyber vulnerabilities in its opponents while patching its own vulnerabilities and defending itself from external cyberattacks.[2] Impressed with the tournament's results, DoD began a new program, Project Voltron, to develop and deploy autonomous cybersecurity systems to scan and patch vulnerabilities throughout the U.S. military.[3]

Even as DoD has begun to implement this technology, potential applications of AI for cybersecurity continue to evolve. The systems in the first Cyber Grand Challenge used rule-based programming and did not make significant use of machine learning. Were a similar competition to be held today, machine learning would likely play a much larger role. Below are several illustrative applications of machine learning in the cybersecurity domain that could be especially impactful for the international security environment.

Increased Automation and Reduced Labor Requirements

Cyber surveillance has tended to be less labor-intensive than the traditional human surveillance methods that it has augmented or replaced. The increased use of machine learning could accelerate this trend, potentially putting sophisticated cyber capabilities that would normally require large corporation or nation-state level resources within the reach of smaller organizations or even individuals.[4] Already there are countless examples of relatively unsophisticated programmers, so-called "script kiddies," who are not skilled enough to develop their own cyber-attack programs but can effectively mix, match, and execute code developed by others. Narrow AI will increase the capabilities available to such

actors, lowering the bar for attacks by individuals and non-state groups and increasing the scale of potential attacks for all actors.

Using AI to Discover New Cyber Vulnerabilities and Attack Vectors

Researchers at Microsoft[5] and Pacific Northwest National Laboratory[6] have already demonstrated a technique for using neural networks and generative adversarial networks to automatically produce malicious inputs and determine which inputs are most likely to lead to the discovery of security vulnerabilities. Traditionally, such inputs are tested simply by randomly modifying (aka "fuzzing") non-malicious inputs, which makes determining those that are most likely to result in new vulnerability discovery inefficient and labor-intensive. The machine learning approach allows the system to learn from prior experience in order to predict which locations in files are most likely to be susceptible to different types of fuzzing mutations, and hence malicious inputs. This approach will be useful in both cyber defense (detecting and protecting) and cyber offense (detecting and exploiting).

Automated Red-teaming and Software Verification and Validation

While there is understandable attention given to new vulnerability discovery, many cyber attacks exploit older, well-known vulnerabilities that system designers have simply failed to secure. SQL-injection, for example, is a decades-old attack technique to which many new software systems still fall prey. AI technology could be used to develop new verification and validation systems that can automatically test software for known cyber vulnerabilities before the new software is operationally deployed. DARPA has several promising research projects seeking to utilize AI for this function.

Automated Customized Social Engineering Attacks

Many major cybersecurity failures began with "social engineering," wherein the attacker manipulates a user into compromising their own security. Email phishing to trick users into revealing their passwords is a well-known example. The most effective phishing attacks are human-customized to target the specific victim (aka spear-phishing attacks) – for instance, by impersonating their coworkers, family members, or specific online services that they use. AI technology offers the potential to automate this target customization, matching targeting data to the phishing message and thereby increasing the effectiveness of social engineering attacks.[7] Moreover, AI systems with the ability to create realistic, low-cost audio and video forgeries (discussed more below) will expand the phishing attack space from email to other communication domains, such as phone calls and video conferencing.[8]

Information Security

The role of AI in the shifting threat landscape has serious implications for information security, reflecting the broader impact of AI, through bots and related systems in the information age. AI's use can both exacerbate and mitigate the effects of disinformation within an evolving information ecosystem. Similar to the role of AI in cyber attacks, AI provides mechanisms to narrowly tailor propaganda to a targeted audience, as well as

increase its dissemination at scale – heightening its efficacy and reach. Alternatively, natural language understanding and other forms of machine learning can train computer models to detect and filter propaganda content and its amplifiers. Yet too often the ability to create and spread disinformation outpaces AI-driven tools that detect it.

Targeted Propaganda and Deep Fakes

Computational propaganda inordinately affects the current information ecosystem and its distinct vulnerabilities. This ecosystem is characterized by social media's low barriers to entry, which allow anonymous actors – sometimes automated – to spread false, misleading or hyper-partisan content with little accountability. Bots that amplify this content at scale, tailored messaging or ads that enforce existing biases, and algorithms that promote incendiary content to encourage clicks point to implicit vulnerabilities in this landscape.[9] MIT researchers' 2018 finding that "falsehood [diffuses] significantly farther, faster, deeper and more broadly" than truth on Twitter, especially regarding political news, further illustrates the risks of a crowded information environment.[10] AI is playing an increasingly relevant role in the information ecosystem by enabling propaganda to be more efficient, scalable, and widespread.[11] A sample of AI-driven techniques and principles to target and distribute propaganda and disinformation includes:

- **Exploitation of behavioral data** – The application of AI to target specific audiences builds on behavioral data collection, with machine learning parsing through an increasing amount of data. Metadata generated by users of online platforms – often to paint a picture of consumer behavior for targeted advertising – can be exploited for propaganda purposes as well.[12] For instance, Cambridge Analytica's "psychographic" micro-targeting based off of Facebook data used online footprints and personality assessments to tailor messages and content to individual users.[13]

- **Pattern recognition and prediction** – AI systems' ability to recognize patterns and calculate the probability of future events, when applied to human behavior analysis, can reinforce echo chambers and confirmation bias.[14] Machine learning algorithms on social media platforms prioritize content that users are already expected to favor and produce messages targeted at those already susceptible to them.[15]

- **Amplification and agenda setting** – Studies indicate that bots made up over 50 percent of all online traffic in 2016.[16] Entities that artificially promote content can manipulate the "agenda setting" principle, which dictates that the more often people see certain content, the more they think it is important.[17] Amplification can increase the perception of significance in the public mind. Further, if political bots are "written to learn from and mimic real people," according to computational propaganda researchers Samuel Woolley and Philip Howard, then they stand to influence the debate. For example, Woolley and Howard point toward the deployment of political bots that interact with users and attack political candidates, weigh in on activists' behavior, inflate candidates' follower numbers, or retweet specific candidates' messaging, as if they were humans.[18] Amplifying damaging or

distracting stories about a political candidate via "troll farms" can also change what information reaches the public. This can affect political discussions, especially when coupled with anonymity that reduces attribution (and therefore accountability) to imitate legitimate human discourse.[19]

- **Natural language processing to target sentiment** – Advances in natural language processing can leverage sentiment analysis to target specific ideological audiences."[20] Google's offer of political interest ad targeting for both "left-leaning" and "right-leaning" users for the first time in 2016 is a step in this direction.[21] By using a systemic method to identify, examine, and interpret emotional content within text, natural language processing can be wielded as a propaganda tool. Clarifying semantic interpretations of language for machines to act upon can aid in the construct of more emotionally relevant propaganda.[22] Further, quantifying user reactions by gathering impressions can refine this propaganda by assessing and recalibrating methodologies for maximum impact. Private sector companies are already attempting to quantify this behavior tracking data in order to vector future microtargeting efforts for advertisers on their platforms. These efforts are inherently dual-use – instead of utilizing metadata to supply users with targeted ads, malicious actors can supply them with tailored propaganda instead.

- **Deep fakes** – AI systems are capable of generating realistic-sounding synthetic voice recordings of any individual for whom there is a sufficiently large voice training dataset.[23] The same is increasingly true for video.[24] As of this writing, "deep fake" forged audio and video looks and sounds noticeably wrong even to untrained individuals. However, at the pace these technologies are making progress, they are likely less than five years away from being able to fool the untrained ear and eye.

Countering Disinformation

While no technical solution will fully counter the impact of disinformation on international security, AI can help mitigate its efficiency. AI tools to detect, analyze, and disrupt disinformation weed out nefarious content and block bots. Some AI-focused mitigation tools and examples include:

- **Automated Vetting and Fake News Detection** – Companies are partnering with and creating discrete organizations with the specific goal of increasing the ability to filter out fake news and reinforce known facts using AI. In 2017, Google announced a new partnership with the International Fact-Checking Network at The Poynter Institute, and MIT's the Fake News Challenge resulted in an algorithm with an 80 percent success rate.[25] Entities like AdVerif.ai scan and detect "problematic" content by augmenting manual review with natural language processing and deep learning.[26] Natural language understanding to train machines to find nefarious content using semantic text analysis could also improve these initiatives, especially in the private sector.

- **Trollbot Detection and Blocking** – Estimates indicate the bot population ranges between 9 percent and 15 percent on Twitter and is increasing in sophistication. Machine learning models like the Botometer API, a feature-based classification system for Twitter, offer an AI-driven approach to identify them for potential removal.[27] Reducing the amount of bots would de-clutter the information ecosystem, as some political bots are created solely to amplify disinformation, propaganda, and "fake news."[28] Additionally, eliminating specific bots would reduce their malign uses, such as for distributed denial-of-service attacks, like those propagated by impersonator bots throughout 2016.[29]

- **Verification of Authenticity** – Digital distributed ledgers and machine speed sensor fusion to certify real-time information and authenticity of images and videos can also help weed out doctored data. Additionally, blockchain technologies are being utilized at non-profits like PUBLIQ, which encrypts each story and distributes it over a peer-to-peer network to attempt to increase information reliability.[30]

Content filtering often requires judgement calls due to varying perceptions of truth and the reliability of information. Thus, it is difficult to create a universal filter based on purely technical means, and it is essential to keep a human in the loop during AI-driven content identification. Technical tools can limit and slow disinformation, not eradicate it.

Economic and Financial Tools of Statecraft

Illicit funds course through the global financial system and support terrorism, money laundering, and WMD proliferation. To counter these flows, U.S. officials have expanded the global network of anti–money laundering and counterterrorist financing tools since 9/11. Yet the United Nations estimates that law enforcement only seizes 1 percent of criminal funds.[31] One potential national security application of AI tools is their use to strengthen counter–illicit–financing operations.

By analyzing and learning from large sets of data, AI could accomplish tasks not possible in a human-centered counter–illicit–financing system. AI's anomaly detection and pattern recognition capabilities could help a system learn from the unstructured data collected by financial institutions. In one case, a regulatory technology company integrating AI tools found a correlation between users who had changed their browser language and a type of fraud.[32] This analysis uncovered a metric not traditionally used by financial investigators and expanded the definition of usable data. Better pattern recognition will also sort information more usefully. Better sorting can reduce false positives that would otherwise result in alerts. For example, AI could reduce false positives in "high-risk" jurisdictions by replacing an imprecise geographic input with a more effective red flag. Fewer alerts will save time and manpower.

Even short of large-scale pattern analysis, AI can improve the counter–illicit–financing framework. Automation could ensure sustained attention to illicit financing threats, even when not prioritized by financial institutions. This feature would allow constant pressure

on potential dangers. It would also reduce stress on financial institutions. Banks would no longer have to shift their attention to respond to changing government priorities – for example, from Iran to North Korea. Automation could also integrate available non-financial information about entities and individuals. Today, a significant amount of publicly accessible information is not automatically part of investigations. Through image recognition, AI programs could use open-source social media information that currently does not inform counter–illicit–financing processes.[33] Changes made to a customer's social media presence or networks mapped out through publicly available images could clarify a customer's risk profile.[34]

AI capabilities could address the counter–illicit–financing framework's major challenges. First, AI could improve efficiency. Human-centered counter–illicit–financing processes generate false positives that detract from investigations and allow threats to go undiscovered or uninvestigated. One study found that 80 to 90 percent of suspicious activity reporting yielded no value.[35] Fewer false positives, through better pattern recognition and data segmentation, will save time and money.

The savings of time and money that AI systems could enable would be particularly important in combating illicit financial flows. Since 9/11, governments have enlisted financial institutions as partners in the fight against illicit finance. Banks have shouldered increasing compliance costs to keep up with growing regulatory requirements, including to counter illicit finance. Between 2011 and 2017, the cost of compliance has increased by over 20 percent for most banks.[36] Lower costs will ensure banks' continued cooperation. Lower costs will also allow smaller, regional banks in high-risk jurisdictions to conduct compliance work that currently only large multinational institutions can afford. Greater participation from smaller banks will reduce vulnerable entry points into the global financial system.

AI could also help governments and financial institutions address data privacy and protection problems. Currently, privacy laws hamper efforts to make the most of collected information. In some cases, financial institutions can struggle even to share information among their own branches in different jurisdictions.[37] These limitations create barriers to integrating information and, more importantly, to learning from past typologies of illicit financing. Dr. Gary Shiffman, CEO of Giant Oak, a data science company that uses algorithms to understand large quantities of data, argues that AI could circumvent this problem. An AI system could learn from analyzing a dataset in one jurisdiction. The system could then move its algorithms to other jurisdictions and learn from a new dataset without moving the underlying data itself.[38] Privacy limitations would no longer hamper the learning.

Though AI could incrementally improve the counter–illicit–financing framework, it could also fundamentally disrupt it. Financial institutions often use static rules to counter illicit funding. For example, a transaction over $10,000 will trigger a currency transaction report. Rogue players, however, can adapt faster than the rules can evolve. For this reason, international standard-setters like the Financial Action Task Force (FATF) urge financial institutions to use risk-based systems that proactively adapt to and mitigate risk. Because

this approach is costly and time-intensive, FATF requires risk-based measures to tackle money laundering and terrorist financing, but not for the financing of proliferation. An AI-based system, constantly learning and incorporating new information, will allow the expansion of risk-based programs to create a more dynamic counter–illicit–finance program across threat categories.[39] An AI system, for example, could spot the patterns used by individuals to evade the $10,000 limit, connect these illicit networks, and potentially block the wires from leaving banks before transferred amounts become too big.[40]

This section has focused on applying AI to flows of finance rather than the infrastructure and markets supporting these flows. The "flash crash" has shown the susceptibility of programmed trading mechanisms to negative interactions and the currently insufficient preparation for this threat.[41] Opponents could use AI to manipulate markets or destabilize currencies. This category of threats, however, falls outside the traditional realm of economic statecraft. Instead, it would be analogous to a malicious cyber attack.

Making the most of financial data will be particularly important going forward. As more and more communication becomes encrypted, financial records will become more important sources of data for investigations and intelligence work. However, the tools to use the data have not yet evolved accordingly. AI offers a way forward.

Defense

Militaries around the globe are already incorporating more robotics and autonomous systems into their forces, a trend that has been the focus of prior CNAS work. Artificial intelligence and machine learning will allow these systems to tackle more challenging tasks in a wider range of environments. Because of the ubiquitous nature of AI technology, non-state groups and individuals will also be able to harness and use this technology.

In combat operations, robots, swarms, and autonomous systems have the potential to increase the pace of combat. This is particularly the case for domains of machine-to-machine interaction, such as in cyberspace or the electromagnetic spectrum. AI could be used not only to create more intelligent robotics, but also to power more advanced sensors, communications, and other key enablers.

- **Situational awareness**: Small robotic sensors could be used to collect information, and AI-enabled sensors and processing could help make better sense of that information. Deep neural networks already are being used for image classification for drone video feeds as part of the Defense Department's Project Maven, in order to help humans process the large volumes of data being collected. While current AI methods lack the ability to translate this into an understanding of the broader context, AI systems could be used to fuse data from multiple intelligence sources and cue humans to items of interest. AI systems also could be used to generate tailored spoofing attacks to counter such sensors and processors.

- **Electromagnetic spectrum dominance**: AI systems could be used to generate novel methods of jamming and communications through self-play, akin to AlphaGo Zero improving its game by playing itself. For example, one AI system could try to send signals through a contested electromagnetic environment while another system attempts to jam the signal. Through these adversarial approaches, both systems could learn and improve. DARPA held a Spectrum Challenge in 2014 with human players competing to send radio signals in a contested environment.[42] DARPA is now using machine learning to aid in radio spectrum allocation,[43] but this concept also could be applied to jamming and creating jam-resistant signals.

- **Decoys and camouflage**: Generative adversarial networks could be used to create militarily relevant deep fakes for camouflage and decoys, and small robotic systems could be used as expendable decoys. As militaries incorporate more AI-enabled sensors for data classification, spoofing attacks against such systems will be increasingly relevant as well.

- **Tactics**: Evolutionary and reinforcement learning methods could be used to generate new tactics in simulated environments, coming up with surprising solutions as they have in other settings.

- **Command and control**: As the pace of battle accelerates and the volume and speed of information eclipses the ability of human warfighters, AI will become increasingly important for command and control. Autonomous systems that have been delegated authority for certain actions can react at machine speed at the battlefield's edge without waiting for human approval. AI can also help commanders process information faster, allowing them to better understand a rapidly changing battlespace. Through automation, commanders can then relay their orders to their forces – human or machine – faster and more precisely.

AI systems can also aid militaries in a range of non-combat support functions. One use of AI will be to help defense leaders better understand their own forces. By analyzing large amounts of data, AI systems may be able to predict stress on the force in various components: when equipment requires maintenance; when programs are likely to face cost overruns or schedule delays; and when servicemembers are likely to suffer degraded performance or physical or psychological injuries. Overall, AI has tremendous potential to help defense leaders improve the readiness of their own forces by assembling and fusing data and doing predictive analysis so that problems can be addressed before they become critical.

AI also is ripe for transforming traditional business processes within military and other government organizations. The U.S. Defense Department, for example, conducts a range of non-military specific business functions, including accounting, travel, medicine, logistics, and other administrative functions. Many of these functions are ripe for automation because they involve routine cognitive or physical labor. In many cases, military organizations may be able to directly import mature and proven technologies from the

commercial sector that can improve efficiencies and reduce personnel costs, such as more automated accounting systems or AI tools in health care. Defense organizations could save substantial sums of money by drawing on these commercial technologies and streamlining their organizations.

Overall, artificial intelligence can help militaries improve understanding, predict behavior, develop novel solutions to problems, and execute tasks. Some applications, such as the use of AI to enable autonomous weapons, raise difficult legal, ethical, operational, and strategic questions. The potential for automation to increase the pace of combat operations to the point where humans have less control over the conduct of war raises profound questions about humanity's relationship with war, and even the nature of war itself.

Intelligence

AI has many uses in intelligence collection and analysis. For collection, the explosion of data that is occurring because of smart devices, the Internet of Things, and human internet activity is a tremendous source of potential information. This information would be impossible for humans to manually process and understand, but AI tools can help analyze connections between data, flag suspicious activity, spot trends, fuse disparate elements of data, map networks, and predict future behavior. This could make clandestine activity more challenging in a number of ways, as the combination of big data, data breaches, and increased open source information could make it more difficult to keep intelligence professionals undercover. For example, facial recognition and biometrics, combined with large surveillance systems, could make operating under aliases increasingly difficult.

At the same time, AI systems may be vulnerable to counter-AI spoofing techniques, such as fooling images, which will have implications for the intelligence community. Deep fakes and the automation of data creation at scale may make it possible to create deep backstories for individuals undercover. AI may even transform verification of human reporting through improvements in systems that can correlate brain imaging to thoughts, with major implications for counter-intelligence and interrogation.[44]

AI also has tremendous potential value in intelligence analysis. AI systems can be used to track and analyze large amounts of data – including open-source data – at scale, looking for indications and warning of suspicious activity. Anomaly detection can help find terrorists, clandestine agents, or indications and warning of potential enemy military activity. AI-based speech-to-text and translation services could greatly increase the scale of processing audio, video, and text-based foreign language information. AI systems could be used to generate simple automated reports, as they do already for some sports games.[45]

AI systems generally perform poorly at reading comprehension, but as they improve they could be used increasingly to write summaries of transcripts, making it easier for human analysts to quickly sift through the ever-growing volumes of information.[46] AI systems also could be increasingly valuable in doing semantic analyses of reports that help link disparate pieces of data that humans might miss. AI systems lack the common-sense reasoning that would allow them to make sense of information, but their ability to operate

with precision at scale will aid human analysts in sorting through massive volumes of information. AI systems will not replace human intelligence analysts, but can aid them by offloading routine tasks and processing data at scale, allowing human analysts to focus on understanding adversaries.

Homeland Security

AI can also aid a variety of border security and homeland security applications. AI-driven perception, processing, and analysis will be essential for collecting, sorting, and interpreting data to better inform human decision-making. The U.S. Department of Homeland Security (DHS) has already started to adopt and implement some of these technological advancements.

Examples of past and current AI-driven DHS initiatives include:

- **Voice recognition algorithms** – The U.S. Coast Guard has used artificial intelligence to analyze voices to build out their physical appearances. This has helped forensically address false distress signals.[47]

- **Open source data for machine learning** – In conjunction with Alphabet Inc.'s Kaggle platform, DHS made data from the Transportation Security Administration available to develop better algorithms to evaluate passenger luggage for illicit and dangerous items.[48]

- **Understanding data** – The Assistant for Understanding Data through Reasoning, Extraction, and Synthesis (AUDREY) AI platform developed by DHS and NASA's Jet Propulsion Laboratory integrates real-time data to make recommendations to firefighters on how to best function as a team.[49]

AI also has broad applicability in a variety of homeland security functions, such as border security.[50] Since the U.S. government cannot be stationed at every mile or inspect every container, AI systems, potentially in combination with UAVs and ground robotics, can aid in monitoring borders through advances in automated surveillance and anomaly detection. Systems that monitor human emotional expression and behavior could aid in recognizing humans that appear nervous or are acting oddly, serving as a "sixth sense" at border crossings. AI systems used for game theory/risk assessment also could be valuable in determining where best to apply scarce resources and how to counter adaptive adversaries, such as drug traffickers. Indeed, such systems already are being used to improve security against poachers in Africa.[51]

Diplomacy and Humanitarian Missions

Advances in artificial intelligence could also reshape the practice of diplomacy. AI technologies in image recognition and information sorting can make diplomatic compounds safer by monitoring personnel and identifying anomalies for potential vulnerabilities. In addition, language processing algorithms will lower language barriers

between countries, allowing them to communicate to foreign governments and publics more easily. More theoretical technologies like political forecasting also remain an option, mining an increasing array of available data to better understand and predict political, economic, and social trends.[52]

However, diplomacy will not be without disruptive challenges. Humans, for the foreseeable future, remain the decision-makers and must properly use the outputs provided by AI technologies. More alarmingly, as efforts to forge testimonies and propagate disinformation abroad are made easier, AI technologies will have to be applied defensively to react to, correct, or even remove malicious content.[53]

International humanitarian operations could also benefit greatly from AI technologies. AI technologies can help monitor elections, assist in peacekeeping operations, and ensure financial aid disbursements are not misused through anomaly detection. Of course, artificial intelligence can also help directly improve the quality of life in less developed nations by increasing productivity, health care, and myriad other economic benefits.[54] Artificial intelligence could also help in avoiding disasters that lead to international intervention. For example, AI technologies that extract significant actionable warning signs from climate and soil patterns will be a boon in agricultural efficiency and disaster preparedness.[55]

Implications

As an enabling technology, AI has many uses across a variety of national security settings. The United States should expand upon nascent efforts within different parts of the government and establish a whole-of-government initiative to harness and rapidly integrate AI tools within government operations. Because many current AI approaches have significant vulnerabilities, the United States should include safety and robustness against adversarial manipulation as key elements of its effort to incorporate AI technology, and employ "red teams" to test AI tools before they are deployed. The ubiquitous nature of AI technology means that the United States will have to move quickly to keep ahead of potential competitors.

THE INDIRECT EFFECTS OF THE ARTIFICIAL INTELLIGENCE REVOLUTION FOR GLOBAL SECURITY

How might AI generate political and societal change relevant for the international security environment beyond direct national security implications? Given the integral link between economic and military power, especially over the medium to long run, understanding how AI innovations will shape the global economy, the information environment, and societies around the world is crucial.

Economic Power and the Future of Work

The clearest connections between AI, the global economy, and economic power are through the effect of AI on the ability of countries and businesses to accumulate capital and the consequences for the future of work. The question is whether the consequences of AI will match, or even exceed, previous large-scale shifts in the economy. For example, in 1820 71 percent of Americans reportedly worked in farming occupations. However, the percentage of Americans working in farming declined significantly over the next century due to industrialization, falling to 30 percent in 1920 and 1 percent by 1988.[56]

There has been a large range of predictions on the way that AI will shape the labor market, and those predictions have a large degree of uncertainty. For example, a recent report by the McKinsey Global Institute suggests that nearly half of current job tasks across industries are automatable, while in six out of ten jobs more than 30 percent of the job tasks are automatable. The midpoint estimate of number of jobs displaced by 2030, according to McKinsey, is 400 million, while the high-end estimate is twice as high – 800 million.[57] These enormous totals, and the wide spread between them, reflect not just the notion that AI will have significant consequences on the labor market, but that those consequences are difficult to predict. Researchers' estimates on the effect of automation vary significantly. Research by Carl Benedict Frey and Michael A. Osborne at Oxford University suggests that 47 percent of U.S. workers might be at risk from automation by about 2030.[58] Another report examining 32 developed countries in the Organisation for Economic Cooperation and Development argues that 14 percent of jobs are at a high risk of automation and another 32 percent of jobs are at significant risk.[59] Meanwhile, a U.S. labor market model by Daron Acemoglu and Pascual Restrepo at the National Bureau of Economic Research, based on data on industrial robotics from 1990–2007, suggests adding "one more robot in a commuting zone reduces employment by 6.2 workers."[60] A Forrester research report, in contrast, argues that only 24.7 million jobs will be displaced by 2027, with 14 million created.[61] And even McKinsey says that only about 5 percent of jobs as they exist today could be fully automated.[62]

However, this is not just a question of how many jobs are displaced versus created, but whether those displaced will be able to find work in the new economy. The process of creative destruction can have significant political consequences even if the macro economic effects are relatively stable.[63] Former Secretary of the Treasury Larry Summers argued in 2017 that automation pressures, in combination with the difficulty of generating new skills

for labor force participation later in life, could result in "a third of men between the ages of 25 and 54 not working by the end of this half century."[64]

The challenge is that the number of jobs created by the cutting-edge companies of today, at the outset of the AI revolution, is already much smaller than the number of jobs created by the leading companies of previous generations. For example, in 2017 Facebook employed a little over 25,000 people, the largest it has ever been. Meanwhile, Ford Motor Company, with a fraction of the size of its peak labor force, still employed 202,000 workers in 2017.[65] The risk is that the optimal economic future for growth is more of a "labor-light economy," as Erik Brynjolfsson and Andrew McAfee argue, where capital generates continuing productivity gains, but workers don't benefit. [66] And workers, in this scenario, would not be just factory workers. They would be lawyers, doctors, investment bankers, and others that currently have middle class, upper middle class, or upper class incomes. All of those jobs have repetitive tasks, no matter how skilled, that narrow autonomous systems may be able to master. In this scenario, workers who perform repetitive physical and cognitive labor become less valued. Even if unemployment is low, reduced wages can be the effect. In fact, Brynjolfsson and McAfee argue that automation has been responsible for stagnant or falling real wages for the median American worker for the past several decades.[67]

Eras with this level of disruption can have significant indirect implications for the balance of power and the security environment. A change in the underlying basis of the economy can lead to industry shifts that benefit some countries at the expense of others. For example, the First Industrial Revolution helped fuel the rise of the United States – the geography of the United States enabled industrialization on a scale that was difficult to achieve in Europe. Government policy to capitalize on these changes can lead to long-lasting shifts in the relative balance of power. The ability of the British government to establish modern financing, in terms of government borrowing and bond markets, enabled Great Britain's creation of the most powerful navy in the world in the late 19th century.[68]

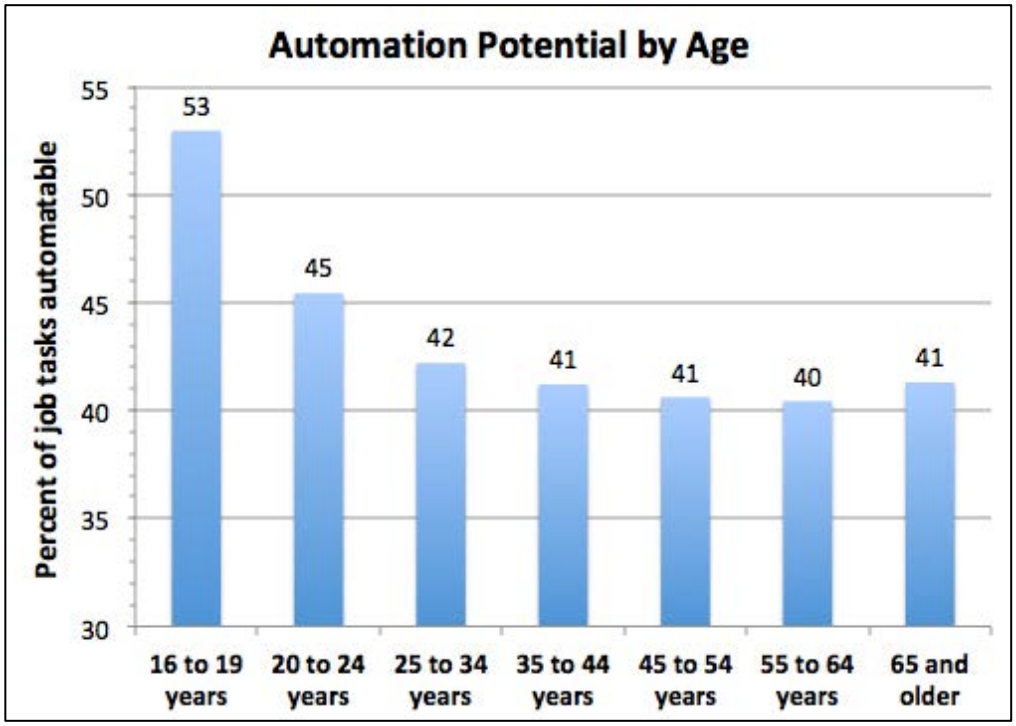
Political and Social Disruption

Economic disruption can also fuel social and political disruption. Large numbers of formerly employed workers, or even just groups that are newly disadvantaged due to economic circumstances, are a recipe for political protest and agitation. Maintaining stability requires a level of political dexterity and bureaucratic competence that can be difficult to achieve at the best of times – and periods of economic instability are hardly the best of times. This is one of the mechanisms through which economic transitions can lead to political conflict that, in the worst case, can make domestic unrest, insurgencies, civil wars, nationalism, xenophobia, and a turn to authoritarianism more likely.

The instability generated by automation is already a potential driving force in the rise of populist nationalist movements around the world. As powerful interest groups such as coal workers experience significant decline, they become ever more radical in their desire to see change to return to an old status quo that is impossible to achieve. This can drive political polarization.

A disparity in how automation affects different demographic groups could conceivably drive internal political conflict. To better understand how automation would affect American workers, the authors compared the analysis done by the McKinsey Global Institute on the effects of automation by sector[69] to the age of workers in each sector, as identified by the Bureau of Labor Statistics.[70] Figure 1 below shows the results.[71]

Figure 1



Automation potential represents the percentage of job tasks being done by workers in each category that could be automated. The youngest workers, who are more likely to be performing routine tasks, are hardest hit by automation. Numbers reflect the percentage of job tasks that are automatable; the percentage of entire jobs eliminated would be much lower. (Paul Scharre/ Data from Bureau of Labor Statistics and McKinsey Global Institute)

The results suggest that automation is likely to hit younger workers the hardest in the United States. This is not surprising, as younger workers are the most likely to be performing routine tasks that are easily automatable. This is particularly true for workers aged 16–19 and 20–24, who are less likely to be highly educated. Worker wages sharply decoupled by education level in the 1980s, with inflation-adjusted wages for those with a college or post-graduate degree rising and wages for high school graduates and dropouts falling.[72] This suggest that one effect of the automation economy will to be magnify the impact of education even more – even of specific majors or disciplines that help prepare people for the jobs of the future.

From the standpoint of managing the consequences of creative destruction, the silver lining is that the workers hardest hit by automation are those who are youngest and have the

most time to gain an education and adapt. One risk, however, is that younger workers who rely on entry-level jobs to pay their way through college and obtain an education could lose the economic opportunities they need to stay relevant in the automation economy. Without policy adjustments to make college and post-graduate education more affordable, the result therefore could be rising inequality.

National Economic Scenarios Regarding AI

Governments are not passive players at the mercy of a tidal wave of automation. Nations have a range of policy options available for responding to the economic pressure that AI will likely generate, from regulating industries to introducing a universal basic income. Countries will undoubtedly want to take advantage of AI as best they can while minimizing its harmful effects. What form that takes will depend on each country's political economy and the national attitudes toward economic growth, unemployment, political unrest, and social welfare.

The consequences of AI plus national policy responses could vary widely. Below are a few illustrative scenarios for how nations might end up after weathering and adapting to a wave of AI-driven creative disruption.

- **Bounty –** The advantages of AI in increasing productivity and prosperity could vastly outweigh the disadvantages to workers, and the outcome could be wealth and abundance for all, even those displaced by automation.

- **Rising inequality –** Even if workers displaced by AI find new jobs, the result could be rising inequality in a labor-light economy, as capital becomes more valuable and the wealthy get wealthier. As inequality widens, social and political instability could result.

- **Resource curse –** AI could lead to an economic paradox, much like the "resource curse" faced by countries abundant in natural resources. Even policy measures like universal basic income could fail to effectively translate to societal well-being and individual happiness.

- **Luddite's revenge –** A dire scenario could be massive unemployment, as the fears of the 19th century Luddites finally come true and machinery eliminates jobs that are not replaced by new ones. One effect of narrow AI could be that humans simply are not as economically valuable as they once were, much like the decline in the role of horses in the global economy following the first and second industrial revolutions.[73]

- **Generational dislocation –** Like the move from the field to the factory, AI could cause a transformation in the labor market that takes a generation to resolve. With a fundamental skills mismatch between the people who have lost jobs and the skills needed for new jobs created by AI, the result could be social and political

disruption lasting a generation. This disruption resolves itself over time as a new generation, educated and trained in the AI economy, dominates the labor market.

- **Fall behind –** Nations that fail to take advantage of AI or even resist it, for fear of potential economic and political disruption, could fall behind other countries, maintaining stability but at the cost of growth and national competitiveness.

Universal Basic Income

Fear of the large-scale dislocation potential of AI, and the enormous social and political consequences that result, is a large driver of recent discussions about the possibility of universal basic income. Universal basic income represents the idea that the government would provide income, sufficient to live on, for everyone. High-profile business leaders such as Richard Branson have argued that universal basic income might become a necessity due to AI.[74] Essentially, if the labor market implications of AI are such that new industries and possibilities for human work do not emerge, huge segments of the population could end up more or less out of work, with capital concentrated ever more in the hands of the ultra-wealthy. This would not necessarily be due to corruption or poor decision-making, just the logic of the marketplace taken to an extreme. Thus, one potential solution is to offer those who are displaced by automation the potential for a guaranteed income, given that they are unlikely to have future workplace options.[75]

Universal basic income raises many questions, of course. Who is paying in for universal basic income, and on what basis? Moreover, what about the possibility for adverse incentives? Universal basic income would essentially lessen the cost of free-riding on the system. It is also possible that universal basic income could reduce the incentive for innovation among people who otherwise would work hard to find new, productive industries where humans would have a comparative advantage over machines in an era of artificial intelligence. These are hard questions, and ones that policymakers will have to consider over the next decades.

Nationalism and International Conflict

As described elsewhere in this report, the clearest national security consequence related to the economics of AI will be the integral link between economic power and military power. It is simply not possible to maintain a leading military over time with a declining economy. The analysis above also suggests, however, that the economic, social, and political dislocation caused by AI could generate additional international security consequences.

Today, there are already political pressures in Western countries such as the United States and Great Britain that are focused on the ways the countries have changed for the worse. Automation and artificial intelligence have not yet received the blame for this, interestingly, despite the evidence presented above about the impact that automation has already had on

the labor market. Instead, political arguments in the West often focus on issues such as immigration, outsourcing, or trade deficits with countries such as China.[76] If job losses, or even just labor force instability, from artificial intelligence accelerate, it could unleash a larger wave of populism and nationalism, as wealth concentration in the hands of a smaller and smaller number of elites generates resentment and political instability. On the global stage, labor force instability at the level AI could generate has in the past led to mass turmoil, coups, and other tension, as well as the type of virulent nationalism that can generate conflict, particularly if populations blame other nations for their economic woes.

The Information Environment

Digital technologies have radically transformed the information environment in the span of only a few decades, democratizing the number of voices, expanding the volume, and accelerating the speed of societal discourse. AI will continue to change the information environment as computers become more capable of targeting information at specific users, amplifying messages, filtering information, and even generating fake audio, images, and videos. The rapid evolution of the internet, social media, and disinformation suggests it is impossible to predict how the information environment will evolve. Below are some challenges, however, that one can anticipate based on existing technology.

The End of Truth

AI has already demonstrated the ability to create audio and visual forgeries. Dr. Hany Farid, a professor of computer science at Dartmouth University who consults for the Associated Press to detect forged images and other media, has described the competition between forgery technology and authentication technology as an "arms race" and an "information war."[77] At the moment, recording and authentication technology has the upper hand, but the trends are not favorable. Society may be only a few years away from such forgeries being able to fool not just the untrained eye and ear, but sophisticated forgery detection experts and systems.[78]

This shift will bring profound implications across domains as diverse as corporate communications, courtroom evidence, journalism, and international security. Take, for instance, the Watergate scandal. President Richard Nixon maintained sufficient support in the Senate to block his removal from office even after two years of aggressive investigative reporting. Only upon the release of the "smoking gun" Oval Office audiotapes – where Nixon can be heard explicitly condoning a criminal cover-up and obstruction of justice – did his support in Congress finally fail. In a world where realistic forgeries were essentially impossible, audiotapes served not just as evidence but as undeniable proof.

AI technology could weaken, if not end, recorded evidence's ability to serve as proof. Some technologies, such as blockchain, may make it possible to authenticate the provenance of video and audio files. These technologies may not mature quickly enough, though. They could also prove too unwieldy to be used in many settings, or simply may not be enough to counteract humans' cognitive susceptibility toward "seeing is believing." The result could

be the "end of truth," where people revert to ever more tribalistic and factionalized news sources, each presenting or perceiving their own version of reality.

AI-enabled forgeries are becoming possible at the same time that the world is grappling with renewed challenges of fake news and strategic propaganda. During the 2016 U.S. presidential election, for example, hundreds of millions of Americans were exposed to fake news. The Computational Propaganda Project at Oxford University found that during the election, "professional news content and junk news were shared in a one-to-one ratio, meaning that the amount of junk news shared on Twitter was the same as that of professional news."[79] A common set of facts and a shared understanding of reality are essential to productive democratic discourse. The simultaneous rise of AI forgery technologies, fake news, and resurgent strategic propaganda poses an immense challenge to democratic governance.

Political Power, Democracy, and Authoritarianism

Due to private and public actors' ability to use AI techniques to shape information flows and perceptions, they could affect democratic processes and the strength of authoritarian regimes while also shaping global public discourse. Potential application areas include:

- **Electoral process influence** – Highly granular voter profiling, enabled by the application of AI technologies, can affect democratic norms through the electoral process. Certain advances are likely to see more narrowly targeted content creation, with bots used to amplify this messaging in targeted sub-groups. For instance, these technologies were used in targeted political ads based on the social media profiles of voters in the 2016 U.S. presidential election and the U.K. Brexit referendum.[80] Mitigation measures for this personalized propaganda – often in private messages so no public data can be gathered and scrutinized – include Facebook's pledge to make all "dark ads" on its platform public.[81]

- **Authoritarian regimes** – Social media allows authorities to manipulate the news environment and control messaging. In China today, reports estimate that the government creates and posts about 448 million social media comments a year.[82] In some cases, bots are utilized to run propaganda efforts both inside and outside a home country, with the aim of creating a strategic advantage in today's crowded information ecosystem.[83]

- **Social media** – The nature of AI makes it liable to concentrate information influence in the hands of a limited number of media platforms. Private companies not only control the data they collect, but can actively promote and demote specific content. Google, for example, de-ranks specific news outlets in its search results and only includes "publishers that are algorithmically determined to be an authoritative source of information" in its fact-checking features.[84]

- **Future targeting efforts** – Uses of AI to target audiences and spread disinformation include the expansion of automated spear phishing and sophisticated targeting of public sector employees with intent to influence government operations (i.e., orders spoofing).[85] Actors could also use AI to create "automated, hyper-personalized disinformation campaigns," in which certain key demographics or areas (i.e., swing districts) are targeted to affect voting behavior at crucial times, potentially resulting in election shaping through sophisticated AI systems.[86]

Because AI tools can be deployed at scale without large numbers of people, these tools could enable small numbers of people to wield outsize political influence, whether through governments, corporations, or other groups. The effect could be to erode the power of the people and democratic institutions and enable new forms of authoritarianism.

CONCLUSION

What world do we end up in? Does AI usher in a new era of prosperity and international peace? Does it lead to shifts in the balance of power on the global stage, with attendant risks of conflict and miscalculation? Could AI lead to massive dislocation and a rise in political unrest, nationalism, and protectionism? Does AI concentrate power to control information in the hands of a few, or continue the democratization of information that computers, networks, and social media have unleashed? Does the cacophony of competing information lead to a turn away from truth to authoritarianism and tribalism, or does the wisdom of the crowds win out with a convergence on truth and centrist policies?

The technological opportunities enabled by artificial intelligence shape the future, but do not determine it. Nations, groups, and individuals have choices about how they employ and respond to various uses of AI. Their policy responses can guide, restrict, or encourage certain uses of AI. In order to manage the challenges ahead, the United States will need to adopt a national strategy for how to take advantage of the benefits of AI while mitigating its disruptive effects.

NOTES

[1] Kevin Kelly, “The Three Breakthroughs That Have Finally Unleashed AI on the World,” *Wired,* October 27, 2014, https://www.wired.com/2014/10/future-of-artificial-intelligence/.

[2] Dustin Fraze, "Cyber Grand Challenge (CGC)," Defense Advanced Research Projects Agency, https://www.darpa.mil/program/cyber-grand-challenge.

[3] Chris Bing, "The tech behind the DARPA Grand Challenge winner will now be used by the Pentagon," Cyberscoop.com, August 11, 2017, https://www.cyberscoop.com/mayhem-darpa-cyber-grand-challenge-dod-voltron/.

[4] Gregory C. Allen and Taniel Chan, "Artificial Intelligence and National Security," Study (Belfer Center for Science and International Affairs, July 2017), 18, https://www.belfercenter.org/sites/default/files/files/publication/AI%20NatSec%20-%20final.pdf.

[5] Mohit Rajpal, William Blum, and Rishabh Singh, "Not all bytes are equal: Neural byte sieve for fuzzing," Preprint, submitted November 10, 2017, https://arxiv.org/abs/1711.04596.

[6] Nicole Nichols, Mark Raugas, Robert Jasper, and Nathan Hilliard, "Faster Fuzzing: Reinitialization with Deep Neural Models," Preprint, submitted November 8, 2017, https://arxiv.org/abs/1711.02807.

[7] John Seymour and Philip Tully, “Weaponizing data science for social engineering: Automated E2E spear phishing on Twitter,” https://www.blackhat.com/docs/us-16/materials/us-16-Seymour-Tully-Weaponizing-Data-Science-For-Social-Engineering-Automated-E2E-Spear-Phishing-On-Twitter-wp.pdf.

[8] Allen and Chan, "Artificial Intelligence and National Security."

[9] Zeynep Tufekci, “YouTube, The Great Radicalizer,” *The New York Times*, March 10, 2018, https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html.

[10] Soroush Vosoughi, Deb Roy and Sinan Aral, “The spread of true and false news online,” *Science Magazine*, 359 no. 6380 (March 9, 2018), 1146-1151.

[11] Miles Brundage et al., “The Malicious Use of Artificial Intelligence: Forecasting, Prevention and Mitigation,” (University of Oxford, February 2018), 16, https://maliciousaireport.com/.

[12] Tim Hwang, “Digital Disinformation: A Primer,” *The Atlantic Council*, September 25, 2017, 7, http://www.atlanticcouncil.org/publications/articles/digital-disinformation-a-primer.

[13] Toomas Hendrik Ilves, “Guest Post: Is Social Media Good or Bad for Democracy?”, Facebook Newsroom, January 25, 2018, https://newsroom.fb.com/news/2018/01/ilves-democracy/; and Sue Halpern, “Cambridge Analytica, Facebook and the Revelations of Open Secrets,” *The New Yorker*, March 21, 2018, https://www.newyorker.com/news/news-desk/cambridge-analytica-facebook-and-the-revelations-of-open-secrets.

[14] Michael W. Bader, “Reign of the Algorithms: How “Artificial Intelligence” is Threatening Our Freedom,” May 12, 2016, https://www.gfe-media.de/blog/wp-content/uploads/2016/05/Herrschaft_der_Algorithmen_V08_22_06_16_EN-mb04.pdf.

[15] Brundage et al., “The Malicious Use of Artificial Intelligence: Forecasting, Prevention and Mitigation,” 46.

[16] Igal Zeifman, “Bot Traffic Report 2016,” Imperva Incapsula blog on Incapsula.com, January 24, 2017, https://www.incapsula.com/blog/bot-traffic-report-2016.html.

[17] Samuel C. Woolley and Douglas R. Guilbeault, "Computational Propaganda in the United States of America: Manufacturing Consensus Online," Working paper (University of Oxford, 2017), 4, http://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2017/06/Comprop-USA.pdf; and Samuel C. Woolley and Phillip N. Howard, "Political Communication, Computational Propaganda, and Autonomous Agents," *International Journal of Communication*, 10 (2016), 4885, http://ijoc.org/index.php/ijoc/article/download/6298/1809.

[18] Woolley and Howard, "Political Communication, Computational Propaganda, and Autonomous Agents," 4885.

[19] Alessandro Bessi and Emilio Ferrara, "Social bots distort the 2016 U.S. presidential election online discussion," *First Monday*, 21 no. 11 (November 2016), 1.

[20] Travis Morris, "Extracting and Networking Emotions in Extremist Propaganda," (paper presented at the annual meeting for the European Intelligence and Security Informatics Conference, Odense, Denmark, August 22-24, 2012), 53-59.

[21] Kent Walker and Richard Salgado, "Security and disinformation in the U.S. 2016 election: What we found," Google blog, October 30, 2017, https://storage.googleapis.com/gweb-uniblog-publish-prod/documents/google_US2016election_findings_1_zm64A1G.pdf.

[22] Morris, "Extracting and Networking Emotions in Extremist Propaganda," 53-59.

[23] Craig Stewart, "Adobe prototypes 'Photoshop for audio,'" Creative Bloq, November 03, 2016, http://www.creativebloq.com/news/adobe-prototypes-photoshop-for-audio.

[24] Justus Thies et al., "Face2Face: Real-time Face Capture and Reenactment of RGB Videos," Niessner Lab, 2016, http://niessnerlab.org/papers/2016/1facetoface/thies2016face.pdf.

[25] Erica Anderson, "Building trust online by partnering with the International Fact Checking Network," Google's The Keyword blog, October 26, 2017, https://www.blog.google/topics/journalism-news/building-trust-online-partnering-international-fact-checking-network/; and Jackie Snow, "Can AI Win the War Against Fake News?" *MIT Technology Review*, December 13, 2017, https://www.technologyreview.com/s/609717/can-ai-win-the-war-against-fake-news/.

[26] "Technology," AdVerif.ai, http://adverifai.com/technology/.

[27] Onur Varol et al., "Online Human-Bot Interactions: Detection, Estimation and Characterization," Preprint, submitted March 27, 2017, 1, https://arxiv.org/abs/1703.03107.

[28] Lee Rainie, Janna Anderson, and Jonathan Albright, "The Future of Free Speech, Trolls, Anonymity, and Fake News Online," (Pew Research Center, March 2017), http://www.pewinternet.org/2017/03/29/the-future-of-free-speech-trolls-anonymity-and-fake-news-online/.; and Alejandro Bessi and Emilio Ferr, "Social bots distort the 2016 U.S. Presidential election online discussion," *First Monday*, 21 no. 11 (November 2016), http://firstmonday.org/ojs/index.php/fm/article/view/7090/5653.

[29] Adrienne Lafrance, "The Internet is Mostly Bots," *The Atlantic,* January 31, 2017, https://www.theatlantic.com/technology/archive/2017/01/bots-bots-bots/515043/.

[30] "PUBLIQ goes public: The blockchain and AI company that fights fake news announces the start of its initial token offering," PUBLIQ, November 14, 2017, https://publiq.network/en/7379D8K2.

[31] United Nations Office of Drugs and Crime, "Estimating Illicit Financial Flows Resulting from Drug Trafficking and Other Transnational Organized Crimes," Research report (October 2011), 11, https://www.unodc.org/documents/data-and-analysis/Studies/Illicit_financial_flows_2011_web.pdf.

[32] Financial services strategy lead executive at artificial intelligence for enterprise firm, interview by Edoardo Saravalle, February 20, 2018.

[33] Jonathan Estreich, "Why the Anti-Financial Crime Community Is Strongly Positioned for a Centralized Cross-Institutional Artificial Intelligence Platform," *ACAMSToday*, September 19, 2017, https://www.acamstoday.org/why-the-anti-financial-crime-community-is-strongly-positioned-for-a-centralized-cross-institutional-artificial-intelligence-platform/.

[34] Kevin Petrasic, Benjamin Saul, and Matthew Bornfreund, "The Emergence of AI RegTech Solutions for AML and Sanctions Compliance," (White & Case, April 25, 2017), https://www.whitecase.com/publications/article/emergence-ai-regtech-solutions-aml-and-sanctions-compliance.

[35] Nick J. Maxwell and David Artingstall, "The Role of Financial Information-Sharing Partnerships in the Disruption of Crime," Occasional paper (Royal United Services Institute, October 2017), vi.

[36] Kevin Dobbs, "Since Dodd-Frank, Compliance Costs Up At Least 20% For Many U.S. Banks," *S&P Global Market Intelligence*, October 27, 2017, https://marketintelligence.spglobal.com/our-thinking/ideas/since-dodd-frank-compliance-costs-up-at-least-20-for-many-u-s-banks.

[37] See, for example: Clare Ellis and Inês Sofia de Oliveira, "Tackling Money Laundering: Towards a New Model for Information Sharing," Occasional paper (Royal United Services Institute, September 2015).

[38] Dr. Gary Shiffman (Chief Executive Officer, Giant Oak) in discussion with Edoardo Saravalle, February 9, 2018.

[39] Carol Stabile, "Using Data Analytics to Identify AML Risk," *ACAMSToday*, September 19, 2017, https://www.acamstoday.org/using-data-analytics-to-identify-aml-risk/.

[40] Stuart Breslow, Mikael Hagstroem, Daniel Mikkelsen, and Kate Robu, "The new frontier in anti–money laundering," (McKinsey & Company, November 2017), https://www.mckinsey.com/business-functions/risk/our-insights/the-new-frontier-in-anti-money-laundering.

[41] See, for example, the discussion about market stability and artificial intelligence: "Artificial intelligence and machine learning in financial services: Market developments and financial stability implications," (Financial Stability Board, November 1, 2017), http://www.fsb.org/wp-content/uploads/P011117.pdf.

[42] "Spectrum Challenge," Defense Advanced Research Projects Agency, http://archive.darpa.mil/spectrumchallenge/.

[43] "The Spectrum Collaboration Challenge," Defense Advanced Research Projects Agency, 2016, https://spectrumcollaborationchallenge.com/.

[44] NeuroscienceNews, "'Mind Reading' Algorithm Uses EEG Data to Reconstruct Images Based on What We Perceive," NeuroscienceNews.co, February 22, 2018, http://neurosciencenews.com/ai-eeg-images-8546/; and Dan Nemrodov, Matthias Niemeier, Ashutosh Patel, and Adrian Nestor, "The Neural Dynamics of Facial Identity Processing:

insights from EEG-Based Pattern Analysis and Image Reconstruction, *eNeuro,* (January 29, 2018), http://www.eneuro.org/content/early/2018/01/29/ENEURO.0358-17.2018.

[45] "AP expands Minor League Baseball coverage," Associated Press, press release, June 30, 2016, https://www.ap.org/press-releases/2016/ap-expands-minor-league-baseball-coverage; and James Kotecki, "Take Me Out To the Ball Game: Ai & AP Automate Baseball Journalism At Scale," AutomatedInsights blog, July 17, 2016, https://automatedinsights.com/blog/take-automated-ball-game-next-chapter-ai-ap-partnership.

[46] Adam Najberg, "Alibaba AI Model Tops Humans in Reading Comprehension," Alizila.com, January 15, 2018, http://www.alizila.com/alibaba-ai-model-tops-humans-in-reading-comprehension/; Allison Linn, "Microsoft creates AI that can read a document and answer questions about it as well as a person," Microsoft's The AI Blog, January 15, 2018, https://blogs.microsoft.com/ai/microsoft-creates-ai-can-read-document-answer-questions-well-person/; and Tom Simonite, "AI Beat Humans At Reading! Maybe Not," *Wired*, January 18, 2018, https://www.wired.com/story/ai-beat-humans-at-reading-maybe-not/.

[47] "Snapshot: Voice Forensics Can Help the Coast Guard Catch Hoax Callers," DHS.gov, September 26, 2017, https://www.dhs.gov/science-and-technology/news/2017/09/26/snapshot-voice-forensics-can-help-coast-guard-catch-hoax.

[48] "Passenger Screening Algorithm Challenge," kaggle.com, https://www.kaggle.com/c/passenger-screening-algorithm-challenge.

[49] "A.I. Could Be a Firefighter's 'Guardian Angel,'" National Aeronautics and Space Administration, https://technology.nasa.gov/features/audrey.html.

[50] "Our Mission," DHS.gov, May 11, 2016, https://www.dhs.gov/our-mission.

[51] Jean Kumagai, "This AI Hunts Poachers," *IEEE Spectrum,* January 6, 2018, https://spectrum.ieee.org/robotics/artificial-intelligence/this-ai-hunts-poachers.

[52] "Prize Challenges: Geopolitical Forecasting Challenge," Intelligence Advanced Research Projects Activity, https://www.iarpa.gov/index.php/working-with-iarpa/prize-challenges/1070-geopolitical-forecasting-challenge.

[53] Justus Thies et al., "Face2Face: Real-time Face Capture and Reenactment of RGB Videos," Niessner Lab, 2016, http://niessnerlab.org/papers/2016/1facetoface/thies2016face.pdf; and for other malicious use of AI, see: Brundage et al., "The Malicious Use of Artificial Intelligence: Forecasting, Prevention and Mitigation."

[54] United Nations Conference on Trade and Development, "The <<New>> Digital Economy and Development," UNCTAD Technical Notes on ICT for Development No. 8 (October 2017), http://unctad.org/en/PublicationsLibrary/tn_unctad_ict4d08_en.pdf.

[55] See, for example, models that would greatly benefit from applied machine learning: United Nations Conference on Trade and Development, "The <<New>> Digital Economy and Development"; and Ernest Mwebaze, Washington Okori, and John A. Quinn, "Causal Structure Learning for Famine Prediction," 2010, http://ai-d.org/pdfs/Mwebaze.pdf.

[56] "Farm Population Lowest Since 1850's," *The New York Times*, 1988, https://www.nytimes.com/1988/07/20/us/farm-population-lowest-since-1850-s.html.

[57] James Manyika et al., "What the future of work will mean for jobs, skills, and wages," Report (McKinsey Global Institute Report, November 2017), https://www.mckinsey.com/global-

themes/future-of-organizations-and-work/what-the-future-of-work-will-mean-for-jobs-skills-and-wages.

[58] Carl B. Frey and Michael A. Osborne, "The future of employment: how susceptible are jobs to computerisation?," *Technological Forecasting and Social Change,* 114 (2017), 254-280.

[59] Ljubica Nedelkoska and Glenda Quintini, "Automation, Skill Use and Training," *OECD Social, Employment and Migration Working Papers*, No. 202, OECD Publishing, Paris, 2018: http://dx.doi.org/10.1787/2e2f4eea-en, 47.

[60] Daron Acemoglu and Pascual Restrepo, "Robots and Jobs: Evidence from US Labor Markets," Working paper (National Bureau of Economic Research, March 17, 2017), 4.

[61] Erin Winick, "Every study we could find on what automation will do to jobs, in one chart," *MIT Technology Review*, January 25, 2018, https://www.technologyreview.com/s/610005/every-study-we-could-find-on-what-automation-will-do-to-jobs-in-one-chart/.

[62] James Manyika et al., "What the future of work will mean for jobs, skills, and wages."

[63] Joseph A. Schumpeter, *Capitalism, Socialism, and Democracy* (New York, NY: Harper & Row, 1942).

[64] Christopher Matthews, "Summers: Automation is the middle class' worst enemy," Axios, June 4, 2017, https://www.axios.com/summers-automation-is-the-middle-class-worst-enemy-1513302420-754facf2-aaca-4788-9a41-38f87fb0dd99.html.

[65] "Number of Facebook employees from 2007 to 2017 (full-time)," Statista.com, 2018, https://www.statista.com/statistics/273563/number-of-facebook-employees/; and "Number of Ford employees from FY 2011 to FY 2017 (in 1,000s)," Statista.com, 2018, https://www.statista.com/statistics/297324/number-of-ford-employees/.

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

[67] Brynjolfsson and McAfee, *The Second Machine Age,* Chapter 9.

[68] Jon T. Sumida, *In Defence of Naval Supremacy: Finance, Technology, and British Naval Policy, 1889-1914* (Winchester, MA: Unwin Hyman, 1989).

[69] *Where machines could replace humans – and where they can't yet*: *Automation by Sector*, (McKinsey Global Institute, 2017), https://public.tableau.com/profile/mckinsey.analytics#!/vizhome/AutomationBySector/WhereMachinesCanReplaceHumans.

[70] *Labor Force Statistics from the Current Population Survey*, Household Data Annual Averages 18b, Employed persons by detailed industry and age (Bureau of Labor Statistics, 2018), https://www.bls.gov/cps/cpsaat18b.htm; and *Labor Force Statistics from the Current Population Survey*, Household Data Annual Averages 18, Employed persons by detailed industry, sex, race, and Hispanic or Latino ethnicity (Bureau of Labor Statistics, 2018), https://www.bls.gov/cps/cpsaat18.htm.

[71] We arrived at an estimate for the automation potential for each age group by taking a weighted average of the automation potential for each age group in each sector of the economy. Additionally, Bureau of Labor Statistics data includes part-time workers while the McKinsey analysis only examined full-time jobs, which could introduce some discrepancies.

[72] Brynjolfsson and McAfee, *The Second Machine Age,* Chapter 9, Figure 9.2.

[73] "Humans Need Not Apply," *YouTube,* August 13, 2014, https://www.youtube.com/watch?v=7Pq-S557XQU.

[74] Catherine Clifford, "Billionaire Richard Branson: A.I. is going to eliminate jobs and free cash handouts will be necessary," CNBC, February 20, 2018, https://www.cnbc.com/2018/02/20/richard-branson-a-i-will-make-universal-basic-income-necessary.html.

[75] James J. Hughes, "A Strategic Opening for a Basic Income Guarantee in the Global Crisis Being Created by AI, Robots, Desktop Manufacturing and Biomedicine," *Journal of Evolution and Technology.* 24 no. 1 (February 2014), 45-61.

[76] For example, see Andrea Cerrato, Francesco Ruggieri and Federico Maria Ferrara, "Trump won in counties that lost jobs to China and Mexico," *The Washington Post's Monkey Cage Blog*, December 2, 2016, https://www.washingtonpost.com/news/monkey-cage/wp/2016/12/02/trump-won-where-import-shocks-from-china-and-mexico-were-strongest/?utm_term=.08a4bfed59fe; and Paul Wiseman, "Mexico taking U.S. factory jobs? Blame robots instead," *PBS News Hour*, November 2, 2016, https://www.pbs.org/newshour/economy/factory-jobs-blame-robots.

[77] Amelia Tait, "How to identify if an online video is fake," NewStatesman.com, February 14, 2018, https://www.newstatesman.com/science-tech/technology/2018/02/how-identify-if-online-video-fake.

[78] Gregory C. Allen, "Artificial Intelligence Will Make Forging Anything Entirely Too Easy," *Wired*, July 1, 2017. https://www.wired.com/story/ai-will-make-forging-anything-entirely-too-easy/.

[79] Rory Clarke and Balazs Gyimesi, "Digging up facts about fake news: The Computational Propaganda Project," Organisation for Economic and Co-operation and Development, 2017, https://www.oecd.org/governance/digging-up-facts-about-fake-news-the-computational-propaganda-project.htm.

[80] Ilves, "Guest Post: Is Social Media Good or Bad for Democracy?"

[81] Rob Goldman, "Update on Our Advertising Transparency and Authenticity Efforts," Facebook Newsroom, October 27, 2017, https://newsroom.fb.com/news/2017/10/update-on-our-advertising-transparency-and-authenticity-efforts/.

[82] Gary King, Jennifer Pan, and Margaret E. Roberts, "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, not Engaged Argument," *American Political Science Review*, 111 no. 3 (2017), 1.

[83] Nicholas Confessore et al., "The Follower Factory," *The New York Times*, January 27, 2018, https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html.

[84] Justin Kosslyn and Cong Yu, "Fact Check now available in Google Search and News around the world," Google's The Keyword blog, April 7, 2017, https://blog.google/products/search/fact-check-now-available-google-search-and-news-around-world/.

[85] John Seymour and Philip Tully, "Weaponizing data science for social engineering: Automated E2E spear phishing on Twitter" (paper presented at the annual meeting for the Black Hat Conference, Las Vegas, Nevada, August 3-4, 2016), 2, https://www.blackhat.com/docs/us-16/materials/us-16-Seymour-Tully-Weaponizing-Data-Science-For-Social-Engineering-Automated-E2E-Spear-Phishing-On-Twitter-wp.pdf.

[86] Brundage et al., "The Malicious Use of Artificial Intelligence: Forecasting, Prevention and Mitigation," 29.