MINES ACTION CANADA
PO Box 4668 Postal Station E Ottawa ON K1S 5H8 CANADA
tél: +1.613.241.3777 fax: +1.613.244.3410 info@minesactioncanada.org www.minesactioncanada.org

Forgotten Communities:

Autonomous Weapons Systems and Bias in Artificial Intelligence

Background:

Machine learning, a type of artificial intelligence, uses datasets to create patterns to understand how the world works. Based on these patterns it will make predictions about how things will act in the future and make decisions based on those predictions.[i] When the datasets used to teach algorithms are not representative of differences in gender, race, (dis)ability, or socio-economic background (i.e. do not take an intersectional approach[1]), a biased algorithm is created. A lack of diversity among programming teams often means that biases go unchecked and are incorporated into the algorithm. Should such problems go unchecked, these problematic biases can increase the margin of error for the accurate identification of individuals in underrepresented groups.

Bias in current technology:

When autonomous weapons systems (AWS) are programed with these biased algorithms, the embedded stereotypes have much more severe consequences. To assess the potential risk posed by biased algorithms in AWS we have to look to current technology. A study at MIT (by Joy Buolamwini and Timnit Gebru) found that 34.7% of dark skinned women were misidentified as men in various face recognition algorithms.[ii] "Facial recognition trials held in England by the London Metropolitan Police have consistently yielded a 98 per cent failure rate."[iii] High error rates of facial recognition technology used by some police forces have raised concerns about the efficacy of using such technology for identifying or targeting suspects in police or military operations.

At times, because of the quality of the video feed, drone operators could mistake a person holding a walking stick for a rifle.[iv] The negative consequences of current image analysis and facial recognition technology should not be ignored, they are the harbingers of future problems. These concerns should be extended to any military use of these technologies, such as with AWS, where the risks are even greater. AWS could potentially have these same difficulties in identifying targets as required by International Humanitarian Law (IHL).

Missing components:

Research on risks of weaponized artificial intelligence has only just begun to consider the implications of bias in artificial intelligence (AI). While gender, race, and socio-economic background are beginning to

[1] In this paper an intersectional approach requires recognizing that individuals experience oppression in varying configurations and in varying degrees of intensity based on interlocking systems of power include race, gender, class, ability, and ethnicity.

be acknowledged, disability is starkly absent. Persons with disabilities make up 15% of the global population and yet the conversation about autonomous weapons has overlooked this diverse community.

In AWS artificial intelligence would use the patterns it creates to help differentiate between combatants and civilians. However, "[t]he data sets assume we are homogeneous doppelgängers of each other, without addressing the micro-variations of people's faces."[v] Also not included would be the experiences of persons with disabilities. By neglecting to include those with disabilities, or those whose appearance or behaviour do not conform to what the programmer identifies as "normal", they are at a higher risk of being misidentified as a combatant and illegally targeted.

There is no way to predict how an autonomous weapon will react if it encounters someone or something that is missing from the data. It is unlikely AWS could appropriately apply IHL principles of distinction and proportionality to a context it does not recognize. It is possible that AWS could assume anything it does not recognize is a hostile agent contrary to IHL. The inability of an AWS to account for the diversity of humanity including persons with disabilities is another reason this technology should be pre-emptively banned.

Recommendations:

High Contracting Parties to the Convention on Certain Conventional Weapons should

- Work towards a legally binding instrument that prohibits fully autonomous weapons.
- Adapt national policy and laws to prevent the development of fully autonomous weapons.
- Work in coordination with civil society and other national stakeholders to support the negotiation of a new treaty. Engage diverse stakeholders.
- Ensure a gender sensitive and inclusive approach is used when creating new international norms and laws concerning lethal autonomous weapons systems to ensure consideration of disability and other facets of human diversity.

---

[i] Spark with Norah Young, "AI's problem with disability and diversity", CBC Radio (2017) <https://www.cbc.ca/radio/spark/362-machine-learning-outliers-smart-device-ownership-and-more-1.4279433/ai-s-problem-with-disability-and-diversity-1.4279444>

[ii] Fussel, Sidney. "Can We Make Non-Racist Face Recognition?", Gizmodo (2018) <*https://gizmodo.com/can-we-make-non-racist-face-recognition-1827639249*>

[iii] Carter, William Michael. "Big Brother facial recognition needs ethical regulations", The Conversation (2018) <*https://theconversation.com/big-brother-facial-recognition-needs-ethical-regulations-99983*>

[iv] Press, Eyal. "The Wounds of the Drone Warrior", New York Times Magazine (2018) <https://www.nytimes.com/2018/06/13/magazine/veterans-ptsd-drone-warrior-wounds.html>

[v] Carter, William Michael. "Big Brother facial recognition needs ethical regulations" <*https://theconversation.com/big-brother-facial-recognition-needs-ethical-regulations-99983*>