May 21, 2018

Rep. Randall Hultgren
Co-Chair
Lantos Human Rights Commission
4150 O'Neill Office Building
200 C Street, SW
Washington, DC 20024

Rep. James McGovern
Co-Chair
Lantos Human Rights Commission
4150 O'Neill Office Building
200 C Street, SW
Washington, DC 20024

RE: **MAY 22 HEARING ON ARTIFICIAL INTELLIGENCE: THE CONSEQUENCES FOR HUMAN RIGHTS**

Dear Chairman Hultgren, Chairman McGovern, and Members of the Commission:

On behalf of Amnesty International[1] and our more than seven million members and supporters worldwide, we hereby submit this statement for the record. Amnesty International is an international human rights organization with major offices around the world, including the U.S. and the U.K.

Amnesty's Artificial Intelligence ("AI") and Human Rights Initiative

Amnesty's AI and Human Rights Initiative tackles human rights challenges posed by AI technologies. A core part of the initiative is the development of ethical principles for the development and use of AI. Amnesty International urges policymakers to enshrine such principles into existing human rights standards. Through our large network of human rights defenders and partner organizations worldwide, Amnesty International aims to facilitate dialogue with diverse global civil society voice about the ethics of AI, in order to ensure that the development of ethical and human rights principles for AI is guided by global human rights perspectives.

Building on our campaigning against the development of 'killer robots',

[1] Amnesty International was awarded the Nobel Peace Prize in 1977.

Amnesty International is also tackling the current use of AI in other fields including the use of AI in policing. This research will inform our work on human rights principles for AI. Amnesty International is also exploring ways in which AI can help solve global human rights challenges.

OVERVIEW

1. AI has the potential to bring positive change for human rights. AI technology could, for example, help widen access to advanced healthcare diagnostics and treatments; automation in industry could take people out of dangerous and degrading work. As Amnesty International Secretary General Salil Shetty stated at the 2017 AI for Good Summit, we have an incredible opportunity to use this technology for good.[2] However, Amnesty believes there are some key issues with AI systems that urgently need to be addressed to respect human rights now and protect rights in future.

2. Global innovation and development of AI technology is being substantially driven by major US-based companies in Silicon Valley. The US is one of a handful of countries expected to experience significant economic gain linked to AI.[3] The US Government therefore has a critical role to play in adopting appropriate policy measures and accountability mechanisms to ensure that AI advances rather than undermines human rights. In addition, US companies must meet their existing responsibility to respect human rights as set out in international human rights standards, when developing and implementing this technology.

3. For the purpose of this paper, Amnesty defines artificial intelligence as advanced computer software and computer-powered hardware that can undertake self-learning computational or physical tasks.

4. Amnesty's chief concerns with current AI systems are that:
   - AI technology is predicted to fuel massive changes to employment globally, particularly through automation of jobs, which will require governmental action to protect workers' rights.

[2] https://www.amnesty.org/en/latest/news/2017/06/artificial-intelligence-for-good/

[3] PwC estimates that China and North America stand to see the biggest economic gains with AI, with 70% of the global economic impact. http://preview.thenewsmarket.com/Previews/PWC/DocumentAssets/476830.pdf

- AI systems collecting and processing vast amounts of personal data create new threats to rights, notably to personal privacy rights on both an individual and group level.
- A growing body of research demonstrates that AI systems are already contributing to discrimination – for example, in policing and criminal justice systems in the US. *The Toronto Declaration* underscores the risks to the right to equality and non-discrimination that are inherent to machine learning, and outlines means of protecting and promoting this right.[4]
- The impact of AI on policing and conflict could have extremely dangerous and irreversible implications on international human rights and humanitarian law.
- A lack of transparency and accountability in current systems denies those harmed by AI-informed decisions adequate visibility of harms and access to effective remedy.
- Innovation in AI technology is being led by powerful corporate actors and has rapidly advanced before appropriate state-based regulatory safeguards have been put in place.

Summary of recommendations

5. Amnesty International recommends that the US government:
   - Considers and acts to protect workers' rights and the right to work where AI technology is predicted to heavily impact employment practices, ensuring a gendered perspective.
   - Ensures that the rights of individuals, including privacy rights, are better protected through stronger data protection laws.
   - Introduces regulation to ensure that AI systems are audited effectively and system developers and users are held accountable, with clear processes of responsibility outlined prior to build and deployment.
   - Supports an international pre-emptive ban on the development, transfer, deployment and use of autonomous weapons systems.

---

[4] Amnesty International and Access Now led the drafting of *The Toronto Declaration on promoting the right to equality and non-discrimination in machine learning systems*, launched 17 May 2018. To date, over ten rights organisations have endorsed the Declaration, including Human Rights Watch and Wikimedia Foundation. Amnesty ultimately hopes that private sector actors and states will endorse the Declaration and acknowledge their existing commitments to the right to equality and non-discrimination. https://www.amnesty.org/en/documents/pol30/8447/2018/en/

- Educates and informs citizens of their rights concerning privacy and data, including in automated decision-making.
- Invests in AI developments in the public sphere to foster AI technology and solutions for the public interest.

6. Amnesty International recommends that US-based companies:
   - Follow a human rights due diligence framework in order to ensure they have taken appropriate measures to avoid causing or contributing to human rights abuses through the use of AI systems.[5]
   - Take practical measures to promote AI systems that favour equity.

IMPACT OF AI ON EMPLOYMENT AND WORKERS' RIGHTS

7. Advanced AI systems will likely increase automation in the workplace. Technological advances and 'efficiency' savings will likely see machines replacing functions previously performed by humans in the workplace, as processes become part or fully automated.

8. The US government needs to approach the impact of technology on workers' rights from a gendered perspective. As more companies try to enforce a lower pay regime and weaker conditions of employment, women are highly likely to be disproportionately affected. The gig economy, if not properly regulated, risks lacking adequate protection for workers' rights thereby facilitating exploitation. At the same time, the expansion of automation is predicted to result in massive job losses, especially in the short-term, and especially at the expense of low-skilled positions, thereby risking further entrenching the social and economic marginalization of women.[6]

9. Authorities must act to regulate the gig economy in order to protect human rights. The growing spread of new forms of casual, on-demand work can prove beneficial, by allowing women to have more

[5] The responsibility of companies to respect human rights and carry out human rights due diligence is set out in the UN Guiding Principles on Business and Human Rights http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf

[6] World Economic Forum, Towards A Reskilling Revolution, January 2018, p 13

flexibility with respect to their work life, whilst supplementing their income. However, when left unregulated this fragmentation and increased fluidity of the labour market can also pose serious risks for the socio-economic rights of women, as their protections are reduced and job and income security, discrimination, and exploitation worsen, thereby further entrenching unequal power relations in the work-place, in the family, and in society.

10. The US government needs to ensure that people can access their employment rights now and in the future, including:
    - Invest in training and reskilling programmes to help those whose jobs could be at risk of automation to stay employable, considering new skills that will be in demand in a tech-driven economy.[7]
    - Enable women to access decent work in the gig economy by implementing best practices such as parental leave, affordable and accessible care services (child, elder, disability); flexible working time arrangements (while respecting working time regulations); social security; basic infrastructure; discrimination protections; equal pay; safe working conditions and pension (particularly in the informal sector).
    - Prepare for an employment landscape that is radically altered by mass unemployment and fully considering the impact on state welfare and benefits systems. This may include exploring the viability and desirability of alternative income models like Universal Basic Income.[8]

PERSONAL DATA – PRIVACY AND PROFILING RISKS

11. Advancements in AI come hand-in-hand with the development of vast economies of personal data – raising concerns about privacy rights. AI systems are developed and trained using extremely large datasets. They are by and large designed to hone their function through continually processing new data – the larger quantities of

[7] The UK Parliament's House of Lords Select Committee on Artificial Intelligence recommended a significant government investment in skills and training to navigate the disruption in the jobs market. See Report of Session 2017–19, April 2018: https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf

[8] For more on the human rights case for exploring Universal Basic Income, see report by Philip Alston, UN Special Rapporteur on Extreme Poverty and Human Rights, delivered to the UN Human Rights Council in June 2017: https://documents-dds-ny.un.org/doc/UNDOC/GEN/G17/073/27/PDF/G1707327.pdf

relevant data that the system can access, the better. (For example, AI software in healthcare diagnostics will in theory perform better over time through collecting and processing live data from a wide source of patients to create more accurate diagnoses).

12. The right to privacy is hugely significant and yet widely abused by states through government mass surveillance programmes. Many governments, including the USA, have ultimately taken advantage of advances in technology to access and store private information on an unprecedented scale. The proliferation of AI systems creates the possibility for system owners to collect detailed and intimate personal information an individual level.

13. There are numerous risks associated with networked systems storing and processing such large amounts of personal data:
    - Use of advanced AI software will dramatically increase the points of personal data collection in terms of both volume and detail. For example, facial-recognition and gait recognition technologies can easily capture and process detailed personal information on a previously unforeseen scale.
    - The networking of interconnected systems – from the internet and telecoms, to systems and sensors in travel, health, logistics, traffic, electricity networks – allows the possibility for cross-referencing data that, if collected previously, used to be held in silos. Networked big data may be used to create intimate and precise personal profiles of individuals, a tactic already widely used for commercial advertising and political marketing during elections.[9] AI software makes profiling on such an intimate individual level much more accessible – with the potential for companies and governments to influence people to a greater degree than ever before, using highly personalised messaging across a range of platforms.
    - Personal data is increasingly being used by systems to inform decision-making processes in all areas of our lives. There is potential for discrimination where information from one aspect of someone's life or previous behaviour is used to inform a decision or access to a service elsewhere. For example,

[9] http://www.bbc.co.uk/news/uk-39171324

insurance providers may use social media data to evaluate an insurance claim without the claimant's knowledge.[10]

AI SYSTEMS MAY PERPETUATE OR FACILITATE DISCRIMINATION

14. The adoption of AI and data-driven processes to aid governance and decision-making across many sectors of society has the potential to facilitate discrimination if proper oversights are not put in place. Working with a group of human rights and machine learning experts, Amnesty International and Access Now have launched *The Toronto Declaration,* which sets out the existing human rights obligations of states and responsibilities of private sector actors to protect the right to equality and non-discrimination in the context of machine learning, and outlines means of protecting these rights. The Declaration also highlights the need for systems (specifically machine learning systems, though the principles apply for related technology) to be visible, to allow individuals or groups means to challenge outcomes. Furthermore, the Declaration outlines existing obligations to ensure individuals and groups of people have access to effective remedy – a challenge for the current state and application of AI systems.

15. *The Toronto Declaration* was in part drafted in response to the serious problem with unconscious bias caused by the lack of diversity in the design of AI systems, which both states and private sector actors must address. The artificial intelligence and wider tech industry has seen a largely homogenous community power the creation and fostering of technology. The expertise and money for developing these systems is concentrated in a small pool of regions (US, North Europe, China). Systems are largely designed and deployed by a group of people with limited diversity in terms of race, culture, gender, caste, and socio-economic backgrounds.

16. As automated systems advance rapidly and are deployed across spheres with a high impact on human rights, there is an urgent need to put safeguards in place to mitigate the risks and guarantee accountability when abuses do occur. Scrutiny of such systems and

[10] Car insurance company Admiral last year attempted to use Facebook data to glean information that would inform insurance decisions: https://www.theverge.com/2016/11/2/13496316/facebook-blocks-car-insurer-from-using-user-data-to-set-insurance-rate

how they work as 'decision support' tools is difficult, given that these systems are usually proprietary. States must create means to regulate AI systems, particularly where they are used in public services, in order to ensure that rights are protected and people have access to effective remedy where rights are harmed.

17. The US Immigration & Customs Enforcement agency's proposed Extreme Vetting Initiative is a case in point.[11] The initiative sought to use automated decision-making, machine learning, and social media monitoring to assist in vetting of visa applicants and to generate leads for deportation. As set out, the program would have been both ineffective and discriminatory, proposing to evaluate whether an individual will become "a positively contributing member of society" or whether he or she "intends to commit criminal or terrorist attacks".[12] In a letter to the US government, 54 leading experts in machine learning and automated decision-making stated that "no computational methods can provide reliable or objective assessments of the traits that ICE seeks to measure" and that the proposed system would likely be inaccurate and biased.[13]

18. Another example is a highly-cited ProPublica investigation that found an algorithm used in the criminal justice systems of several US states to calculate a 'risk score' for prison inmates' likelihood of reoffending to be highly discriminatory.[14]

---

[11] In July 2017, ICE held an industry day in which it sought input from the private sector about an "overarching vetting contract that automates, centralizes and streamlines the current manual vetting process effort." ICE has since reportedly abandoned the proposal : https://www.washingtonpost.com/news/the-switch/wp/2018/05/17/ice-just-abandoned-its-dream-of-extreme-vetting-software-that-could-predict-whether-a-foreign-visitor-would-become-a-terrorist/?noredirect=on&utm_term=.6c56e8c72620

[12] Open letter to US Department of Homeland Security signed by 56 non-governmental organisations, November 2017: https://www.brennancenter.org/sites/default/files/Coalition%20Letter%20to%20DHS%20Opposing%20the%20Extreme%20Vetting%20Initiative%20-%2011.15.17.pdf

[13] Letter to US Department of Homeland Security signed by 54 computer scientists, engineers, mathematicians, and other experts in the use of automated decision-making, November 2016: https://www.brennancenter.org/sites/default/files/Technology%20Experts%20Letter%20to%20DHS%20Opposing%20the%20Extreme%20Vetting%20Initiative%20-%2011.15.17.pdf

[14] ProPublica, *Machine Bias*, May 2016

19. Predictive policing tools also carry a high risk of perpetuating discrimination. One research study from the Human Rights Data and Analysis Group (HRDAG) developed a replica of a predictive policing algorithmic programme that is used by police forces in numerous US states, and ran it as a simulation on crime data in Oakland.[15] They concluded that the programme reinforced existing racial discrimination within the police. This was because the system was built using already biased data that recorded higher crime rates in parts of the city with a higher concentration of black residents. The algorithm therefore predicted more crime in those areas, dispatching more frontline police officers, who subsequently made more arrests. The new data was fed back into the algorithm, reinforcing its decision-making process and creating a pernicious feedback loop that would contribute to over-policing of black neighbourhoods in Oakland.

20. Amnesty International has carried out research into the "Gangs Matrix" Database by the Metropolitan Police Service in London, UK, which uses an automated system to assign risk scores to individuals suspected of being 'gang members'.[16] The Matrix itself and the process for adding individuals to it, assigning 'risk scores' and sharing data with partner agencies appears to be ill-defined with few, if any, safeguards and little oversight. As a result, the matrix has taken on the form of digital profiling: 78% of individuals on the database are black, a number which is disproportionate both to the black population and the percentage of black people responsible for serious youth violence in London. In this context, the introduction of automated risk-scoring on top of an already deeply flawed data collection policy with no effective oversight and safeguards in place raises significant human rights concerns.

AUTONOMOUS WEAPONS SYSTEMS

21. Developments in AI over the last decade mean that it will be possible to develop and deploy fully autonomous weapons systems (AWS) which, once activated, can select, attack, kill and wound

---

[15] HRDAG, *To predict and serve?*, October 2016

[16] Amnesty International, *Trapped in the Matrix: Secrecy, stigma, and bias in the Met's Gangs Database*, May 2018

human targets, without effective and meaningful human control. Amnesty believes that these developments pose a very serious threat to human rights in the field of conflict and policing, and calls for an international pre-emptive ban on the development, transfer, deployment and use of autonomous weapons systems.

22. The use of AWS in law enforcement operations would be fundamentally incompatible with international human rights law, and would lead to unlawful killings, injuries and other violations of human rights. Effective policing is much more than just using force; it requires the uniquely human skills of empathy and negotiation, and an ability to assess and respond to often dynamic and unpredictable situations, which AWS would be incapable of. Decisions by law enforcement officers to use minimum force in specific situations require direct human judgement about the nature of the threat and meaningful control over any weapon.

23. Similarly, the use of lethal AWS would be incompatible with the three pillars of international humanitarian law; namely distinction, proportionality and taking reasonable precautions. AWS would lack the ability to analyse the intentions behind people's actions, or make complex decisions about the proportionality or necessity of an attack.

24. China, Israel, Russia, South Korea, the UK, and the USA, are among several states currently developing systems to give machines greater autonomy in combat. The history of weapons development suggests it is only a matter of time before this could spark another hi-tech arms race. This would cause these systems to proliferate widely, and end up in the arsenals of unscrupulous governments and eventually in the hands of non-state actors, including armed opposition groups and criminal gangs.

25. AWS also raises important issues related to transparency and accountability for human rights violations and individual criminal responsibility. Use of AWS would pose serious challenges to bringing accountability for crimes under international law. Under international human rights law, states have an obligation to investigate allegations of human rights violations and bring the perpetrators to justice as part of the right to an effective remedy – a

right which is applicable at all times.

26. In the case of lethal and less-lethal AWS, it is not possible to bring a machine to justice and no criminal sanctions could be levelled against it. However, actors involved in the programming, manufacture and deployment of AWS, as well as superior officers and political leaders, should be accountable for how AWS are used. But the nature of AWS is such that it would be impossible foresee or programme how an AWS will react in every given circumstance, given the countless situations it may face.

27. Furthermore, without effective human oversight, superior officers would not be in a position to prevent an AWS from committing unlawful acts, nor would they be able to reprimand it for misconduct. AWS, are by their very nature, autonomous agents that have no individual accountability. Deploying them in combat or for the use of force in civilian environments would be a perilous step for humanity, taking away one of the strongest deterrents against the unlawful use of violence.

TRANSPARENCY AND ACCOUNTABILITY

28. The inability to scrutinise the workings of all current deep learning systems (the 'black box phenomenon') creates a huge problem with trusting algorithmically-generated decisions.[17] Where AI systems deny someone their rights, understanding the steps taken to deliver that decision is crucial to deliver remedy and justice.

29. Provisions for accountability need to be considered before AI systems become widespread – practically, this may occur at multiple points, including in developing software, using training data responsibly, executing decisions. To what extent will any automated decision be able to be 'overridden', and by whom?

30. Restricting the use of deep learning systems in some cases may be required, where such systems make decisions that directly

[17] See for example, Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, 2015

impact individual rights. The US government should encourage the development of explainable AI systems, which would be more transparent and allow for effective remedies.[18] For example, a draft bill before New York City council advocates for transparency for all systems where algorithms are generating decisions in government services.[19]

38. Systems need transparency, good governance (including scrutiny of systems and data for potential bias), and accountability measures in place before they are rolled out into public use – especially where AI systems play a decisive and influential role in public services (policing, social care, welfare, state healthcare). It is vital that AI systems are not rolled out in areas of public life where they could discriminate or generate otherwise unfair decisions without the ability for interrogation and accountability.

39. There are also widely-applicable opportunities offered by AI systems in supply chain management, supported by blockchain technology for product identification, including provenance tracking and secure transfer of custody to provide transparency and accountability from product source to distribution. These include ensuring the tracking and movement of conflict-free goods and minerals.

40. Where there are potential adverse consequences for human rights, there must be higher transparency standards applied, with obligations both on the developers of the AI and the institutions using the AI system. This includes:
    - Detecting for and correcting for bias in design of the AI and in the data used.
    - Effective mechanisms to guarantee transparency and accountability in use, including regular audits to check for discriminatory decisions and access to remedy when individuals are harmed.

---

[18] An expert group of AI researchers has recommended that core public agencies, such as those responsible for criminal justice, healthcare, welfare, and education (e.g "high stakes" domains) should no longer use "black box" AI and algorithmic systems. See AI Now Institute, AI Now 2017 Report

[19] https://www.nytimes.com/2017/08/24/nyregion/showing-the-algorithms-behind-new-york-city-services.html

- Not using AI where there is a risk of harm and no effective means of accountability.

CORPORATE ACTORS

31. Government and civil society have struggled to keep up with the myriad of challenges to privacy and freedom of expression posed by developments in internet technologies: laws and public policies are still catching up with technologies that have been in wide use for years, if not decades. At the same time, there is a tension for policy-makers between the imperative to get to grips with and regulate the development and use of AI systems, and the appeal of these systems – which promise to 'modernize' and 'increase efficiency' across the public sector, while reducing cost. The overwhelming majority of AI systems are developed by private technology companies – systems which governments then may purchase to use in public services. As the uses of powerful AI technologies start to permeate all aspects of life, it is crucial that civil society and governments do not lag behind in responding to AI developments as they did with the development of the internet.

32. Amnesty is concerned that proprietary AI systems built by private actors will be in widespread use, including across the public sector, before human rights risks have been fully considered and addressed, and appropriate regulatory safeguards put in place. This presents a major barrier to ensuring transparency and accountability of such systems. Corporate actors themselves have a responsibility to respect human rights that exists independently from state's obligations. States need to ensure that positive developments in AI technologies, for example in healthcare, are not restricted by intellectual property practices.

CONCLUSION

33. To ensure personal data collection and use by AI systems does not impact negatively on the rights of people in the USA and around the world, the government must:
    - Create and uphold adequate regulation of private companies, including, for example, by mandating independent audits of AI systems where their use cases mean they can potentially have a significant impact on human rights.
    - Give greater powers to regulatory bodies that provide oversight

and accountability on the use of AI and big data, particularly where AI systems could adversely affect rights.

- Ensure that the rights of individuals, including privacy rights, are strengthened and upheld through stronger data protection laws, similar to the EU's General Data Protection Regulation (GDPR).
- Advocate for a pre-emptive international ban on the development, transfer, deployment and use of Autonomous Weapons Systems.[20]
- Ensure that AI systems in public service use are designed in a manner compatible with human rights standards, such as being non-discriminatory and providing means to pursue effective remedy.
- Invest in AI development in the public sphere to ensure development of AI technology and solutions for the public interest, and that it does not solely follow the commercial interests of private companies.
- Educate and inform citizens of their rights concerning privacy and data, including in automated decision-making.
- Restrict the use of AI systems that can't be interrogated in use cases where those systems make automated decisions that affect an individual or a groups' enjoyment of their human rights.

34. Companies and other private sector actors that develop and deploy AI systems and applications should:

- Follow a human rights due diligence framework to ensure they have taken appropriate measures to avoid causing or contributing to human rights abuses through the use of AI systems.
- Take practical measures to promote systems that favour equity, by investing in programmes that promote diversity of staff at development and deployment stage, and ensure that marginalised groups and individuals are not adversely affected by intentional or inadvertent discrimination.[21]

---

[20] Amnesty International urges the US government to engage in a comprehensive debate around the multiple challenges posed by AWS in order to develop and articulate a national policy on AWS (including less-lethal AWS) that takes full account of the state's obligations to respect and ensure international human rights law and international humanitarian law. This must be done in consultation with a broad range of stakeholders, including by meaningful and substantive engagement with non-governmental organizations and relevant experts, including AI and robotics experts and industry leaders.

[21] See *The Toronto Declaration* for suggested means of promoting equity and preventing discrimination in

For more information, please contact Joanne Lin, National Director for Advocacy and Government Affairs, at 202/509-8151 or jlin@aiusa.org.

Sincerely,

Joanne Lin
National Director
Advocacy and Government Affairs
Amnesty International USA

Anna Bacciarelli
Researcher/Advisor
Technology and Human Rights
Amnesty International

Joe Westby
Researcher/Advisor
Technology and Human Rights
Amnesty International

machine learning systems.