The Rise of Artificial Intelligence: Risks from the Perspective of Mass Atrocities

Published by the Budapest Center for Mass Atrocities Prevention, September 2020

Authors: Olivia Dudley *(former Research Assistant)*, Madalena Ausshauer *(former Research Assistant)*, Marine Milard *(Research Assistant)* and Sophie Smith *(Research Assistant)*

Abstract. Artificial intelligence (AI) is becoming commonplace in every aspect of society at an accelerating rate. However, the rise of AI does not come without challenges. The aim of this paper, therefore, is to raise awareness of some risks AI represents in the context of mass atrocities. The list is non-exhaustive. It includes the exploitation of sensitive data, the question of ethics, cybercrime and warfare, programmes and models run by AI. The paper is written by the team of interns in the Budapest Centre and it mainly targets young people to introduce them to the topic of mass atrocities in relation to AI; however, we also welcome academics to utilize this work for their purposes.

The authors hope that the document will also contribute to the research planned by the Budapest Centre within the Initiative "Multipolar Task Force."[1]

[1]The Budapest Centre for Mass Atrocities Prevention, "Towards a Multipolar Task Force to Prevent Mass Atrocities in The Era of Geopolitical Change, Climate Threats to Security and Digital Revolution," *The Budapest Centre for Mass Atrocities Prevention*, July 22, 2020. https://www.genocideprevention.eu/climate-threats-to-security-and-digital-revolution/.

Over the past decade, digital transformations have significantly altered and re-invented the social order around the world. New technologies have permeated all aspects of society, ranging from communication and education to global governance. Within this framework, artificial intelligence (AI) is finding an ever-larger role to play as one of the fastest growing areas in sciences and technology. It is hard to foresee its implications, but we must be ready for addressing its challenges in all fields. The technology has certainly been a beneficial tool for many industries and governments. At the same time, it has the potential to become a dangerous weapon which may instigate and exacerbate crimes of mass atrocities.[1]

From the perspective of mass atrocities, this study identifies several key risks as follows to the use of AI (the list is not exhaustive):

Manipulation of data

Given that AI is able to collect vast amounts of personal data, AI technologies may collect and use sensitive data for manipulation of populations, for example in cases of polls and elections. Big data collected through AI can be used for micro-targeting groups of the population to maximize the effectiveness of campaigns. This was evident in the US presidential election of 2016 in which citizens received different adverts depending on their individual psychology, as well as were subject to fake news and social media 'bots' engineered, according to numerous public sources, by Russia. The use of AI for such purposes allows for undermining the sovereignty of states and democracy systems, as well as enables foreign actors to influence individuals to act against the interests of their nations.[2]

Production of biased results

AI can create problems when they deliver biased results. This can happen, for example, if a population is underrepresented in the data used to train the model. In turn, this may exacerbate existing societal inequalities and polarization within society.[3] This has been documented in an analysis of the facial recognition systems in the United States where 200 facial recognition algorithms had worse performance in their ability to successfully recognize and match faces of minorities. Most systems developed in the US had a "higher rate of false positive matches for Asian and African American faces over Caucasian faces, sometimes by a factor of 10 or even 100". Put more simply, the facial recognition systems were more likely to find a match even when there was not one. The false positives have a strong impact on prosecuting alleged criminals and create a high risk for these minorities of falsely being accused of and prosecuted for crimes in a country with the highest incarceration rate in the world.[4]

---

[2] United Nations General Assembly, Rome Statute of the International Criminal Court (July 17, 1998), https://www.icc-cpi.int/NR/rdonlyres/ADD16852-AEE9-4757-ABE7-9CDC7CF02886/283503/RomeStatutEng1.pdf.

[3] Benjamin Cheatham, Kia Javanmardian and Hamid Samandari, "Confronting the risks of artificial intelligence," McKinsey & Company, April 26, 2019, accessed July 7, 2020. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/confronting-the-risks-of-artificial-intelligence.

[4] Karon Hao, "A US government study confirms most face recognition systems are racist," MIT Technology Review, December 20, 2019, accessed July 7, 2020. https://www.technologyreview.com/2019/12/20/79/ai-face-recognition-racist-us-government-nist-study/.

Polarization of populations by algorithms

There have been several instances where AI algorithms have proven prejudiced against certain groups. For example, search engines, such as Google, and social media websites, such as Facebook, have shown a racial bias against colored people when analyzing images.[5] A former Amazon recruitment engine presented a bias against women, rating male candidates more highly.[6] Such dynamics generate divide and may fuel prejudices and hatred along identity lines within societies. Moreover, the misuse of the algorithms may facilitate the marginalization of disadvantaged groups and create a conducive environment to generate hatred and conflicts with extreme consequences.

Contribution to inequalities in societies

As AI technologies will be able to replace workers, particularly in low-skilled jobs this may generate substantial unemployment among lower income citizens who, on the top of that lack usually the opportunities for retraining.[7] High income groups, on the other hand, often hold professions that are less vulnerable to AI and have greater access to retraining so they remain less affected by the innovation.[8] That will likely lead to increasing income inequalities that amplify social inequalities which represent fertile soil for tensions within societies and could be misused for fueling hatred along identity lines.

Tool against "unwanted" groups

Systems of mass surveillance using AI, such as the IJOP enable governmental authorities to use "mass DNA collection and facial recognition cameras" as well as microphones in the street to keep groups of the population under control.[9] Through everyday use of phones, pulling data from geolocation, contacts, internet searches, and other functions, the activities of specific, "unwanted" groups can be monitored. The IJOP application can classify and report to the governmental authorities 36 "suspicious" activities ranging from having contact with criminals to excessive use of electricity and owning too many books. For example, the Chinese government has been using AI to identify, track and follow the Uyghur population and the Muslim population in Xinjiang,

---

[5] Nicolas Kayser-Bril, "Google apologizes after its Vision AI produced racist results," *Algorithm Watch,* April 7, 2020, accessed September 28, 2020, https://algorithmwatch.org/en/story/google-vision-racism/; Olivia Solon, "Facebook ignored racial bias research, employees say," NBC News, July 23, 2020, accessed September 28, 2020. https://www.nbcnews.com/tech/tech-news/facebook-management-ignored-internal-research-showing-racial-bias-current-former-n1234746.

[6] Jeffry Daston, "Amazon scraps secret AI recruiting tool that showed bias against women," *Reuters,* October 11 2018, accessed September 28, 2020. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G.

[7] Anton Korinek and Joseph E. Stiglitz, "Artificial Intelligence and Its Implications for Income Distribution and Unemployment," *National Bureau of Economic Research,* December 2017, https://www.nber.org/papers/w24174.

[8] Anton Korinek and Joseph E. Stiglitz, "Artificial Intelligence and Its Implications for Income Distribution and Unemployment."

[9] Konstantin Salomatin and Shura Burtin, "Beyond Orwell's Worst Nightmares: How China Uses Artificial Intelligence to Commit Genocide," Bylinetimes, November 18, 2019, accessed July 7, 2020, https://bylinetimes.com/2019/11/18/beyond-orwells-worst-nightmares-how-china-uses-artificial-intelligence-to-commit-genocide/.

West China, and with reference to the results, individuals belonging to these groups are prosecuted and sent to so called "re-education" camps.[10] The example illustrates that AI can be misused and applied as a primary accomplice for transforming an intention against an „unwanted" group into persecution, detention that may lead ultimately, even to mass extermination.

New threats by hacking and programming

Gaining possession of AI technology can enable extreme individual actors or organizations to manipulate military programs for example, through changing the initial targets or objectives so that intentionally hurt civilians or even its own military personnel. The Franchiser Foundation assesses AI as a new challenge for national security concepts and points that hacking and collecting top-secret, sensitive data or manipulating AI technologies and programs enable the users to target specific groups and/or minorities of civilians.[11]

Security risks also depend on whether AI is programmed for harmful or beneficial purposes. The response to the risk is relatively simple if a program is known clearly to be destructive. In that case, its use can be controlled, limited or even prohibited. The risk becomes serious, however, when AI is intended to be beneficial but the manner in which it was programmed ends up being inadvertently destructive.[11]

Ethical question of responsibility and liability in military operations

When AI is used in military operations, the elimination of opponents is oftentimes no longer solely a human decision, which, in turn, makes it challenging to determine the responsible individual or party in the aftermath of such attacks. With incorporating AI technologies into military technologies, the risk of dehumanizing death comes in the forefront to an even larger extent. Given the lack of the ability to make critical and empathic decisions in military operations when applying AI, the new technology could cause the death of masses of innocent civilian lives and ultimately, act against the military's objective. An example of that is the use of drone warfare in the US military operations in Afghanistan or Iraq.[12]

---

[10] Alexander Ma, "China uses an intrusive surveillance app to track its Muslim minority, with technology that could be exported to the rest of the world. Here's how it works," *Business Insider France*, May, 11, 2019, accessed July 7, 2020, https://www.businessinsider.fr/us/how-ijop-works-china-surveillance-app-for-muslim-uighurs-2019-5.

[11] Ewa Dönitz, Erduana Shala, Timo Leimbach, Antje Bierwisch, Sonja Grigoleit, Joachim Klerx, *Complete narrative threat scenarios produced through the scenario development of Task 4.4* (Oslo: Fraunhofer Institute for Systems and Innovation Research ISI, 2013), https://www.isi.fraunhofer.de/content/dam/isi/dokumente/ccv/2013/ETTIS_Deliverable_4_4_Catalogue%20of%20Threat%20Scenarios.pdf.

[12] Michael Bruch and Ralf Grosser, *The rise of artificial intelligence - Future outlooks and emerging risks* (Munich: Allianz, 2018), https://www.agcs.allianz.com/news-and-insights/reports/the-rise-of-artificial-intelligence.html.

References

Cheatham, Benjamin, Kia Javanmardian and Hamid Samandari. "Confronting the risks of artificial intelligence," McKinsey & Company, April 26, 2019. Retrieved from: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/confronting-the-risks-of-artificial-intelligence.

Daston, Jeffry. "Amazon scraps secret AI recruiting tool that showed bias against women." *Reuters,* October 11 2018. Accessed September 28, 2020. Retrieved from: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G.

Dönitz, Ewa. Erduana Shala, Timo Leimbach, Antje Bierwisch, Sonja Grigoleit, Joachim Klerx. C*omplete narrative threat scenarios produced through the scenario development of Task 4.4.* Oslo: Fraunhofer Institute for Systems and Innovation Research ISI, 2013. Retrieved from: https://www.isi.fraunhofer.de/content/dam/isi/dokumente/ccv/2013/ETTIS_Deliverable_4_4_Catalogue%20of%20Threat%20Scenarios.pdf.

Hao, Karon. "A US government study confirms most face recognition systems are racist." MIT Technology Review, December 20, 2019. Retrieved from: https://www.technologyreview.com/2019/12/20/79/ai-face-recognition-racist-us-government-nist-study/.

Kayser-Bril, Nicolas. "Google apologizes after its Vision AI produced racist results." *Algorithm Watch,* April 7, 2020. Accessed September 28, 2020. Retrieved from: https://algorithmwatch.org/en/story/google-vision-racism/.

Korinek, Anton and Joseph E. Stiglitz. "Artificial Intelligence and Its Implications for Income Distribution and Unemployment." *National Bureau of Economic Research,* December 2017. Retrieved from: https://www.nber.org/papers/w24174.

Ma, Alexander. "China uses an intrusive surveillance app to track its Muslim minority, with technology that could be exported to the rest of the world. Here's how it works," Business Insider France, May, 11, 2019. Retrieved from: https://www.businessinsider.fr/us/how-ijop-works-china-surveillance-app-for-muslim-uighurs-2019-5.

Salomatin, Konstantin and Shura Burtin. "Beyond Orwell's Worst Nightmares: How China Uses Artificial Intelligence to Commit Genocide." Bylinetimes, November 18, 2019. Retrieved from: https://bylinetimes.com/2019/11/18/beyond-orwells-worst-nightmares-how-china-uses-artificial-intelligence-to-commit-genocide/.

Solon, Olivia. "Facebook ignored racial bias research, employees say." *NBC News*, July 23, 2020. Accessed September 28, 2020. Retrieved from:

https://www.nbcnews.com/tech/tech-news/facebook-management-ignored-internal-research-showing-racial-bias-current-former-n1234746.

The Budapest Centre for Mass Atrocities Prevention. "Towards a Multipolar Task Force to Prevent Mass Atrocities in The Era of Geopolitical Change, Climate Threats to Security and Digital Revolution." *The Budapest Centre for Mass Atrocities Prevention*, July 22, 2020. https://www.genocideprevention.eu/climate-threats-to-security-and-digital-revolution/.

United Nations General Assembly, Rome Statute of the International Criminal Court. July 17, 1998. https://www.icc-cpi.int/NR/rdonlyres/ADD16852-AEE9-4757-ABE7-9CDC7CF02886/283503/RomeStatutEng1.pdf.