Title: Artificial Intelligence and Human Rights in an Age of Exponential Growth

Thomas Sarsfield
“I propose to consider the question, ‘Can machines think?’” So asked Alan Turing in his landmark essay “Computing Machinery and Intelligence” in 1950. Today, increasingly powerful ‘thinking machines’ exert a discrete yet powerful influence over our politics, economy, and day-to-day lives. These artificial intelligence (AI) technologies will not only become more sophisticated with time, but do so at an exponential rate of advance. In order for the Universal Declaration of Human Rights (UDHR) to remain relevant in the 21st century, the international community will need to address the unique challenges posed by AI to the rights of humankind.

If oil and coal were the lifeblood of industrial economies, then data and algorithms are the lifeblood of our contemporary digital economy. Even as you read this paper, the phone in your pocket is collecting a swathe of information about you, your colleagues, and where you are (Fussell 2022). Facebook is identifying people near you in order to make friend recommendations. TikTok is using audio data to spatially map the room you are in, while Google is using your location to determine traffic conditions. This passive collection of vast amounts of personal data is a defining feature of ‘surveillance capitalism’, which is “the unilateral claiming of private human [experiences] as free raw material for translation into behavioral data” (Laidler 2019).

This mass collection of data poses challenges for Article 12 of the UDHR, which establishes a right against “arbitrary interference” with one’s privacy. For instance, China is collecting massive amounts of data provided by a network of CCTV cameras and facial recognition software. This information is then analyzed by AI in order to create a ‘social credit system’ which is used to penalize or reward citizens based on their adherence to principles of the Chinese Communist Party. Such AI-enabled privacy violations increase the ease with which governments can violate other articles of the UDHR. Article 18 ensures a person’s right to
freedom of movement, but the Chinese government can deny citizens the ability to travel if their social credit score reaches a low enough level (Kobie 2019).

Privacy is not the only concern for the UDHR, however. Article 2 of the UDHR prohibits discrimination. Despite this racial bias is often manifested in code. For example, when the term ‘black girls’ is entered into Google images, users are more likely to see sexually explicit images (Noble 2018). AI has also been used by courts to make risk assessments during sentencing hearings. These algorithms are more likely to recommend stiff sentences for minority offenders (Angwin et al, 2016). When machines think, they are susceptible to human faults because their ‘thinking’ is determined by humans with their own biases.

To combat these challenges, governments and international organizations need to view the UDHR through the lens of the 21st century. To tackle privacy and discrimination issues, governments need to strengthen data protection laws in order to ensure AI is not misused by powerful actors. Currently, data protection is largely achieved through a “notice-and-consent” model where companies inform users of how their data may be collected and used through lengthy terms and conditions banners consumers rarely read (Kerry 2020). However, focusing instead on the types of data companies can collect and the manner in which it is used would provide regulators with more flexibility to directly tackle privacy and discrimination issues.

However, other challenges require an update to the UDHR itself. First, an article which protects against emotional manipulation by AI should be added to the UDHR. In 2014, it was revealed Facebook used algorithms to affect users’ moods by priming them with specific types of content (Goel 2014). No company or government should be able to design a powerful, sophisticated tool that allows them to autonomously control the feelings of other human beings. AI manipulation is more dangerous than traditional forms of propaganda in the sense that it is far
more subtle. While a discerning citizen can decide whether or not to take state-sponsored media at face value, it is comparatively much more difficult to know if your social media feed is being quietly manipulated to make you sad, angry, or happy.

Second, an article to enshrine the right to human interaction should also be included. This has two powerful effects. First, at a time when loneliness and mental health are increasingly prominent issues, it ensures nobody will be forced to go through a day or receive essential services without human interaction. For example, the United Kingdom’s National Health Service has invested money into researching the use of therapeutic AI chatbots for patients experiencing mental distress (Browne 2022). While this may or may not be an adequate short-term solution for patients as they wait to see a provider, a healthcare system should never be allowed to only offer therapeutic services via AI chatbot. The second effect would reinforce Article 23 of the UDHR, which guarantees a right to employment. As AI becomes more sophisticated, administrative tasks in the customer service sector will become increasingly automated (Birnbaum 2019). A universal right to human interaction ensures someone going to the store will be able to request human assistance, which will protect workers from displacement in industries vulnerable to AI-driven automation

John F. Kennedy once said, "The rights of every man are diminished when the rights of one man are threatened." His words are more relevant in a hyper-globalized world than ever before. Technology not only evolves and spreads at an exponential rate due to interconnected global supply chains. An AI system used to oppress people in one country will inevitably be exported to others. In order to ensure AI works in humanity’s best interests, governments need to apply and expand the UDHR framework to meet emerging challenges presented by new technologies.
Works Cited

Angwin, Julia, et al. “Machine Bias.” ProPublica, 23 May 2016,

Forbes, Forbes Magazine, 19 June 2019,

Browne, Grace. “The Problem with Mental Health Bots.” WIRED UK, 1 Oct. 2022,
https://www.wired.co.uk/article/mental-health-chatbots.


