Tech Global Institute
8 Brunswick Street
Brampton, ON L6X 4Y6
www.techglobalinstitute.com

July 25, 2023

Submission to the Department of Industry, Science and Resources of the Government of Australia on the discussion paper on Safe and Responsible AI in Australia

Tech Global Institute welcomes the opportunity to provide comments on the important topic of formulating comprehensive actions on the regulation and governance of artificial intelligence (AI) in Australia. Tech Global Institute is a global policy lab with a mission to reduce equity and accountability gaps between internet technologies and the Global Majority. We are a community of senior policy and legal experts, trust and safety professionals, AI systems researchers and human rights specialists who collectively have decades of experience in building, scaling and governing the products at leading technology companies, academia and multilateral organisations. We focus on elevating the voice of underserved communities around the world in the design, development, deployment and governance of technologies that impact their lives.

With the emergence of new branches of large-scale consumer technologies, governments around the world, including Australia, are at a crucial crossroads to develop proportionate and effective guardrails that promote innovation while safeguarding democracy, privacy and human rights. Often these conversations do not adequately account for the *lived* experiences of historically marginalised groups, including immigrant and indigenous communities, that interact with technologies. Specifically, the conversations are limited to topics on access and inclusion, although evidence on the use of social media and other products from the past decade has shown people everywhere are using them for diverse reasons, such as e-commerce, healthcare, news and agriculture. As such, because of historically unaddressed biases and power dynamics, harms from these same technologies have disproportionately fallen on underrepresented communities. Our work specifically focuses on developing policy evidence on the impact of technologies on underrepresented groups and advising governments on how best to address gaps. Through updating its regulations and policies to address AI risks, the Australian Government has the unique advantage of leading the world in the current and next phases of AI development, ensuring risks are mitigated early on and all parts of society can equitably reap its benefits.

In our submission below, we address critical considerations in how AI impacts the lives of historically marginalised and underserved groups, the appropriate regulatory frameworks to address differential power dynamics in society and the role of the Australian Government to develop and set precedent on responsible, equitable and inclusive AI systems. We welcome this opportunity to share evidence and look forward to future engagements on this critical topic.

Sincerely,

Sabhanaz Rashid Diya
Founder

Shahzeb Mahmood
Associate General Counsel

Abdullah Hasan Safir
Leverhulme Centre for the Future of Intelligence
University of Cambridge

On behalf of Tech Global Institute

OUR RESPONSE TO THE CONSULTATION QUESTIONS

2. What potential risks from AI are not covered by Australia's existing regulatory approaches? Do you have suggestions for possible regulatory action to mitigate these risks?

For over two decades, internet exceptionalism[1] – i.e., the notion that internet is unique, and therefore should be regulated by its own norms and tailored legal and regulatory frameworks[2] – has dissuaded governments from proactively regulating cyberspace with specialised bodies of rules. Instead, liability-limiting shields were implemented to 'facilitate the robust development and worldwide expansion of electronic commerce, communications, research, development and education in the digital age'.[3] It was thought that existing legal constructs – including traditional contract, tort, criminal and administrative concepts – were flexible enough to counteract internet abuses.[4] However, the recognition of widespread harms resulting from the laissez-faire non-interventionist approach in recent years has led to a growing demand for regulation, particularly with the accelerated development and adoption of AI and automated decision-making (ADM) technologies worldwide.

We believe that the existing legal constructs should be retained and applied to new technologies, but as the use of existing constructs meets their limitations, regulatory instrumentalism is necessary to align new technologies, such as advanced AI and ADM, to societal values. Given the ever-evolving nature of technologies and the potential perils of far-reaching harm, legal protection under one legislation may not be adequate in and of itself. To be effective, laws will have to be proactive, and adopt a novel approach, combining data protection, consumer rights, competition, anti-discrimination and online safety legislations, within the broader architecture of a human-centric and rights-focused AI regulation. It is crucial that benefits of technocracy and efficiency must yield to core values of human rights and democracy.

---

[1] Brian Holland, 'In Defense of Online Intermediary Immunity: Facilitating Communities of Modified Exceptionalism' (2008) 56(2) *Kansas Law Review* 369, 394. See also Neil Fried, 'The Myth of Internet Exceptionalism: Bringing Section 230 into the Real World' (2021) 5(2) *American Affairs* <https://americanaffairsjournal.org/2021/05/the-myth-of-internet-exceptionalism-bringing-section-230-into-the-real-world/>.

[2] David Harvey, *Collisions in the Digital Paradigm: Law and Rule-Making in the Internet Age* (Hart Publishing, 2017) 113.

[3] YouTube Inc., 'Memorandum of Law in Support of Defendants' Motion for Summary Judgment', Submission in *Viacom International Inc., et al v YouTube, Inc., et al* and *The Football Association Premier League Limited v YouTube, Inc., et al*, Civil No. 07-CV-2103 (LLS) and Civil No. 07-CV-3582 (LLS), 11 March 2010, 16.

[4] Roger Brownsword, 'Smart Transactional Technologies, Legal Disruption and the Case for Network Contracts', in Larry A. DiMatteo, Michel Cannarsa, and Cristina Poncibò (eds.), *Cambridge Handbook on Smart Contracts, Blockchain Technology and Digital Platforms* (New York: Cambridge University Press, 2020), 313, 322.

We believe that strengthening the existing laws on competition, consumer rights and data protection, complemented by a new AI regulation, can ensure the regulatory framework is future-proof, technology-neutral, and fit-for-purpose, and can address foreseeable harms associated with new technologies. Our assessments and recommendations are elaborated below.

Competition

Big technology companies tend to dominate the market,[5] and Google and Meta – both prolific users of AI driven technology[6] – exemplify this trend. Google, for instance, held over 95% market share in general search engine services and search advertising in Australia in 2020,[7] maintaining over 93% market share in both segments for over a decade.[8] Similarly, according to independent findings of the Productivity Commission and the Australian Communication and Media Authority, Facebook accounted for at least 95% of Australian social media users,[9] with 91% representing youth consumers and 97% being elderly consumers.[10]

Both companies operate multi-sided platforms, offering free services to users while monetising user data to subsidise their operations.[11] An investigation by the Federal Cartel Office in Germany concluded in February 2019 found Meta's data policy abusive, as the company amassed vast amounts of user data from its own platform, affiliated services (such as Instagram and WhatsApp) and unaffiliated third-party sites, and then combining them to create detailed profile of the users, without obtaining effective and voluntary

[5] Australian Competition & Consumer Commission, *Digital Platform Services Inquiry* (Interim Report No. 5, September 2022) 7 ('*DPSI*') <https://www.accc.gov.au/system/files/Digital%20platform%20services%20inquiry%20-%20September%202022%20interim%20report.pdf>.
[6] Francina Cantatore and Brenda Marshall, 'Safeguarding Consumer Rights in a Technology Driven Marketplace' (2021) 42(2) *Adelaide Law Review* 467, 474.
[7] *DPSI* (n 5) 17.
[8] Australian Competition & Consumer Commission, *Digital Platform Inquiry* (Final Report, June 2019) 65 ('*DPI*') <https://www.accc.gov.au/system/files/Digital%20platforms%20inquiry%20-%20final%20report.pdf>.
[9] Productivity Commission, *Data Availability and Use* (Final Report, 31 March 2017) 573 <https://www.pc.gov.au/inquiries/completed/data-access/report/data-access.pdf>.
[10] Australian Communications and Media Authority, *Communications and media in Australia: The Digital lives of younger Australians* (Report, May 2021) 7 <https://www.acma.gov.au/sites/default/files/2021-05/The%20digital%20lives%20of%20younger%20Australians.pdf>; Australian Communications and Media Authority, *Communications and media in Australia: The Digital lives of older Australians* (Report, May 2021) 5 <https://www.acma.gov.au/sites/default/files/2021-05/The%20digital%20lives%20of%20older%20Australians.pdf>.
[11] *Facebook Inc v Australian Information Commissioner* [2022] FCAFC 9, 3 (Allsop CJ). See also Jacques Crémer, Yves-Alexandre de Montjoye and Heike Schwei, 'Competition policy for the digital era' (Final Report, 2019) 44; Maximilian N. Volmar and Katharina O. Helmdach, 'Protecting consumers and their data through competition law? Rethinking abuse of dominance in light of the Federal Cartel Office's Facebook investigation' (2018) 14(2-3) *European Competition Journal* 195, 212.

consent.[12] It enabled Meta to gain an unfair and unlawful competitive advantage to the detriment of consumers and competitors.[13] Similarly, the Australian Competition and Consumer Commission (ACCC) noted that 'while users of Facebook may expect a certain level of data to be generated through their use of the main Facebook services as a quid pro quo for their use of the service, users may not expect Facebook to be collecting data on their interactions on other seemingly unrelated sites and apps, and using that data to assist it sell ad inventory'.[14]

Data accumulation strategies and systemic privacy infringement can lead to antitrust harm.[15] Big data, characterised by its volume, velocity, variety, verifiability and value, has strengthened digital platforms' dominance in the data-driven markets, enabling exclusionary conduct that harms competition and consumers.[16] We note that the ACCC has necessary tools in its legislative armoury, including in particular section 46 of the Competition and Consumer Act 2010, to hold the companies accountable. This provision allows the regulator to sanction conduct of corporations with substantial market power if it has the purpose, effect, or likely effect of substantially lessening competition in the relevant market. Notably, the provision is open-ended, enabling privacy harms to be considered when assessing substantial lessening of competition without establishing actionable infringement.[17] Yet, since the law was amended in November 2017 to adopt a more flexible effects-based test, the ACCC exercised its new powers under this provision only once, and never against a digital platform.[18]

According to the ACCC, the existing framework is ill-suited for effectively and efficiently addressing anti-competitive harms in the dynamic markets for digital platform services, given the complexity and protracted nature of the investigations and court proceedings, and its limitations in addressing harms retrospectively.[19] Instead, the regulator proposed introducing a new sector-specific ex ante regulation,

[12] 'Bundeskartellamt prohibits Facebook from combining user data from different sources', *Bundeskartellamt* (News, 7 February 2019) <https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2019/07_02_2019_Facebook.html>.
[13] Dzhuliia Lypalo, 'Can Competition Protect Privacy? An Analysis Based on the German Facebook Case' (2021) 44(2) *World Competition* 169, 177.
[14] *DPSI* (n 5) 35.
[15] Giuseppe Colangelo and Mariateresa Maggiolino, 'Antitrust Über Alles. Whither Competition Law after Facebook?' (2019) 42(3) *World Competition* 355, 366.
[16] Roberto Augusto Castellanos Pfeiffer, 'Digital Economy, Big Data and Competition Law' (2019) 3(1) *Market and Competition Law Review* 53, 55; Erika M. Douglas, 'The New Antitrust/Data Privacy Law Interface' (2021) 130 *The Yale Law Journal Forum* 647, 659. See also Nikolas Guggenberger, 'Essential Platforms' (2021) 24(2) *Stanford Technology Law Review* 237, 244.
[17] Stephen G. Corones and Arlen Duke, Corones' Competition Law in Australia (Lawbook Co., 7th ed, 2019) 501, 511.
[18] Google, Submission to the Australian Competition and Consumer Commission, *Discussion Paper for the fifth interim report under the Digital Platforms Services Inquiry* (8 April 2022) 22.
[19] *DPSI* (n 5) 48.

similar to the ones introduced in Germany,[20] the European Union[21] and Japan.[22] Additionally, recommendations for the introduction of mandatory codes of conduct were made to regulate large digital platforms.[23] Using these ex ante regulations and codes, companies could be forced to share access to its data to enhance competition, while imposing limitation measures on data combining certain activities. We support the decision to introduce sector-specific regulatory regime, but it should be complementary and not an alternative to the existing enforcement tools, as the current ex post framework can address a range of conducts, including anti-competitive conduct arising from the entrenched market positions of the digital platforms.

Consumer Protection

While the Australian Consumer Law (ACL) under Schedule 2 of the Competition and Consumer Act 2010 provides valuable consumer protections in various areas, it was not designed to address the unique challenges and potential harms posed by AI and ADM systems, and its existing framework may be deficient in fully safeguarding against AI-related harms. In 2018, the ACCC acknowledged the significance of algorithms in optimising data usage but expressed concern about its impact on the consumer experience.[24]

Under the ACL, there are several effective tools to address consumer protection issues in the digital marketplace, including prohibition on misleading or deceptive conduct, false representations about services (including data handling practices), and unconscionable actions.[25] It also requires compliance with unfair contract term provisions for consumer-facing terms of use and privacy policies.[26] And, some of these provisions have been used to successfully target non-compliant companies. For instance, in December 2020, Meta faced a lawsuit for making false, misleading and deceptive representations about

[20] 'Bundeskartellamt determines Google's paramount significance for competition across markets' *Bundeskartellamt* (Case Summary, 5 January 2022) <https://www.bundeskartellamt.de/SharedDocs/Entscheidung/EN/Fallberichte/Missbrauchsaufsicht/2022/B7-61-21.pdf?__blob=publicationFile&v=6>.
[21] *Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act)* [2022] OJ L 265/1; *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act)* [2022] OJ L 277/1.
[22] 'Digital Platforms', *Ministry of Economy, Trade and Industry* <https://www.meti.go.jp/english/policy/mono_info_service/information_economy/digital_platforms/index.html>.
[23] *DPSI* (n 5) 111.
[24] Australian Competition and Consumer Commission, 'ACCC to Further Increase Enforcement Work' (Media Release No 145/18, 3 August 2018) <https://www.accc.gov.au/media-release/accc-to-further-increase-enforcement-work>.
[25] *Competition and Consumer Act 2010 (Cth)* sch 2 s 18, ch 2 pt 2-2, pt 3-1 div 1.
[26] Ibid ch 2 pt 2-3.

the security, privacy and secrecy of users' personal data through the use of Onavo, a virtual private network application acquired by the company in October 2013.[27] The application was used to monitor user activities and collect extensive personal data, which were then used to glean market insight into the popularity of WhatsApp that resulted in its multibillion-dollar acquisition in 2014.[28] Similarly, in August 2022, Google was fined AU$ 60 million by the federal court for making misleading representations about the collection and use of location data without the customers' informed choice.[29] Online directories and comparison websites – including Service Seeking, HealthEngine, iSelect, and Trivago – were also prosecuted for tampering with customer reviews, manipulating algorithms and engaging in abusive data practices.[30]

However, the prohibitions are limited to precluding overtly misleading, deceptive or unconscionable conducts and false representations about services, which, while suitable as a catch-all safety net,[31] are deficient in at least two respects. Firstly, it falls short in promoting fair and honest practices (e.g., in communication strategies, algorithmic transparency, human oversight). Second, and perhaps more importantly, it is retrospective in nature.

Currently, the law mostly responds to harms once they have occurred, whereas effective consumer protection in the digital ecosystem warrants a legal architecture that proactively places normative limits on conduct that are foreseeably harmful. Often, online services are designed to nudge and manipulate consumer choice, confuse consumers, foster screen addiction and social isolation, amplify harmful content and online echo chambers, unlawfully extract personal data, and target vulnerable consumers.

It is worth noting that the overarching objective of a consumer protection law is to scrutinise the full life of a transaction and safeguard consumers in the face of information asymmetry and market power imbalances. With the advancements in AI and ADM technologies, these challenges are exacerbated as corporations are further consolidating their positions in the market and presenting new threats to

[27] *Australian Competition & Consumer Commission v Facebook, Inc.* (Federal Court of Australia, NSD1339/2020) <https://www.accc.gov.au/system/files/ACCC%20v%20Facebook%20Inc%20%26%20Ors_%20Concise%20Statement_0.pdf>.
[28] Elizabeth Dwoskin, 'Facebook's willingness to copy rivals' apps seen as hurting innovation', *Washington Post* (online, 10 August 2017) <https://www.washingtonpost.com/business/economy/facebooks-willingness-to-copy-rivals-apps-seen-as-hurting-innovation/2017/08/10/ea7188ea-7df6-11e7-a669-b400c5c7e1cc_story.html>.
[29] *Australian Competition and Consumer Commission v Google Australia Pty Ltd* (Federal Court of Australia, NSD1760/2019) 65.
[30] *Australian Competition and Consumer Commission v Service Seeking Pty Ltd* [2020] FCA 1040; *Australian Competition and Consumer Commission v HealthEngine Pty Ltd* [2020] FCA 1203; *Trivago NV v Australian Competition and Consumer Commission* (2020) 384 ALR 496; *Australian Competition and Consumer Commission v iSelect Ltd* [2020] FCA 1523.
[31] Jeannie Marie Paterson and Yvette Maker, 'AI in the Home: Artificial Intelligence and Consumer Protection Law', in Ernest Lim and Phillip Morgan (eds), *The Cambridge Handbook of Private Law and Artificial Intelligence* (Cambridge University Press, 2021) 4.

consumers. For instance, the integration of self-learning algorithms in big data analysis presents companies with an unprecedented opportunity to gain comprehensive insights into customers' personal circumstances, behaviour patterns and personality traits. Leveraging this information, companies can customise their advertising strategies and fine-tune pricing and contract terms to suit each customer profile.[32] Thus, individuals may be induced to purchase goods they do not need, overspend, engage in risky financial transactions, or indulge in their weaknesses (e.g. gambling or drug addiction). Increasingly, companies are employing web beacons, pixel tags, device and browser fingerprinting, facial recognition, cross-device tracking, audio beaconing and dark pattern techniques to extract vast trove of data, irrespective of whether data subjects uses their services or not, to create their profiles and commercially target them. The distinctive attributes of the AI technologies, characterised by opacity, complexity, unpredictability and semi-autonomous behaviour, introduce novel risks and liability concerns for consumers.[33] To effectively address these evolving risks, a proactive and protective framework is essential to adequately address foreseeable risks in the rapidly evolving digital landscape.

We therefore recommend that the ACL should be amended to specifically incorporate foreseeable harms to consumers arising from AI and ADM, including by mandating algorithmic transparency, enabling correction of inaccuracies in training datasets to avoid bias and discrimination, restricting manipulative advertising and exploitative microtargeting, providing a right to explanation of automated decisions and human review of such decisions, and protecting vulnerable consumers such as children, elderly individuals and individuals with disabilities. Furthermore, the ACCC should be empowered to formulate ex ante codes, so that the emerging and evolving consumer harms arising from the use of AI can be proactively addressed without the need for protracted statutory amendments. In addition, unfair terms in standard form contracts are not sanctioned by a penalty regime, meaning that such terms may be voided without any statutory penalties.[34] We recommend amending the statute to introduce penalty regimes, in order to deter companies from incorporating unfair terms in the public-facing documents such as terms of use and privacy policy.

Privacy and Data Protection

In the age of digital transformation, users of internet services are not only consumers but also producers of data, which in turn becomes a valuable input to the production of goods and services. However,

[32] Martin Ebers and Susana Navas, 'Artificial Intelligence and Consumer Protection', *FIFTEEN EIGHTY FOUR* (Blog Post, 10 September 2020) <https://www.cambridgeblog.org/2020/09/artificial-intelligence-and-consumer-protection/>.
[33] Ibid.
[34] *Competition and Consumer Act 2010 (Cth)* sch 2 s 23(1); *Australian Securities and Investments Commission Act 2001 (Cth)* s 12BF(1).

according to the Australian Community Attitudes to Privacy Survey 2020, privacy is a major concern for 70% of Australians, with nearly 60% having experienced problems with how their personal information was handled in the past 12 months, and almost 9 in 10 want more choice and control over their personal information.[35] Another research by the Consumer Policy Research Centre concluded that Australian consumers perceive themselves to be 'uninformed, unprotected and powerless' in relation to the collection of personally identifying information.[36] It also confirms that the current legislative framework is ineffective, inadequate and incapable of delivering privacy protections.

It has become apparent that the existing consent mechanism – involving lengthy policy disclaimers and click-through agreements – has rendered notice-and-choice ineffective in protecting individuals' privacy. Conversely, the widespread use of AI has raised privacy concerns due to its ability to use automated and algorithmic systems to process vast amounts of data, re-identify individuals by converting non-identifying information into identifying information, learn, and make predictions without always having transparent and explainable processes.[37] For instance, if an energy company sells to third-parties de-identified customer demographic data, such as postcode and energy consumption, it can be used by an AI agent to combine with other datasets to make specific inferences about age, lifestyle and income ranges, as well as re-identify individuals. Such data can then be used to give valuable insights to retailers about home appliance products used by customers and allow insurance companies to re-price their insurance products based on re-identification. And all of this was done with the consent of the customer when they installed a smart metre at their house.[38] Consumer data is collected not only by technology companies that consumers directly interact with, but also by data brokers who collect and sell data to third-parties. Because of the difficulties with anticipating potential data misuse and understanding its broader societal impacts, especially with the exponential advancements in the scope and speed of AI, stratospheric rise of digital platforms and vast amounts of personal data available in the public domain, the monitoring of personal data has become ubiquitous. A survey by the European Consumer Organisation shows that the

[35] Office of the Australian Information Commissioner, *Australian Community Attitudes to Privacy Survey 2020* (Report, September 2020) < https://www.oaic.gov.au/__data/assets/pdf_file/0015/2373/australian-community-attitudes-to-privacy-survey-2020.pdf>
[35] Francina Cantatore and Brenda Marshall, 'Safeguarding Consumer Rights in a Technology Driven Marketplace' (2021) 42(2) *Adelaide Law Review* 467, 474.
[36] Consumer Policy Research Centre, 'Research: Australian consumers 'soft targets' in Big Data economy' (Report, 13 May 2018) <https://cprc.org.au/2018/05/13/research-australian-consumers-soft-targets-big-data-economy>.
[37] Karen Yeung, 'Five Fears about Mass Predictive Personalization in an Age of Surveillance Capitalism' (2018) 8(3) *International Data Privacy Law* 258, 259
[38] Jillian Carmody, Samir Shringarpure, Gerhard Van de Venter, 'AI and privacy concerns: a smart meter case study' (2021) 19(4) *Journal of Information, Communication and Ethics in Society* 492, 496.

increased uptake in AI across industries and government agencies not only magnified the use of personal data, it also increased apprehension that AI will lead to more abusive data practices.[39]

While the digital economy has generated significant benefits due to widespread adoption of AI, including consumer convenience, improved efficiencies and new employment opportunities, it has also resulted in large amounts of information about people being generated, used, disclosed and stored, and often in a manner that results in discriminatory outcomes. The data-driven, algorithmic processes associated with AI enable new, much more refined and systematic, forms of stereotyping and differentiation.[40] Often, the algorithmic outcomes are influenced by the assumptions and biases of its human programmer. AI and ADM systems can be trained, tested and deployed using inaccurate or misleading information, or historical data that is affected by prejudice, such as underrepresentation of minorities and marginalised communities in datasets. Biased decisions made by such systems can create feedback loops, reinforcing existing biases and further perpetuating unfair treatment and discrimination.

In Australia, the Privacy Act 1988 requires federal government agencies and private sector businesses to collect and manage personal information in a manner that is consistent with the 13 Australian Privacy Principles, but the general consensus is that the law is below par compared to the United States, the United Kingdom and the European Union.[41] Significantly, there is no direct right of action and statutory tort for serious invasions of privacy, with consumers' right limited to complaining to the relevant company and then to the Office of the Australian Information Commissioner.[42] As such, the ACCC found that 'existing regulatory frameworks for the collection and use of data have not held up well to the challenges of digitalisation and the practical reality of targeted advertising that rely on the monetisation of consumer data and attention'.[43] We note that the Australian Government is reviewing the legislation and issued a report in February 2023, aimed at making the law 'fit for purpose' to 'adequately protect Australians' privacy in the digital age'.[44] However, it does not sufficiently address the risks associated with AI. Broadly consistent with the findings of the report, we make the following recommendations:

---

[39] BEUC, *Artificial Intelligence: what consumers say* (Report) < https://www.beuc.eu/sites/default/files/publications/beuc-x-2020-078_artificial_intelligence_what_consumers_say_report.pdf>.
[40] Policy Department for Economic, Scientific and Quality of Life Policies, *New aspects and challenges in consumer protection* (Report, April 2020) 10 <https://www.europarl.europa.eu/RegData/etudes/STUD/2020/648790/IPOL_STU(2020)648790_EN.pdf>.
[41] Samson Yoseph Esayas and Angela Daly, 'The Proposed Australian Consumer Data Right: A European Comparison' (2018) 2(3) *European Competition and Regulatory Law Review* 187, 195, 200.
[42] Gerard Goggin et al, 'Data and Digital Rights: Recent Australian Developments' (2019) 8(1) *Internet Policy Review* 1, 6-8.
[43] *DPI* (n 8) 3.
[44] Sam Buckingham-Jones et al, '$2200 per customer: Big business faces hefty privacy reform bill', *Financial Review* (online, 16 February 2023) <https://www.afr.com/technology/2200-per-customer-big-business-faces-hefty-privacy-reform-bill-20230216-p5cl0p>.

1. The Information Commissioner should be empowered to make ex ante codes in public interest and issue emergency declarations, so that the emerging and evolving privacy concerns, and significant and imminent privacy harms, arising from the use of AI in both public and private sectors can be proactively addressed without the need for lengthy statutory amendments. While reducing uncertainty, introducing consistent privacy standards and encouraging responsible data practices, it will enable the regulator to formulate tailored sectoral codes for different industries and provide additional guidance on important regulatory issues, such as algorithmic transparency and explainability, and privacy impact assessments and its standards.

2. The statute should have extraterritorial effect to ensure compliance with privacy and data protection obligations under the Australian framework by offshore companies. Due to the growing global nature of data processing in general, and the AI-driven data processing in particular, it is important that the statutory safeguards are extended to the personal information of Australian citizens, regardless of where it is processed or stored, or where the companies are situated.

3. Online privacy settings should reflect the privacy by default framework, ensuring that strong privacy settings and stringent data protection measures are embedded in the system, and personal information is safeguarded without requiring users to take additional steps when they interact with the service. This approach empowers users to make informed decisions about their data, while reducing the risk of unintentional privacy breaches, data misuse, intrusive data collection activities or unauthorised sharing of personal information.

4. The collection, use and disclosure of personal information should be fair and reasonable in the circumstances, assessed on an objective standard and in consideration of the potential risks of adverse impact on privacy in proportion to the benefits derived. As AI technologies are becoming more pervasive and data processing increasingly sophisticated, the data stewardship test will incorporate fairness, transparency and accountability, along with data minimisation principles and guardrails for algorithmic biases, in the data processing activities. This test promotes privacy-conscious practices, ensures entities' handling of personal information is within individuals' reasonable expectations, and militates against adverse and unfair use of data.

5. A right to object to the collection, use, or disclosure of personal information, as well as the rights to request erasure and correction of such information, should also be introduced in the statute, subject to appropriate limitations. It will ensure their privacy preferences are respected and empower individuals

to have greater control over their information, by enabling them to withhold consent or opt-out from certain data processing activities, and correcting inaccurate, misleading or outdated information.

6. A right to request meaningful information about how the system operates, without disclosing trade secrets or proprietary information, should be incorporated in the statute. This will enable individuals to gain clarity on the factors considered and logic involved in decision-making processes, the weight assigned to different data points and a description of the likely outcome for the process. AI systems often employ complex algorithms and machine learning techniques to make decisions that affect individuals, such as determining eligibility for services, credit scores or job opportunities, and information about the system will enhance transparency, accountability and individual empowerment.

7. Exemptions afforded to small businesses should be removed to ensure that all businesses targeting consumer data in Australia, irrespective of their size or turnover, are compliant with privacy and data protection obligations under the statute. Current exemption creates loopholes that leave consumer data vulnerable to misuse by small businesses that have access to substantial personal information. It will ensure that consumer data is protected uniformly and increase consumer trust that their data is being handled responsibly, regardless of the company they interact with. It will also ensure alignment with international privacy laws that do not provide exemptions based on business size, such as the General Data Protection Regulation, facilitating cross-border data transfers and international data protection cooperation.

8. A privacy impact assessment and audit should be undertaken by companies mandatorily for activities with high privacy risks, to evaluate the potential impacts of those activities on individuals' privacy, identify measures to address and minimise those vulnerabilities and implement processes to ensure that privacy considerations are integrated into the design, development and deployment stages. For the regulatory burden to be proportionate, the extent of the assessments and audits should be appropriate to the risk of harm. For instance, facial recognition technology and other uses of biometric information, location tracking and social engineering tools that enable creation of deepfake content, clearly pose high levels of risks to individual privacy to warrant extended impact assessments and regular audits. Because of the difficulties in foreseeing AI outcomes and reverse-engineering algorithmic decisions, no single measure can be completely effective in avoiding perverse effects. Paired with proactive risk assessments, auditing outcomes of algorithmic decision-making can help match foresight with hindsight, and enable watchdogs and consumers to know where to look out for untoward outcomes.

9. Individuals should have a direct right of action and statutory tort for serious invasions of privacy. Not only is the absence of such a right depriving individual of a legal avenue to seek remedies, redress, and other recourses for privacy invasions, without a statutory tort, companies may be less motivated to invest in robust privacy safeguards and data protection. Introducing a direct right of action and statutory tort contributes to the development of a privacy-centric culture in companies, and will facilitate investments in privacy-by-design practices.

Human Rights

Given that human rights are interdependent and interrelated, AI affects potentially nearly all human rights. For instance, criminal risk-assessment software widely adopted in the U.S. criminal justice system was inaccurate in forecasting potential future crimes and heavily biased against black defendants, implicating the right to liberty and security and to a fair trial.[45] Autonomous weapon systems, if not programmed to consider nuances of a particular situation or react appropriately to unexpected circumstances, may result in potential high error rates with fatal casualties, posing a potential threat to the right to life.[46] AI-enabled content moderation on social media will primarily impact freedom of expression and right to information, but also freedom of religion if it is used to identify and remove religious content, press freedom if it targets news content, and freedom of association when it results in the removal of online groups, pages and content that facilitate the gathering of individuals. Google's image recognition algorithms, for instance, accidentally categorised two black persons in a photo as gorillas and Amazon's algorithm appeared to discriminate against women,[47] contrary to principles of equality and non-discrimination. AI applications programmed to advertise different treatment recommendations based on the insurance status or income of patients may adversely impact right to healthcare, and in certain cases right to life, especially for marginalised communities. AI applications currently being used could amplify the impact on human rights, democracy and the rule of law at scale, affecting larger parts of society and more people at the same time.

[45] Julia Angwin et al, 'Machine Bias', *ProPublica* (Blog Post, 23 May 2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Lauren Goode, 'Facial recognition software is biased towards white men, researcher finds', *The Verge* (11 February 2018) <https://www.theverge.com/2018/2/11/17001218/facialrecognition-soft ware-accuracy-technology-mit-white-men-black-women-error>.
[46] Charline Daelman, 'AI through a Human Rights Lens. The Role of Human Rights in Fulfilling AI's Potential' in Jan De Bruyne and Cedric Vanleenhove (eds), *Artificial Intelligence and the Law* (Cambridge University Press, May 2021) 126.
[47] James Vincent, 'Google 'fixed' its racist algorithm by removing gorillas from its image-labeling tech', *The Verge* (12 January 2018) <https://www.theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai>.

Australia is the only developed democracy in the world without a bill of rights or national legislation protecting human rights.[48] Instead, the government introduced the *Human Rights (Parliamentary Scrutiny) Act 2011*, which enables 'parliamentary scrutiny of new laws for consistency with Australia's [international] human rights obligations and to encourage early and ongoing consideration of human rights issues in policy and legislative development'.[49] However, this statute advisory in nature, merely allowing parliamentarians introducing bills to issue a statement of compatibility with human rights and its review by an independent Parliamentary Joint Committee on Human Rights, without any mechanism to contest for non-compliant laws or individuals' right to seek remedies. Absent an enforceable domestic human rights instrument, we recommend that the AI framework should expressly incorporate human rights consideration as a baseline requirement for the design, development and deployment of AI in Australia, aligned with the international human rights instruments that the country has ratified.

Consistent with the draft Data Privacy Guidelines for the development and operation of Artificial Intelligence solutions issued by the UN Special Rapporteur on the Right to Privacy and the UN Guiding Principles on Business and Human Rights, we also recommend that human rights assessment and audits should be adopted throughout the lifecycle of AI, including the planning, testing, correction, and implementation phases. External certification of an approved auditor in data privacy with expertise in AI should also be considered. Unless there is developed a specific international law mechanism for settling jurisdictional issues, AI solutions operating across multiple jurisdictions should implement and operate as a multinational federation of individual single jurisdiction AI solutions, while respecting conflict of law principles.

3. Are there any further non-regulatory initiatives the Australian Government could implement to support responsible AI practices in Australia? Please describe these and their benefits or impacts.

Despite having far reaching opportunities, adoption of AI has major ramifications for vulnerable groups including children, persons with disabilities as well as it may often bring potential threats to human rights, social inclusion, and cultural values.[50] However, some advocates of AI may claim that AI regulation by the government will slow innovation.[51] We propose two crucial non-regulatory measures to pre-empt

[48] George Williams and Lisa Burton, 'Australia's Exclusively Parliamentary Model of Rights Protection' (2013) 34(1) *Statute Law Review* 58, 59

[49] Commonwealth, Parliamentary Debates, House of Representatives, 30 September 2010, 271 (Robert McClelland, Attorney-General).

[50] Toby Walsh et al, 'The effective and ethical development of artificial intelligence: an opportunity to improve our wellbeing' *Australian Council of Learned Academies* (Report, 2019) 98 <https://acola.org/wp-content/uploads/2019/07/hs4_artificial-intelligence-report.pdf>.

[51] Roger Clarke 'Regulatory alternatives for AI' (2019) 35(4) *Computer Law & Security Review* 398, 402.

responsible deployment of AI so that industries, designers and developers in Australia can enhance AI innovations without violating the commitment to ensuring an equitable, inclusive and fair society for its citizens.

According to the OECD Repository, there exists more than 800 policy documentations globally as of July 2023 about responsible, fair, transparent and accountable AI which provide an ethical lens to AI policymaking.[52] However, ethics alone cannot guarantee reliable, safe and trustworthy AI systems, it needs actual design decisions that reflect the human-centred AI (HCAI) approach.[53] While multinational companies such as Google, Meta, IBM, Apple and Microsoft have their own ethical design guidelines, it is important for the Australian Government to formulate a mandatory design guideline for all companies dealing with AI design, development and deployment observing the following non-exhaustive principles.

1. *Bringing social and cultural situatedness and justice at the centre of design:* Designing AI technologies must respond to socio-technically situated plurality,[54] which means that designers need to care about the situations of the individuals whose lives are going to be affected by such designs. Approximately 3.2% of Australia's population is made up of Aboriginal and Torres Strait Islander people,[55] and AI's potential impacts on such vulnerable populations could lead to further marginalisation.[56] AI algorithms can inherit biases present in training data, resulting in discriminatory outcomes for these underrepresented groups. Biased decisions in such areas as law enforcement, immigration or counterterrorism could disproportionately impact such groups and undermine trust in national security institutions. However, clearly developed guidelines for how to design, develop and deploy AI systems by avoiding such biases and harms needs to be in place.

   More than 250 indigenous languages, including 800 dialects, are spoken throughout Australia,[57] and AI systems, including generative AI, needs to be sensitive to such multilingualism and cultural diversity. AI designers and developers must address that marginalised communities, including indigenous people, do not exist in self-reproducing, self-definition worlds of meaning in Australia;

[52] 'National AI policies & strategies' *OECD.AI* (online, 2021) <https://oecd.ai/en/dashboards/overview/policy>.
[53] Ben Shneiderman, *Human-centered AI* (Oxford University Press) 15.
[54] Tobias Matzner 'Plural, Situated Subjects in the Critique of Artificial Intelligence' in Andreas Sudmann (ed), *The Democratization of Artificial Intelligence* (2019) 109, 119.
[55] 'Census of Population and Housing - Counts of Aboriginal and Torres Strait Islander Australians' *The Australian Bureau of Statistics* (website, 31 August 2022) <https://www.abs.gov.au/statistics/people/aboriginal-and-torres-strait-islander-peoples/census-population-and-housing-counts-aboriginal-and-torres-strait-islander-australians/2021>.
[56] Nani Jansen Reventlow, 'How Artificial Intelligence Impacts Marginalised Groups' *Digital Freedom Fund* (online, 29 May 2021) <https://digitalfreedomfund.org/how-artificial-intelligence-impacts-marginalised-groups/>.
[57] 'Languages Alive' *The Australian Institute of Aboriginal and Torres Strait Islander Studies* (online, website) <https://aiatsis.gov.au/explore/languages-alive>.

rather, they live in complicated, contentious intercultural realms,[58] and this should be reflected in their designs. For example, the common crawler technology that is used in most AI systems uses only textual languages on the internet, which is predominantly in English. This means that most content is in one language and in a limited format. These limitations on how the data is collected need to be further assessed to ensure AI systems are not skewed to one part of the population, and ineffective, or even harmful, for the rest. Language models developed through other modalities (e.g., image, voice-based and visual) could be more helpful as a design alternative to avoid exclusionary practices.

Design guidelines should also include ways to remove bias from datasets across class, gender, sexual identity, nationality, religion, ethnicities and other protected characteristics in Australian societies. AI models, such as large language models, can rapidly learn demeaning language and harmful stereotypes about groups who are frequently marginalised. As already mentioned, training data often reflects historical patterns of systemic injustice, which can be compounded for certain intersectionalities, around gender and race. The sophistication of AI algorithms makes it increasingly challenging to distinguish between real and synthetic media, raising concerns about its potential discrimination. AI design and development should therefore be evaluated not only on the fairness of the technical systems, but also the environments in which it is developed, the diversity of the AI development team and their responsiveness to languages and multi-linguality. Australian companies must therefore recognise the plurality of worldview and ethics existing in the society while drawing on ontological and epistemological distinctions to design responsible and equitable AI systems that do not favour a specific or dominant worldview.

2. *Promoting rights-based design as core design philosophy:* AI and AI-empowered systems and technologies can help particular groups in various ways, but these systems are prone to have bias and they reproduce systemic inequalities in social fabrics. AI systems, therefore, need to prioritise human rights, making it a key design consideration for the developers since inclusion, safety, transparency and fairness come as its derivatives. We propose to incorporate a human rights-based approach to the AI design guideline focusing on promoting and defending human rights and normatively based on international human rights standards of the United Nations.[59]

[58] David Martin, 'Rethinking the design of Indigenous organisations: the need for strategic engagement' *Centre for Aboriginal Economic Policy Research* (Report, No 248/2003) iv <https://openresearch-repository.anu.edu.au/bitstream/1885/41861/3/2003_DP248.pdf>.
[59] 'Human Rights-Based Approach' *UN Sustainable Development Group* (online) <https://unsdg.un.org/2030-agenda/universal-values/human-rights-based-approach>.

We note that the existing discrimination laws, including the *Disability Discrimination Act 1992*, the *Racial Discrimination Act 1975*, the *Sex Discrimination Act 1984* and the *Age Discrimination Act 2004*, are inadequate in addressing discriminations caused surreptitiously or unconsciously, or by "black box" decision-makers. These laws lack the structural capacity to mitigate risks of discrimination resulting from assumptions and biases of individuals, inaccuracies in information, or gaps in historical data that are embedded in the datasets used by the algorithmic systems to train itself. Therefore, the rights of children and persons with disabilities within the realm of AI and how safety by design is critical to address such vulnerable groups merits special consideration.

Among Australia's about 5.1 million children, around 1.5 million are under the age of four, 2.2 million are between the ages of 5 and 12, and 1.4 million are between the ages of 13 and 17.[60] We propose that Australia should reflect on and draw from already existing frameworks such one developed by UNICEF,[61] and the United Nations Convention on the Rights of the Child, to formulate design guidelines for AI with children and youth. These frameworks can work as baselines to establish rights of minors in relation to AI development and deployment, but we suggest the Australian Government should collaborate with global bodies such as UNICEF, Amnesty International and the Office of the United Nations High Commissioner for Human Rights (OHCHR) to develop deeper understanding about how AI systems can protect, provide for and empower children and youth, including those coming from the First Nations families.

We note that concerns have been raised by the Australian eSafety Commission regarding the possibility for predators to automate child grooming.[62] We reiterate that Australian AI systems should incorporate safety by design and privacy by design to reduce likely harms to minors, including production of child sexual abuse materials. Rapid development of AI systems without adequate guardrails on data collection and labelling for training data pose disproportionate risks for minors, especially in terms of their privacy, exposure to harmful or age-inappropriate content, bias and discrimination and exploiting their vulnerabilities to be targets of unethical or manipulative advertising. Specifically, diffusion models, such as Stable Diffusion and OpenAI's Dall-E, are being used to produce child sexual abuse materials.[63] An emerging concern in the field of child safety is not

[60] 'Face the facts: Children's Rights' *The Australian Human Rights Commission* (online) <https://humanrights.gov.au/our-work/education/face-facts-childrens-rights#:~:text=About%20Australian%20children,between%2013%20and%2017%20years>.
[61] 'Policy guidance on AI for children' *UNICEF* (Report, 2021) <https://www.unicef.org/globalinsight/media/2356/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf>.
[62] Josh Butler, 'AI tools could be used by predators to 'automate child grooming', eSafety commissioner warns' *The Guardian* (online, 20 May 2023) <https://www.theguardian.com/technology/2023/may/20/ai-tools-could-be-used-by-predators-to-automate-child-grooming-esafety-commissioner-warns>.
[63] David Thiel, Melissa Stroebel and Rebecca Portnoff, 'Generative ML and CSAM: Implications and Mitigations' *Standford University* (Report, June 24, 2023) <https://fsi.stanford.edu/publication/generative-ml-and-csam-implications-and-mitigations>.

only the creation of illegal synthetic media of real children but also to make sexual abuse materials of children that do not exist. We note that the *Online Safety Act 2021*, designed to improve and promote online safety for Australians, can address certain content-related harms – for instance, those arising from deepfake, non-consensual intimate images, or cyber-bullying materials – using a range of complaint and control mechanism ranging from content removal and blocking to access-control systems. However, the statute primarily focuses on content regulation and applies only to certain service providers, possibly excluding many entities involved in intrusive data collection and processing activities and having limited application in AI contexts. We propose that the statute is adequately updated to adapt to AI systems.

We are also concerned about the 4.4 million Australians with disabilities,[64] since AI systems can discriminate against facial differences, gestures, speech impairment and other forms of disability. Since recruiting processes increasingly use algorithms to screen out candidates, the immediate effect of these biases and exclusionary processes will be, in many situations, to impede the job rights of disabled persons. Many countries have already started making important decisions regarding who receives public services like social safety benefits or health insurance using predictive AI models, which in many cases, as highlighted by the Australian Human Rights Commission, are exacerbating biases and can disproportionately affect people with disabilities.[65] Using AI for biometrics could put disabled individuals in dangerous positions, especially when it decides whether or not people should move across nations or regions during emergencies and humanitarian crises. In order to protect the rights of disabled individuals, we recommend that the Australian Government develop design guidelines based on the United Nations Convention on the Rights of Persons with Disabilities, while developing and implementing AI tools that directly affect those at-risk groups, such as biometrics, facial recognition technology, and emotional recognition technology. In addition, the Australian Government should work with the OHCHR and other international human rights organisations to adopt more accessible rules and guidelines for the businesses who create AI tools and to compel those businesses to make the reasonable adjustments that people with disabilities require.

It is therefore essential that AI ecosystems, including generative AI, incorporate safety and privacy by design principles throughout the product lifecycle to ensure thoughtful development and deployment.

---

[64] 'Disability, Ageing and Carers, Australia: Summary of Findings' *The Australian Bureau of Statistics* (website, 24 October 2019) <https://www.abs.gov.au/statistics/health/disability/disability-ageing-and-carers-australia-summary-findings/2018>.
[65] Edward Santow, 'Commissioner's Foreword: Using Artificial Intelligence to Make Decisions: Addressing the Problem of Algorithmic Bias', *Australian Human Rights Commission* (November 2020) <https://humanrights.gov.au/our-work/rights-and-freedoms/publications/using-artificial-intelligence-makedecisions-addressing>. See also Sheridan Wall and Hike Schellman, 'Disability Rights Advocates are Worried about Discrimination in AI hiring Tools', *MIT Technology Review* (online, 21 July 2021) <https://www.technologyreview.com/2021/07/21/1029860/disability-rights-employment-discrimination-ai-hiring/>

Specifically, systems should remove harmful content for minors from training data, for example through hashing and matching techniques, as well as detect harmful input and output prompts. Additionally, AI systems should adopt robust data protection policies, which includes measures ensuring strong encryption and anonymisation techniques to protect individuals' privacy, and enable individuals to access, rectify and erase personal information, especially if it concerns a minor. To facilitate the exercise of these rights, AI systems should have user-friendly interfaces that allow users to easily navigate privacy settings and manage their personal information. However, building minimum universal standards for ethical AI for these groups might not be enough, they need to be ratified strongly to block high-risk AI applications that pose possible risks of violating the rights of vulnerable groups.

14. Do you support a risk-based approach for addressing potential AI risks? If not, is there a better approach?

We support a risk-based approach for addressing potential AI risks, recognising the nascency of comprehensive literature available in this field and the role of a risk-based approach as a meaningful starting point in understanding and establishing a baseline for AI systems and their underlying concerns. The European Commission's proposed Artificial Intelligence Act proposes a similar risk-based approach to classify AI systems based on their threat to health, fundamental rights and safety. That said, a risk-based approach is only effective if (a) it does not limit itself to use cases of artificial generative intelligence (AGI) rather takes into account the broader AI ecosystem, including use cases, significant feature modifications, third-party use cases and societal impact of these systems, and (b) it is complemented by a public impact assessment. Further, what constitutes a "high risk" versus "low risk" use should not be predetermined, rather assessed on the basis of periodic and systematic review of the systems, their use cases and their impact.

AI is in incipient stages of development and it is not possible to accurately and adequately estimate and establish the wide range of risks it poses or conclusively determine efficacy of the risk mitigation strategies. There is no existing common language or taxonomy of AI risks.[66] To future-proof any regulatory framework, it is therefore necessary that it does restrict itself to existing use cases and corresponding mitigation approaches, rather take a systemic view and mandate periodic review of the entire system, including the use cases, access to its application programming interface (API), the underlying policies and processes, and a risk-based impact assessment. It is also necessary that the

[66] Bernd Wirtz, Jan Weyerer and Ines Kehl, 'Governance of artificial intelligence: A risk and guideline-based integrative framework' (2022)
39(4) *Government Information Quarterly* 101685, 101685.

impact-based risk assessment is updated at a regular frequency, given the rapidly evolving nature of AGI technologies.

Additionally, there are needs to include independent ethics-based audit mechanisms to supplement a risk-based approach that will specifically focus on the impact of AI systems and ensure the interpretation of risk by individual companies are neither subjective nor downplayed them. This requires regulatory frameworks to mandate privacy-preserving access to data for researchers and independent auditors to supplement company-led assessments and institute public impact assessments. To be feasible and effective, an ethics-based auditing should be continuous and constructive, approach ethical alignment from both a systems and policy perspective, and be aligned with public policies and incentives for ethically desirable behaviour.[67]

We further recommend that the Australian Government establishes a clear throughline between an impact-based risk assessment and the enforcement framework, which would require a risk-based approach to be complemented with necessary revisions in existing regulatory frameworks. This would address complexities with agility, wherein technological developments outpace legislative processes. An updated framework for existing laws, as described in response to Question 2 above, to address "old-fashioned abuses" – for example, employment discrimination, surveillance and data breach – will be critical to move fast and enforce on "known risks".

15. What do you see as the main benefits or limitations of a risk-based approach? How can any limitations be overcome?

First, the most established challenge for developing a comprehensive risk-based approach is the subjective and contextual nature of the risks and its differential impact on specific groups. Further, these risks are not constant, rather rapidly evolving, therefore, a standardised approach is likely to fall short of addressing the broad range of risks posed by AI systems.[68] A risk-based based approach threatens to exacerbate gaps on explainability, a critical tenet of well-established AI ethics frameworks, because the process will only identify the "what" and not the "why".[69]

[67] Jakob Mökander and Luciano Floridi, 'Ethics-based auditing to develop trustworthy AI.' (2021) 31(2) *Minds and Machines* 323, 327.
[68] Hervé Corvellec, 'Organizational risk as it derives from what managers value: A practice-based approach to risk assessment' (2010) 18(3) *Journal of Contingencies and Crisis Management* 145, 154.
[69] Omar Khadeer Hussain, 'The process of risk management needs to evolve with the changing technology in the digital world' (2022) 16(3) *Service Oriented Computing and Applications* 143, 145.

Second, a risk-based approach leaves assessments to the interpretation of individual companies, which could be downplayed and biased towards the company's existing capabilities, thereby threatening lower investments in building more robust integrity tools. In absence of open-source models, comprehensive documentation and privacy-preserving researcher access to data, the government does not have the ability to audit or confirm if the risks were accurately and adequately identified.

AI systems are in nascent stages of development, and even routine tasks pose can compound to unforeseen, damaging consequences. Therefore, a risk-based classification of AI systems would inadvertently result in potentially harmful systems "falling through the cracks" and risk assessments alone would not necessarily confirm if the AI systems are compliant with international human rights standards.

To address some of these limitations, the Australian Government should work with international partners to mandate impact and fairness assessments of AI systems that have been developed outside of its territory. Before implementing AI-driven projects at scale, it should be mandatory for companies to undertake impact assessment processes that are grounded in international human rights principles and publicly share results, including inviting comments from stakeholders. This is important because external AI systems can dominate, exploit, and even dehumanise vulnerable populations within Australian territory in an imbalanced way and introducing repairs and post-hoc fixes can often be ineffective to resist those harms. These assessments should prioritise the issues that are critical for underserved societies, specifically among ethnic, sexual and religious minorities, the First Nations populations as well as low-and- middle-income communities. Attention should be paid to topics on digital sovereignty, infrastructural and regulatory capacities, harms associated with the labour and material supply chains of AI technologies, beta-testing of new features and products, and exploitative commercial use. Recent investigative reports[70] reveal how the global AI industry is exploiting nefarious strategies to violate such citizen rights for their profit making. Foreign corporations in South Africa use AI technologies to monitor its people, exporting surveillance data for racial control. Venezuela's AI industry seeks cheap labour during economic crises, while Jakarta's Gojek taxi drivers fight back by strengthening worker power through community actions. The Australian Government, in collaboration with its international partners, should put increasing pressure on corporations to achieve required ethical standards and public impact assessment reports will play a critical and concrete role in this regard.

As a major global power, Australia should be aware of the questions of geopolitical power imbalance in any technology-transfer partnership since AI development and applications will be deeply divisive if the

[70] Karen Hao, 'Artificial intelligence is creating a new colonial world order' *MIT Technology Review* (Blog Post, 19 April 2022) <https://www.technologyreview.com/2022/04/19/1049592/artificial-intelligence-colonialism/>.

power relations among the countries are not equitable. A risk-based approach would be effective and comprehensive if countries can agree on the existing structural inequality in developing governance and legal frameworks for AI and should find ways of mutual cooperation to reduce them. Actors and decision-makers from the respective countries must have control over the strategic decisions of application of AI technologies in their own contexts so that AI practices cannot cause more harm than benefits for those stakeholders. Australia should partner with the U.S. and other major economies for mobilising research funds and resources to address both innovation and ethical implications of AI systems worldwide. It should also play a leading role in identifying mechanisms and protocols that ensure meaningful participation of underserved stakeholders and reduce institutional barriers for such participation. At the same time, Australia should play its ethical role to enable the Global Majority stakeholders to engage in formally defined roles to collectively participate in the global AI technology transfer and governance processes for a more inclusive AI future of the world.

There is no existing common taxonomy for risks which results in challenges with standardisation and scaling of a risk-based approach. It is critical that governments do the groundwork of collaborating with academia, civil society and industry to develop a shared vocabulary which will provide the foundation for future proofing regulatory frameworks. Australia's AI Ethics Principles draw on the OECD's Principles and it is already working with G7, G20 and OECD countries on AI ethics and research and development.[71] It should enhance such collaborations and cooperation with the United Nations, International Organization for Standardization, the World Trade Organization and other international organisations to develop agreeable frameworks on AI. Working together with partners will benefit Australia in diverse ways including making national standards and norms for AI consistent to the international scenario. This will reduce fragmentation and conflicting regulations across different jurisdictions, facilitating smoother international trade and cooperation. Given the transnational nature of AI systems, it is imperative to ensure the process of developing such frameworks is diverse and inclusive, factoring in the unique societal, privacy, cybersecurity and accountability challenges of different regions. Specifically, by engaging with the WTO, Australia can contribute to the establishment of rules and regulations that reduces barriers to trade, protects intellectual property rights, addresses privacy concerns and facilitates market access for compliant AI systems and services. As part of this process, it is critical to engage with diverse civil society and academic institutions to ensure transparency, accountability, diversity of perspectives and ownership in the formulation and implementation of any international AI frameworks or acts.

[71] Time Bradley, 'How Australia takes a global approach to human-centred AI for maximum benefits' *OECD.AI* (online, 11 May 2020) <https://oecd.ai/en/wonk/how-australia-takes-a-global-approach-to-human-centred-ai-for-maximum-benefits>.