August 2023

Supporting Responsible AI- Submission by SafeGround

As AI is used in new and varied ways, it is essential to understand the risks and mitigate these to avoid harms. As the brief for these consultations states, “Australians need to be able to trust that AI will be used ethically, safely and responsibly.” In order to support the safe and responsible use of AI, multiple dimensions must be addressed through the development of both tools, frameworks and principles, and regulatory measures. This submission will focus on digital dehumanisation and the weaponisation of AI, and give recommendations for government actions in these areas.

Digital dehumanisation:

The risks and potential harms

Digital dehumanisation is the process whereby humans are reduced to data, which is then used to make decisions and/or take actions that negatively affect their lives. The harms of automated decision-making technology have been highlighted recently in Australia through the RobotDebt scheme. AI-enabled decisions are emerging in a variety of domains in the civil sphere, from social welfare, banking and loans, and the justice system. The process in automated decision-making deprives people of dignity, demeans individuals’ humanity, and removes or replaces human involvement or responsibility through the use of these technology. Marginalisation of groups is also exacerbated as factors like race, ethnicity, gender, age, socio-economic status and so forth may result in automated decisions that further entrench inequalities experienced by people from these groups, when such systems fail to evaluate wider context and instead reflect metrics that reduce people to these dehumanising data points.

Suggested measures

It is imperative that the government understands these harms, and regulates the use of automated-decision-making technologies within its own departments and the private sector. The Australian government can implement comprehensive strategies that foster responsible AI advancements and prevent digital dehumanisation.

Conducting risk assessments across public and private sectors can focus on identifying the merits, concerns, and potential risks tied to any applications employing automated decision-making technologies. Evaluations should incorporate risk prevention, mitigation and monitoring measures.

As applications of AI are being increasingly used, there is a pressing need to identify groups that are most susceptible to digital dehumanisation. This should include gauging the socioeconomic impacts of AI systems and guaranteeing that their deployment does not create a digital divide between those who benefit from these applications and those who are at the forefront of unintended consequences or even victims to malicious uses of AI.

AI systems and applications identified as bearing potential risk of digital dehumanisation should be rigorously tested and evaluated by AI, ethics, human rights and policy experts in cross-disciplinary processes that consider the technologies limitations and potential societal and human harms before implementing them or introducing them into the market. Standards or frameworks/guidelines could

be developed to ensure this is comprehensive and consistent. The Government and private companies should collaborate to devise protocols for due diligence and oversight. These will identify, prevent, mitigate and account for impacts of AI systems, helping to ensure responsible and safe AI. Criteria and standards should also be applied to grant applications and contract awarding processes. Applications should have to explicitly articulate how AI will be harnessed to prevent harm and mitigate risks.

Public awareness and comprehension of AI should also be championed. This can be achieved through accessible educational resources, civic participation avenues, proficiency in digital skills, instruction on AI ethics, and media literacy. Civil society can play a leading role in training. Given that artificial intelligence's influence basically transcends age barriers, AI courses and training should cater for various age groups. Courses and education should enhance understanding of risks and impacts, and empower individuals to recognise and take action against digital dehumanisation. Educational frameworks need to account for linguistic, social, societal and cultural diversity. Providing financial support to community-led programs aimed at educating diverse Australian communities about digital dehumanisation will facilitate an environment where AI is harnessed for good. Such public awareness and education is essential for addressing a whole range of other issues, concerns and risks relating to responsible AI beyond digital dehumanisation.

Weaponisation of AI:

Global and Australian Context

AI, among its many applications, has become a critical technology in the defence domain. AI can be used in many innovative ways by the military, including assisting minefield clearance, delivery of supplies or aid. Many governments are beginning to pursue autonomy in weapons. The development landscape in Australia is extensive, including the Trusted Autonomous Systems Defence Cooperative Research centre fostering projects and collaboration between Defence, private companies and industry, partnerships between defence and large arms developers, and an array of Australian private companies. Cooperation on AI is also a key aspect within the advanced capabilities sharing as part of the AUKUS security alliance. The newly announced Strategic Advanced Strategic Capabilities Accelerator also includes AI as a priority area. It is imperative the Australian government ensures all development is both lawful and ethical.

To ensure the responsible use of AI, it is essential there are clear limits on autonomy and that certain autonomous weapons be ruled out. Over 90 states have called for a new legally binding instrument to address the risks of autonomous weapons, as well as the United Nations Secretary-General, the International Committee of the Red Cross, AI experts in academia and industry and civil society.

Suggested regulation

The Australian government should support negotiations for a new legally binding instrument and work with other states to establish prohibitions and obligations to address the legal, ethical and security risks posed by increasing autonomy in weapons including:

- A prohibition on weapons that apply force to humans based on sensory data
- A prohibition on weapons that are insufficiently predictable
- A general obligation to maintain meaningful human control over the use of force

- Regulation of the development and use of other autonomous weapons, including through a combination of:
    - limits on the types of target, such as constraining them to objects that are military objectives by nature
    - limits on the duration, geographical scope and scale of use, including to enable human judgment and control in relation to a specific attack
    - limits on situations of use, such as constraining them to situations where civilians or civilian objects are not present
    - Specific requirements for human–machine interaction, notably to ensure effective human supervision, and timely intervention and deactivation.

Australia should also adopt clear policy that rules out the development and use of those autonomous weapons to be prohibited. This is guided by the general principle to be recognised by the government, that weapons must not apply force to humans based on the processing of sensory data. This is imperative for the responsible development of AI, its use in military applications and in the civil domain - especially because of the dual-use nature of AI technology.

Further, the government should establish a commitment to ensure meaningful human control over all weapons including over the critical functions of selecting targets and the decision to attack. This should apply to weapons developed by the Department of Defence, ADF and related government branches as well as private companies in Australia. This could be incorporated in the existing 'System of Control', utilised by the Department of Defence as the framework for weapons development, or in a standalone policy document to address AI and autonomy in weapons.

Public and civil society views on autonomous weapons should be taken into account to ensure trust from the public in how AI is used and regulated. As such, the government should also support civil society participation in international meetings related to global efforts to address autonomous weapons and invite and facilitate civil society onto their delegations at international meetings. This will enhance transparency of the government's actions with regard to the regulation of autonomous weapons, the most acute end of the spectrum of digital dehumanisation.