Center for a
New American
Security

May 22, 2018

# Testimony before the Tom Lantos Human Rights Commission

# Artificial Intelligence: The Consequences for Human Rights

Paul Scharre, Senior Fellow and Director
Technology and National Security
Center for a New American Security

Chairman Hultgren, Chairman McGovern, and distinguished members, thank you for inviting me to testify today.

Recent years have seen rapid advances in artificial intelligence and machine learning. AI tools are now coming out of research labs and into the real world, and are reshaping a variety of industries – medicine, transportation, finance, cybersecurity, and more. AI will similarly have important applications to human rights.

Artificial intelligence is a general-purpose enabling technology, much like electricity, computers, or networks. AI will be used by state and non-state actors for a variety of purposes, some of which will no doubt include suppressing human rights. Other uses may help to enhance human rights or fight against repressive regimes. It is not my intention today to estimate what the net effect of AI technology will be for human rights. Rather, I would like to walk through some features of the technology and some potential use cases to help illustrate some of the possibilities as AI technology becomes more widely used.

## Applications of Artificial Intelligence

AI tools can be used for a variety of applications. Some examples include:

- **Data classification**, such as identifying images, classifying song genres, or arriving at medical diagnoses.[1] Given a sufficiently large set of training data, algorithms can be trained to classify data extremely accurately, often better than humans.
- **Anomaly detection**, such as finding fraudulent financial transactions or new forms of malware.[2] Traditional methods of anomaly detection require looking for known signatures. However, new AI tools can find anomalies whose signatures are not yet known by analyzing routine patterns of data and then identifying new data that is outside the norm. These systems can be used to monitor large data streams, such as financial transactions, at scale and in real-time in ways that would not be feasible for humans.

- **Prediction**, such as making statistical predictions about future behavior based on large datasets. Systems of this type are already widely used commercially, such as recommendation algorithms in Netflix and Amazon and search engine auto-fills. Other uses raise difficult ethical issues, such as predictive policing or predicting patient longevity in end-of-life care.[3]
- **Optimization,** such as improving performance and efficiency in industrial systems. Given a known goal, such as saving energy or reducing costs, AI systems can often find novel solutions to problems.[4]

## AI and Threats to Human Rights

In the hands of a repressive state with access to large datasets about its population, these tools could be used to further increase state control. Automated facial recognition technology combined with security cameras could make *1984*-style continuous monitoring feasible in metropolitan areas. Combined with other readily available digital data collected through computers and smartphones, AI tools could be used to comprehensively monitor a person's behavior, communications, likes, and desires at a scale not even Orwell could have imagined.

People living in the digital age create a cornucopia of data: smartphone geolocation, browser history, web search history, online purchases, contacts, social media engagements, email and text message content, telephone calls, and more. Whoever has access to this data has tremendous insight not only into a person's past, but also the ability to predict their future behavior. Without AI tools, though, it is hopelessly impractical to manage this data at scale.

AI makes much of this data more discoverable through data classification tools that can recognize faces, identify human emotions, translate voice to text, translate languages, and process language. AI tools also make it feasible to analyze and process this data at scale. This means that the kind of intrusive monitoring that would in the past have been extremely time-consuming and resource-intensive can now be done quickly and at scale, allowing far more extensive and intrusive monitoring of a population.

Moreover, large datasets can be aggregated to generate statistically valid predictions. By learning from data across an entire population and then applying this to readily available information about an individual, AI tools could be used to make predictions about that individual's preferences or behavior – political, financial, sexual, or other. AI tools could be used to not only monitor a population, but predictively crack down on would-be dissidents.

## AI to Assist Human Rights

At the same time, there are a number of features of AI tools that would make them powerful allies for those fighting repressive regimes. AI systems embed expertise within the software itself, lowering the bar the skills needed for a given capability. One does not need to spend years learning chess anymore to play at the level of a grandmaster; one can merely download a chess app for free. Similarly, AI systems will put greater abilities in the hands of non-state groups and individuals. Smartphones already turn surveillance tools back against the state, allowing citizens to record abuses by authorities. AI tools such as embedded object recognition or facial recognition in the hands of everyday citizens could make it even easier to identify abusers and hold perpetrators to account.

A core feature of information technology is that it renders the costs of copying and transmitting information close to zero. One of the consequences of this is that it is difficult to keep information secret. While this is true for personal information, it is also true for state secrets. Individuals have accessed and released large tranches of government secrets on a scale that was impossible in the pre-digital era. The ease with which information freely flows in the digital age is a hindrance to repressive regimes that thrive on secrecy.

AI tools will make it easier for individuals and non-state organizations to process and analyze this data. In January 2018, a student at the Australian National University pointed out that "heat maps" of jogging routes from runners wearing geo-locating Fitbits could be used to identify secret military and intelligence bases overseas.[5] Journalists quickly discovered that they could de-anonymize the data and actually identify specific users who had run routes as well as previous locations they had visited.[6] This analysis was done manually, but AI tools could make it easier to process this data at scale, including linking it with other datasets such as social media profiles.

Embedding expertise within the software allows for greater automation, which can expand the scale at which smaller groups can achieve effects. For example, a few individuals have been able to cause significant internet disruption for short periods of time using botnets to infect Internet of Things (IoT) devices and launch distributed denial of service (DDoS) attacks.[7] Automation may allow small groups to achieve outsize effects, which levels the playing field against powerful actors and may be helpful in combatting repressive states.

## Conclusion

A key question for any new technology is whether it concentrates power in the hands of a few or democratizes power towards the many. AI has features of both. At present, large datasets are needed to train AI systems. Additionally, the most cutting-edge advances in AI require significant computing resources.[8] At the same time, many AI tools are freely available for download online,[9] and much data is openly available. Artificial intelligence will enable actors who both seek to enhance human rights and those who aim to repress them.

## CNAS Funding

CNAS is a national security research and policy institution committed to the highest standards of organizational, intellectual, and personal integrity. The Center retains sole editorial control over its ideas, projects, and productions, and the content of its publications reflects only the views of their authors. In keeping with its mission and values, CNAS does not engage in lobbying activity and complies fully with all applicable federal, state, and local laws. Accordingly, CNAS will not engage in any representation or advocacy on behalf of any entities or interests and, to the extent that the Center accepts funding from foreign sources, its activities will be limited to bona fide scholastic, academic, and research-related activities, consistent with applicable federal law. A full list of CNAS supporters and the center's funding guidelines can be found at https://www.cnas.org/support-cnas.

PAUL SCHARRE is a Senior Fellow and Director of the Technology and National Security Program at the Center for a New American Security. He is the author of *Army of None: Autonomous Weapons and the Future of War*, published in April 2018.

From 2008-2013, Mr. Scharre worked in the Office of the Secretary of Defense (OSD) where he played a leading role in establishing policies on unmanned and autonomous systems and emerging weapons technologies. Mr. Scharre led the DoD working group that drafted DoD Directive 3000.09, establishing the Department's policies on autonomy in weapon systems. Mr. Scharre also led DoD efforts to establish policies on intelligence, surveillance, and reconnaissance (ISR) programs and directed energy technologies. Mr. Scharre was involved in the drafting of policy guidance in the *2012 Defense Strategic Guidance*, *2010 Quadrennial Defense Review*, and Secretary-level planning guidance. His most recent position was Special Assistant to the Under Secretary of Defense for Policy.

Prior to joining OSD, Mr. Scharre served as a special operations reconnaissance team leader in the Army's 3rd Ranger Battalion and completed multiple tours to Iraq and Afghanistan. He is a graduate of the Army's Airborne, Ranger, and Sniper Schools and Honor Graduate of the 75th Ranger Regiment's Ranger Indoctrination Program.

Mr. Scharre has published articles in *The New York Times, Wall Street Journal, CNN, TIME, Foreign Policy, Foreign Affairs, Politico,* and *The National Interest,* and has appeared on CNN, MSNBC, Fox News, NPR, and the BCC.  He has testified before the House and Senate Armed Services Committees and has presented at the United Nations, NATO, the Pentagon, the CIA, and other national security venues. Mr. Scharre is a term member of the Council on Foreign Relations. He holds an M.A. in Political Economy and Public Policy and a B.S. in Physics, cum laude, both from Washington University in St. Louis.

## Notes

[1] Christopher Mims, "Using Neural Networks to Classify Music," *MIT Technology Review*, June 03, 2010, https://www.technologyreview.com/s/419223/using-neural-networks-to-classify-music/; Tom LH. Li, Antoni B. Chan, and Andy HW. Chun, "Automatic Musical Pattern Feature Extraction Using Convolutional Neural Network" (paper presented at the proceedings of the International MultiConference of Engineers and Computer Scientists 2010, Hong Kong, March 17-19, 2010), http://www.iaeng.org/publication/IMECS2010/IMECS2010_pp546-550.pdf; and Dave Fornell, "How Artificial Intelligence Will Change Medical Imaging," ImagingTechnologyNews.com, February 24, 2017, https://www.itnonline.com/article/how-artificial-intelligence-will-change-medical-imaging.

[2] Efstathios Kirkos, Charalambos Spathis, and Yannis Monolopoulous, "Data Mining Techniques for the detection of fraudulent financial statements," *Expert Systems with Applications,* 32 (2007), 995-1003, http://delab.csd.auth.gr/papers/ESWA07ksm.pdf; and "DeepArmor: A cognitive approach to endpoint protection," SparkCognition.com, https://www.sparkcognition.com/deeparmor-enterprise/.

[3] Randy Rieland, "Artificial Intelligence Is Now Used to Predict Crime. But Is It Biased?," Smithsonian.com, March 05, 2018, https://www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968337/; and Anand Avanti et al., "Improving Palliative Care with Deep Learning," Preprint, submitted November 17, 2017, https://arxiv.org/pdf/1711.06402.pdf.

[4] Richard Evans and Jim Gao, "DeepMind AI Reduces Google Data Centre Cooling Bill by 40%," DeepMind blog, July 20, 2016, https://deepmind.com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-40/.

[5] Jeremy Hsu, "The Strava Heat Map and the End of Secrets," *Wired*, January 29, 2018, https://www.wired.com/story/strava-heat-map-military-bases-fitness-trackers-privacy/.

[6] Matt Burgess, "Strava's data lets anyone see the names (and heart rates) of people exercising on military bases," *Wired*, January 30, 2018, http://www.wired.co.uk/article/strava-military-bases-area-51-map-afghanistan-gchq-military; Alex Hern, "Strava suggests military users 'opt out' of heatmap as row deepens," *The Guardian*, January 29, 2018, https://www.theguardian.com/technology/2018/jan/29/strava-secret-army-base-locations-heatmap-public-users-military-ban.

[7] Josh Fruhlinger, "The Mirai botnet explained: How teen scammers and CCTV cameras almost brought down the internet," CSOonline, March 9, 2018, https://www.csoonline.com/article/3258748/security/the-mirai-botnet-explained-how-teen-scammers-and-cctv-cameras-almost-brought-down-the-internet.html; AFP, "Disgruntled Gamer 'Likely' Behind October US Hacking: Expert," November 17, 2016, https://www.securityweek.com/disgruntled-gamer-likely-behind-october-us-hacking-expert; United States Attorney's Office, District of New Jersey, "Justice Department Announces Charges And Guilty Pleas In Three Computer Crime Cases Involving Significant Cyber Attacks," December 13, 2017, https://www.justice.gov/usao-nj/pr/justice-department-announces-charges-and-guilty-pleas-three-computer-crime-cases.

[8] OpenAI, "AI and Compute," OpenAI blog, May 16, 2018, https://blog.openai.com/ai-and-compute/.

[9] For example, see "TensorFlow," https://www.tensorflow.org/.