



Reply to the Civil Society Questionnaire to Inform the Special Rapporteur on the Rights of Persons with Disabilities' Thematic Report on Artificial Intelligence and Disability

November 2021

The Centre for AI and Digital Ethics (CAIDE) is a cross-disciplinary research centre with four University of Melbourne member faculties – Melbourne Law School, Engineering and IT, Arts and Science. CAIDE brings a cross-disciplinary perspective on the ethical, regulatory and legal issues relating to Artificial Intelligence (AI) and digital technologies. Our research and teaching seek to explore the impact, deployment and governance of this emerging technology across society. We examine issues of fairness, privacy, accountability and transparency in this emerging technology to further our understanding, but also to guide the development and appropriate policy settings for effective use across society.

The Melbourne Social Equity Institute (MSEI) supports interdisciplinary research on social equity issues across the full spectrum of social life including health, law, education, housing, work and transport. We bring together researchers from across the University of Melbourne to identify unjust or unfair practices that lead to social inequity and work towards finding ways to ameliorate disadvantage. We facilitate researchers working with government and community organisations and help with the dissemination and translation of research for public benefit. Our research themes include Mental Health and Society and Digital Access and Equity.

The Melbourne Disability Institute (MDI) is an interdisciplinary research institute established by the University of Melbourne in 2018 to build a collaborative, inclusive, interdisciplinary and translational research program to improve the lives of people with disability. MDI works with academics, people with disability, policymakers, the disability sector and broader community to expand and apply much-needed evidence to address the complex problems facing people with disability and their families.

This submission was prepared by Dr Yvette Maker, Dr Piers Gooding, Dr Sue Olney, Ms Kirsten Deane, Ms Gabby Bush and Prof Jeannie Paterson. For enquiries, please contact y.maker@unimelb.edu.au and p.gooding@unimelb.edu.au.

We welcome the opportunity to respond to the Special Rapporteur on the Rights of Persons with Disabilities' questionnaire for NHRIs, OPDs, human rights defenders and civil society. This submission draws on research conducted by the authors and supported by three University of Melbourne bodies, the Centre for AI and Digital Ethics, the Melbourne Social Equity Institute and the Melbourne Disability Institute. This includes research on the legal, social and technological issues arising in relation to the use of AI, ML or ADM – as well as broader rights issues associated with emerging technology and digitisation – in three key areas:

1. the law and politics of data concerning mental health and disability;

2. the allocation of personal budgets for care and support services in Australia’s National Disability Insurance Scheme; and
3. the adequacy of existing regulation of AI, especially non-binding statements of AI ethics, in addressing disability human rights concerns.

As we discuss in this submission, attention to the design, use and regulation of AI, ML and ADM in relation to the human rights of person with disabilities is required in multiple respects. It is also important to note that narratives of innovation can be used to distract from longstanding problems with *existing* policies and practices, or to give the appearance of being cutting-edge, and can divert attention and institutional resources from where they are needed most.

1. Please provide information about the extent to which technologies such as AI, ML and ADM are used in engagements between the individual and State bodies and private bodies in a way that allows persons with disabilities to better engage positively in society. The following relationships are of particular interest:

- (i) the individual and State bodies

We note three examples of apparently rights-affirming technological practices with potential to improve relationships between persons with disabilities and state bodies. One concerns **refugee connectivity and access to services**. In 2018, researchers at the Data & Society research institute undertook a survey of mobile phones, mental health, and privacy at a Syrian refugee camp in Greece.¹ The research has relevance to persons with psychosocial disability or any disability in situations of forced migration, particularly given the risk of disability-based discrimination in national immigration policies. Although it is not clear the extent to which AI, ML and ADM were used in the connectivity efforts – which perhaps suggests the value of a broader enquiry into ‘data-driven technologies’ – the authors demonstrated ways that refugees’ phones were essential to aid, survival and well-being. The survey design simultaneously employed two distinct methodologies: one concerned with mobile connectivity and *mental health*, and a second concerned with mobile connectivity and *privacy*. The research was premised on a view that privacy can be essential to easing distress and mental health, both in terms of receiving support, and in the lives of refugees and asylum seekers more generally, particularly those at risk of persecution.

Another apparently positive practice concerns **monitoring of healthcare services in disability and aged care settings**. In 2018, Lisa Pont and colleagues used basic information technology to monitor prescription medicine use in 71 residential aged care facilities in Australia, with the aim of identifying systemic problems, such as prescribing errors and medication misuse. A major concern they sought to address was excessive prescription of sedating psychiatric pharmaceuticals. Where excessive prescription patterns could not be easily explained, their system flagged the need for further regulatory checks. Similarly, Johanna Westbrook and colleagues produced some evidence indicating that commercial electronic prescribing or ‘e-prescribing’ systems could reduce hospital in-patient prescribing error rates, including in a psychiatric ward in a large Australian hospital, mainly by reducing the number of incomplete, illegal, or unclear medication orders.² From a rights-based perspective, such measures offer preventative monitoring capabilities for upholding the right to the enjoyment of the highest attainable standard of health, the right to physical and mental integrity, and the right to be free from cruel, inhuman or degrading treatment.

¹ Mark Latonero, Danielle Poole and Jos Berens, *A Survey of Mobile Phones, Mental Health, and Privacy at a Syrian Refugee Camp in Greece* (Harvard Humanitarian Initiative and the Data & Society Research Institute, 2018) 47.

² Johanna I Westbrook et al, ‘Effects of Two Commercial Electronic Prescribing Systems on Prescribing Error Rates in Hospital In-Patients: A Before and After Study’ (2012) 9(1) *PLOS Medicine* e1001164.

Another example of **data-driven technologies in systemic advocacy and the monitoring of state-run services** (again, not AI, ML or ADM) occurred in 2021 in Canada, when a coalition of open data advocacy groups in collaboration with disabled people's organisations aimed to crowd-source a database of congregate institutions for disabled people. They aimed to trace the impact of COVID-19 on persons with disabilities in Canada and advocate for vaccination priority. A collaboration between open data groups and disabled peoples' organisations led to a public event in which members of the public could join an online initiative to 'Hack the Data Gap' and create an up-to-date database of relevant residential facilities.³ Although such initiatives should not replace robust state mechanisms for gathering vital statistical data, in line with Article 31 of the CRPD, the process demonstrated how civil society organisations can marshal public data-gathering as a democratic resource to protect the rights of persons with disabilities.

(v) individuals and their right to live independently and being included in the community

Persons with disabilities have a right to supports and services that are designed for their situations, needs and goals. Data created when someone engages with government services can provide a useful base of information to do this. However, **many datasets are disconnected, as they are organised around different government agencies and services.** In Australia this is further complicated by the different information collected by the different tiers of government, and differences in services provided (and therefore data gathered) by individual state governments. The National Disability Data Asset (NDDA) aims to link data from different government agencies across Australia to better understand the life experiences of persons with disabilities, while protecting people's privacy.⁴ The NDDA is currently in an 18-month pilot phase. If established, it would be Australia's most comprehensive human services data asset, covering areas such as education, jobs, income, health, safety and disability supports.

However, integrating data from different sources is not a straightforward process. **Research is exposing significant risks for particular groups of citizens in how their data, or their invisibility in some datasets, is used to inform policy and practice and the allocation of resources.**⁵ During the consultation process held as part of the pilot project, persons with disabilities expressed significant **concerns about privacy protections and the way the data could be used to curtail rather than expand services and supports.** A failure to include the full diversity of persons with disabilities in datasets can also cause, or perpetuate, bias or discrimination against persons with disabilities when these datasets are used to train AI systems.⁶

2. Please provide information about the extent to which technologies such as AI, ML and ADM pose a risk to the rights of persons with disabilities when deployed in relation to the areas highlighted in question 1.

(i) the individual and State bodies

The **rise of the 'digital welfare state'**⁷ is directly relevant to discussions of the rights of persons with disabilities. In Australia, crude ADM was used in a scheme that became known as 'Robodebt', which ran between 2015 and

³ Megan Linton, 'Invisible People and Institutions: No Data About Custodial Institutions for Disabled People in Canada?' (blog) *Datalibre.ca* <<http://datalibre.ca/2021/02/18/invisible-people-and-institutions-no-data-about-custodial-institutions-for-disabled-people-in-canada/>>.

⁴ *National Disability Data Asset* (Web Page, 2021) <<https://ndda.gov.au/>>.

⁵ Virginia Eubanks, *Automating Inequality: How High-tech Tools Profile, Police, and Punish the Poor* (St Martin's Press, 2018); Sue Olney, 'Not My Problem: The Impact of Siloed Performance Management on Policy Design and Implementation' in D Blackman et al (eds) *Handbook on Performance Management in the Public Sector* (1st ed, Edward Elgar 2021)28 <doi.org/10.4337/9781789901207.00008>.

⁶ Anhong Guo et al, 'Toward Fairness in AI for People with Disabilities: A Research Roadmap' [2019] (125) *SIGACCESS Newsletter* 1, 3.

⁷ United Nations General Assembly, *Report of the Special Rapporteur on Extreme Poverty and Human Rights 26 September* (No A/73/396, 2018) <<https://undocs.org/pdf?symbol=en/A/73/396>>.

November 2019.⁸ The scheme was designed to recover alleged overpayments of social assistance by using a simple algorithm (which did not use AI or ML) to identify discrepancies between an individual's declaration of income to taxation authorities and to the national social security agency, the latter of which can affect the rate of social assistance payments to which a person is entitled. Discrepancies identified by the program were treated as evidence of undeclared or under-reported income (and hence overpayment of social assistance) and automatically generated debt notices were sent to affected individuals. The program relied on flawed income calculations to assert the debts.⁹ The government's claims that its long-standing policy of data matching was now supported by 'improved automation'¹⁰ highlighted the risks of such processes being used to obscure poor policy-making. The scheme was deemed to be unlawful and, following a class action, the government settled out of court. According to the Australian Broadcasting Commission, the scheme raised \$1.73 billion in debts against 433,000 people, of which \$751 million was wrongly recovered from 381,000 people.¹¹ The federal court of Australia approved the settlement, which was reportedly worth at least \$1.8 billion.

Philip Alston, former UN Special Rapporteur on Extreme Poverty and Human Rights, pointed to the Robodebt fiasco as an example of the way the digitisation of welfare services has:

very often been used to promote deep reductions in the overall welfare budget, a narrowing of the beneficiary pool, the elimination of some services, the introduction of demanding and intrusive forms of conditionality, the pursuit of behavioural modification goals, the imposition of stronger sanctions regimes, and a complete reversal of the traditional notion that the state should be accountable to the individual.¹²

Alston's report suggests that many, if not most, automated welfare systems are primarily designed, deployed and justified to save costs and identify fraud, and not to improve care or expand access. Given the reciprocal relationship between poverty and disability, this trend has major implications for the rights of persons with disabilities. Persons with disabilities are typically over-represented among those receiving social security benefits.

Further, and importantly for the purposes of this submission, the Robodebt scheme was implemented in ways that **discarded previous safeguards** designed to protect social security recipients who had psychosocial and intellectual disabilities.¹³ This suggests the existence of multiple layers of discrimination against persons with disabilities and highlights the use of (albeit crude) automation to extend punitive neoliberal logic in social security systems.

Despite the adverse findings regarding Robodebt, the Australian government and the agency that administers Australia's federal system of individualised disability support funding, the National Disability Insurance Scheme (NDIS), also recently proposed the introduction of a series of reforms that would include **a new process for allocating individual funding that relied on the use of algorithms**. Under the proposed changes, participants in the scheme would be assessed by contracted allied health staff using short, standardised assessment tools such as WHODAS and PEDiCAT. These 'independent assessments' would then be used to create 400 'personas'. Participants would be matched to one of the personas to determine how much funding they would receive in

⁸ Terry Carney, 'Robo-Debt Illegality: The Seven Veils of Failed Guarantees of the Rule of Law?' (2019) 44(1) *Alternative Law Journal* 4.

⁹ Peter Whiteford, 'Debt by Design: The Anatomy of a Social Policy Fiasco – Or Was It Something Worse?' (2021) 80 *Australian Journal of Public Administration* 340.

¹⁰ *Ibid* 353.

¹¹ 'Robodebt Condemned as a "Shameful Chapter" in Withering Assessment by Federal Court Judge', *ABC News* (online, 11 June 2021) <<https://www.abc.net.au/news/2021-06-11/robodebt-condemned-by-federal-court-judge-as-shameful-chapter/100207674>>.

¹² United Nations General Assembly (n 7).

¹³ Christopher Knaus, 'Centrelink's New Robodebt Trial Bypasses Previous Safeguard for Mentally Ill', *The Guardian* (online, 14 August 2018) <<https://www.theguardian.com/australia-news/2018/aug/15/centrelinks-new-robodebt-trial-bypasses-previous-safeguard-for-mentally-ill>>.

their individual budgets.¹⁴ This would constitute a major departure from the existing system, where applicants and participants provide information about their goals and needs and evidence from their own treating specialists. While it is unclear whether ML or other AI technology was intended for use in the proposed system, the proposal effectively drew on mechanistic and algorithmic approaches to assessing disability support needs. Concerns were raised about the quality of the data being collected to build the personas, including the small dataset on which they would be based, and the predominance of data from the ‘independent’ health assessments rather than participants’ own perspectives. Two pilots of the assessment process had been conducted, and the personas were under construction, when the reforms were abandoned following a determined community campaign against them by persons with disabilities, their families, their representative organisations, allied health professionals, lawyers and academics. While the government claimed that the new system would produce more equitable and flexible budgets,¹⁵ persons with disabilities argued strongly that the process would not adequately capture their individual needs and expressed their concern that funding would be cut as a result.¹⁶

Another concern on this front is that **disability support and other services may become conditional on forms of monitoring and surveillance enabled by automation**. For example, ‘Electronic Visit Verification’ is being used in the US to log the precise times at which mental health service home-visits begin and end.¹⁷ The initiative serves as an auditing tool for service providers, including to prevent fraud against Medicaid. However, the surveillance of caregiver labour also appears functionally indistinguishable from surveillance of care recipients, creating what Jacob Metcalf describes as a ‘deceptively intrusive tracking of the lives of Medicaid recipients’.¹⁸ Metcalf also warns that information collected could be re-purposed to leverage against users’ Medicaid claims.

Finally, common themes have emerged in government inquiries into various aspects of Australia’s NDIS and its intersection with mainstream/universal services, highlighting competing priorities, jurisdictional ambiguities, information asymmetry, resource asymmetry, inconsistent criteria, and the importance of stewardship of the emerging market to ensure equity of access to information, services and support under the scheme.¹⁹ More broadly, only ten percent of Australians with disability are NDIS participants and little is known about how those outside the NDIS are faring. This raises questions about **risks associated with automated decision-making concerning disability based on skewed or incomplete datasets**.

¹⁴ Joint Standing Committee on the National Disability Insurance Scheme, *Independent Assessments* (Parliament of Australia, 2021) <https://parlinfo.aph.gov.au/parlInfo/download/committees/reportjnt/024622/toc_pdf/IndependentAssessments.pdf;fileType=application%2Fpdf>.

¹⁵ National Disability Insurance Agency, ‘What Happens to Information From the Pilot?’, *NDIS* (Web Page, 21 January 2021) <<https://web.archive.org/web/20210324231855/https://www.ndis.gov.au/participants/independent-assessments/second-independent-assessment-pilot/what-happens-information-pilot>>.

¹⁶ Helen Dickinson and Anne Kavanagh, ‘NDIS Assessments Are Now Off The Table: That’s A Good Thing, The Evidence Wasn’t There’ *The Conversation* (online, 10 July 2021) <<https://theconversation.com/ndis-independent-assessments-are-off-the-table-for-now-thats-a-good-thing-the-evidence-wasnt-there-164163>>; Every Australian Counts, ‘Disability Sector Statement On The Australian Government’s Planned Reforms To The National Disability Insurance Scheme (NDIS)’ (Web Page, 11 March 2021) <<https://everyaustraliancounts.com.au/ndis-sector-statement/>>; Luke Henriques-Gomes ‘Not The NDIS We Fought For: Disability Advocates Fear New Assessments Driven By Cost Cutting’ *The Guardian* (online, 11 March 2021) <<https://www.theguardian.com/australia-news/2021/mar/11/not-the-ndis-we-fought-for-disability-advocates-fear-new-assessments-driven-by-cost-cutting>>.

¹⁷ *21st Century Cures Act 2016* (US) s 12006.

¹⁸ Jacob Metcalf, ‘When Verification Is Also Surveillance’, *Data & Society: Points* (27 February 2018) <<https://points.datasociety.net/when-verification-is-also-surveillance-21edb6c12cc9>>.

¹⁹ Parliament of Australia, ‘Joint Standing Committee on the National Disability Insurance Scheme’ (Web Page, Australian Government, 2021) <https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/National_Disability_Insurance_Scheme>.

(ii) consumers and providers of goods and services

The **role of digital labour platforms in transforming the world of disability care and support work** is a crucial disability rights issue, and one that may be overlooked if inquiries focus overly narrowly on AI, ML and ADM. Digital labour platforms can be designed in ways that institutionalise the exploitation of care and support labour, turning the interests of recipients against (generally low-paid) staff.²⁰ Even as the platforms may help to equitably distribute care and support labour, they raise legitimate concerns, particularly in relation to health and safety, insurance, unpaid work and the long term training needs of the workforce.²¹ Some platforms are not subject to robust government procurement policies or lack the type of regulatory oversight that would secure the rights of either care workers or support recipients.²² Government policies of personalised disability funding, such as the NDIS in Australia and the English adult social care system, commonly seek to reduce costs to the state by encouraging recipients to negotiate lower payment rates with their support providers.²³ This policy incentive drives down the wages of care and support workers, a dynamic that is reinforced by digital labour platforms. While flexibility and choice of good-quality support services is essential for persons with disabilities' enjoyment of the right to live independently and be included in the community, low-paid and insecure employment is unlikely to produce meaningful choice of high-quality supports.²⁴ Privatised platforms also divert resources away from government support to profit-seeking enterprise. This outsourcing can also erode the democratic system as private companies begin to make public decisions for citizens.²⁵ Some independent, not-for-profit advocacy organisations in Australia have become shareholders in digital care labour platforms, raising questions about the independence and accountability of civil society organisations.²⁶

Regarding **ad-tech and the protection of sensitive personal information concerning disability**, in 2019, Privacy International analysed more than 136 popular webpages across France, Germany and the UK related to depression.²⁷ The websites were chosen to reflect those that people would realistically find when searching for help online. The authors found that over three quarters of web pages contained third-party trackers for marketing purposes, which could enable targeted advertising and marketing from large companies like Google, Amazon and Facebook. Most websites, according to the authors, failed to comply with the EU General Data Protection Regulation in upholding individuals' privacy (acknowledging that the UK is no longer part of the EU). In 2020, a follow-up study found concluded that '[g]enerally, most websites analyses haven't taken action to

²⁰ International Labour Office, *World Employment and Social Outlook 2021: The Role of Digital Labour Platforms in Transforming the World of Work* (ILO, 2021).

²¹ Department of Premier and Cabinet, *Report of the Inquiry into the Victorian On-Demand Workforce*, Victorian Government, 2020 s5.6 <https://s3.ap-southeast-2.amazonaws.com/hdp.au.prod.app.vic-engage.files/4915/9469/1146/Report_of_the_Inquiry_into_the_Victorian_On-Demand_Workforce-reduced_size.pdf>.

²² Rick Morton, 'Exclusive: Gov's \$5.8m Aged-Care App Offers "No Duty of Care"', *The Saturday Paper* (online, 23 May 2020) <<https://www.thesaturdaypaper.com.au/news/politics/2020/05/23/exclusive-govs-58m-aged-care-app-offers-no-duty-care/15901560009868>>.

²³ Fiona Macdonald and Sara Charlesworth, 'Cash for Care under the NDIS: Shaping Care Workers' Working Conditions?' (2016) 58(5) *Journal of Industrial Relations* 627; Iain Ferguson, 'Increasing User Choice or Privatizing Risk? The Antinomies of Personalization' (2007) 37(3) *The British Journal of Social Work* 387.

²⁴ Kirstein Rummery and Michael Fine, 'Care: A Critical Review of Theory, Policy and Practice' (2012) 46(3) *Social Policy and Administration* 321, 337.

²⁵ Sarah Valentine, 'Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control' (2019) 46(2) *Fordham Urban Law Journal* 369.

²⁶ Georgia Wilkins, "'A Question of Integrity": Ageing Lobby Group Invests in Controversial Aged-Care App', *Crikey* (online, 14 September 2020) <<https://www.crikey.com.au/2020/09/14/mable-aged-care-app/>>.

²⁷ Privacy International, *Your Mental Health for Sale?* (6 November 2020) <https://www.youtube.com/watch?v=Sbsw51OrvBU&list=UUwyzKZWhsD2YFg8huOaO3IOg&ab_channel=PrivacyInternational>.

limit data sharing [meaning]... personal data are still shared for advertising purposes' with hundreds of third parties with no clear indication of the potential consequences.²⁸

In other contexts, **devices that generate data about behaviour are being used to infer cognitive impairment.**²⁹ Similar to data concerning mental health, and regardless of its accuracy, this information could allow companies to advertise specific treatments, services, or financial products. It could also be sold to other interested parties, such as insurers.

More broadly, in ad-tech, Ben Hutchinson and colleagues at Google demonstrated that social attitudes casting disability as bad and even violent – particularly in regard to mental health – were encoded in AI systems designed to detect hate speech and identify negative/positive sentiment in written text.³⁰ The 'machine-learned model to moderate conversations', according to Hutchinson and colleagues, classifies texts which mention disability and particularly references to mental health conditions as more 'toxic' while 'a machine-learned sentiment analysis model rates texts which mention disability as more negative'.³¹ Such studies highlight how biased datasets create biased algorithms, which has major implications for people's lives, as the next example shows.

(iii) workers/employees and employers

Discriminatory automated hiring programs raise clear human rights concerns. To give one example, Mr Kyle Behm, a high-achieving university student in the US, was reportedly refused a minimum-wage job in 2016 after being 'red-lighted' by the automated personality test he had taken as part of his job application.³² Mr Behm had previously accessed mental health services and had a mental health condition or psychosocial disability. He only became aware of the 'red-lighting' after being informed by a friend who happened to work for the employer. Mr Behm applied for several other minimum-wage positions but was again seemingly 'red-lighted' following automated personality testing. Mr Behm's father, a lawyer, publicised the widespread use of the job applicant selection program and launched a class-action suit alleging that the hiring process was unlawful. He argued that the process violated the *Americans with Disabilities Act of 1990* by being equivalent to a medical exam, for which its use under the ADA for hiring purposes would be illegal. In November 2017, the US retailer Lowe's announced a change to online application processes for retail employees 'to ensure people with mental health disabilities can more readily be considered for opportunities with Lowe's'.³³ Mr Behm was seemingly harmed due to data to which he was never given access. Nor does it appear that Mr Behm had an easily accessible opportunity to contest, explain or investigate the test outcome. Cathy O'Neil argues that this type of algorithmic 'red-lighting' has the potential to 'create an underclass of people who will find themselves increasingly and inexplicably shut out from normal life'.³⁴

²⁸ Privacy International, 'Mental Health Websites Don't Have to Sell Your Data. Most Still Do', *Privacy International* (7 October 2021) <<http://privacyinternational.org/report/3351/mental-health-websites-dont-have-sell-your-data-most-still-do>>.

²⁹ Rolfe Winkler, 'Apple Is Working on iPhone Features to Help Detect Depression, Cognitive Decline', *Wall Street Journal* (online, 21 September 2021) <<https://www.wsj.com/articles/apple-wants-iphones-to-help-detect-depression-cognitive-decline-sources-say-11632216601>>.

³⁰ Ben Hutchinson et al, 'Social Biases in NLP Models as Barriers for Persons with Disabilities' [2020] *arXiv:2005.00813 [cs]* <<http://arxiv.org/abs/2005.00813>>.

³¹ Ibid.

³² Cathy O'Neil, 'How Algorithms Rule Our Working Lives', *The Guardian* (online, 1 September 2016) <<https://www.theguardian.com/science/2016/sep/01/how-algorithms-rule-our-working-lives>>.

³³ 'Lowe's Announces Changes to Online Application Process for Retail Employees' *Letter from Lowes and Bazelon Center for Mental Health Law*, 17 November 2017 <<http://www.bazelon.org/wp-content/uploads/2017/11/Joint-Statement-with-Lowes.pdf>>.

³⁴ O'Neil (n 32).

(iv) learners and the providers of education

In China, the advocacy group Article 19 recently surveyed 27 companies whose **emotion recognition technologies are being trialled in three areas: public security, driving safety, and educational settings**.³⁵ Companies like Taigusys Computing and EmoKit refer to autism, schizophrenia and depression as conditions they can diagnose and monitor using 'micro-expression recognition'. These are pseudoscientific claims with dangerous implications for persons with disabilities (including imputed disabilities). In addition, the authors of the Article 19 report argued that 'firms that purportedly identify neurological diseases and psychological disorders from facial emotions fail to account for how their commercial emotion recognition applications might factor in these considerations when assessing people's emotions in non-medical settings, like classrooms'.³⁶ AI Now Institute, an interdisciplinary research centre examining artificial intelligence and society, have called for a ban on technology designed to recognise people's emotions in certain cases, particularly concerning 'important decisions that impact people's lives and access to opportunities',³⁷ and there is cause to echo these concerns from the perspective of disability justice. There are many other examples of intrusive, extractive data-driven technology being presented as a self-evidently virtuous innovation for the sick and disabled, in what Mara Mills describes as the 'assistive pretext'.³⁸

3. Please provide information on any formal complaints, official investigations, and related jurisprudence arising from allegations of discrimination against persons with disabilities in relation to the matters highlighted in response to question 2.

Regarding the Robodebt scandal, the material noted in response to question 2(i) above includes information in footnotes on formal complaints, official investigations and other relevant matters.

4+5. Please provide information about all relevant criminal and civil laws, codes, regulatory mechanisms, cases and other determinations. Please explain the effectiveness of these laws, codes and regulatory mechanisms in relation to the protection and advancement of the rights of persons with disabilities.

There is ongoing debate about data governance and the broader governance of the information economy. **More attention to the perspectives and requirements of persons with disabilities is required** in these debates. Regarding 'welfare surveillance' and the leveraging of information concerning disability by governments, **attention is needed to the use of secondary personal data**; that is, personal information collected by government agencies or others that was originally collected for other purposes. Finland has made an innovative attempt to improve public governance over the use by government agencies of secondary personal data, including personal data related to disability. The Finnish Parliament approved a new general 'Act on the Secondary Use of Social Welfare and Health Care Data' in March 2019.³⁹ The aim is to ensure flexible and secure use of data by establishing a centralised electronic licence service and a licensing authority for the secondary use of health and social data. The data is dispersed in several information systems managed by different authorities and the new law would clarify regulatory oversight and streamline the processing of data requests, ostensibly allowing faster access to data, improved data security, and clearer application of principles of administrative law, such as improved accountability, transparency, and public participation. Such examples may

³⁵ Article 19, 'Emotional Entanglement: China's Emotion Recognition Market and Its Implications for Human Rights' (January 2021) <<https://www.article19.org/wp-content/uploads/2021/01/ER-Tech-China-Report.pdf>>.

³⁶ Ibid.

³⁷ Kate Crawford et al, *AI Now 2019 Report* (AI Now Institute, December 2019) 6 <ainowinstitute.org/AI_Now_2019_Report.html>.

³⁸ Mara Mills, 'Deaf Jam: From Inscription to Reproduction to Information' (2010) 28(1) *Social Text* 39 <doi.org/10.1215/01642472-2009-59>.

³⁹ Act on the Secondary Use of Social Welfare and Health Care Data (Laki sosiaali- ja terveystietojen toissijaisesta käytöstä, based on government proposal HE 159/2017) March 2019.

warrant further attention to establish if they achieve their stated aims, including with regards to citizen data concerning disability.

Further attention is also required to the **adequacy of statements of AI ethics and other ‘soft law’ responses to concerns about the human rights impacts of AI-informed technology for persons with disabilities**. In recent years, industry, research bodies, governments, international and intergovernmental organisations and professional associations have produced such statements of principles to guide the design and deployment of AI-enabled technology.⁴⁰ These statements are non-binding, reflecting the predominance of industry self-regulation in this area.⁴¹ Concepts associated with social justice and human rights, such as ‘equity’, ‘justice’ and ‘access’, are sometimes included in statements of AI ethics and, in some instances, compliance with human rights principles and laws is identified as an essential feature of ethical AI.⁴² However, **AI ethics statements and codes do not generally address the rights of persons with disabilities**, despite growing calls for the development and use of AI, ML and ADM that protects and promotes disability human rights.⁴³

There is growing agreement that human rights principles offer ‘a sound normative framework to steer AI-systems towards the good’.⁴⁴ They can also provide at least some guidance on the kinds of laws, policies and other measures that might be necessary to enforce them at the local, national and international levels.⁴⁵ In Australia, industry participants and their representative bodies have expressed concern that poorly conceived human rights-based approaches to designing or regulating AI-informed technology may stifle innovation and have adverse consequences for both producers and consumers.⁴⁶ One issue here is that, like statements of AI ethics, **existing guidance for developing human rights-compliant AI generally offer broad principles to be applied across technologies and uses**. For example, in its recent report on ‘Human Rights and Technology’, the Australian Human Rights Commission proposed a ‘human rights by design’ approach to the development and use of AI and other new and emerging technologies, and further work is now required to detail how this should be applied in relation to particular user groups (such as persons with disabilities) and particular technologies and applications.⁴⁷

In its report on *Principled Artificial Intelligence*, which sought to ‘map consensus in ethical and rights-based approaches to principles for AI’,⁴⁸ authors at the Berkman Klein Institute concluded that many existing discussions of principled AI from the business sector appear favourable to human rights language. Governments, by contrast, were less likely to use human rights language. The authors did not discuss why they thought this might be the case. At least three possibilities are that (1) the human rights framework has conceptual limitations in the extent to which it can address the deeper foundations of the information economy and the political and

⁴⁰ Anna Jobin, Marcello Lenca and Effy Vayena, ‘The Global Landscape of AI Ethics Guidelines’ (2019) 1(9) *Nature Machine Intelligence* 389.

⁴¹ Australian Human Rights Commission, *Human Rights and Technology* (Final Report, 2021).

⁴² See, e.g. AI High-Level Expert Group, *Ethics Guidelines for Trustworthy AI* (European Commission, 2019) 4; IEEE, *Ethically Aligned Design: A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems* (1st ed, 2019) 18; Department of Industry, Science, Energy and Resources, *Australia’s Artificial Intelligence Ethics Framework* (Australian Government, 2019) <<https://www.industry.gov.au/data-and-publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles>>.

⁴³ Aspen Lillywhite and Gregor Wolbring, ‘Coverage of Ethics within the Artificial Intelligence and Machine Learning Academic Literature: The Case of Disabled People’ (2021) 33(3) *Assistive Technology* 129; Carine Marzin, *Plug and Pray? A Disability Perspective on Artificial Intelligence, Automated Decision-Making and Emerging Technologies* (European Disability Forum, 2018); Australian Human Rights Commission (n 41).

⁴⁴ Nathalie A Smuha, ‘Beyond a Human Rights-Based Approach to AI Governance: Promise, Pitfalls, Plea’ [2020] *Philosophy & Technology* 1 (advance) <doi.org/10.1007/s13347-020-00403-w>.

⁴⁵ Mark Latonero, *Governing Artificial Intelligence: Upholding Human Rights and Dignity* (Data & Society, October 2018).

⁴⁶ Australian Human Rights Commission (n 41) 46, 127, 130.

⁴⁷ *Ibid* 11, 19; Department of Industry, Science, Energy and Resources (n 42).

⁴⁸ Jessica Fjeld et al, *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI* (SSRN Scholarly Paper No ID 3518482, Social Science Research Network, 15 January 2020) 64-5 <<https://papers.ssrn.com/abstract=3518482>>.

economic might of Big Tech firms, (2) the human rights framework currently lacks the specificity to impose concrete requirements that would better secure human rights, or (3) governments, which are increasingly consumers and users of AI-based technology, are reluctant to subject their own use of technology to human rights-based scrutiny.

A further issue with **human rights-based approaches** is that they **do not tend to involve detailed consideration of the features and limitations of technologies**. For example, while there is clear evidence the data used to train automatic speech recognition systems is not representative of the full diversity of persons with disabilities, it is unclear whether *all* AI-enabled speech recognition systems can be made accessible and accurate for all speakers where only a small number of people have a particular speech variation, or where the overall extent of variation in the population is great.⁴⁹ Any such limitations do not mean that discriminatory outcomes are unavoidable or that other human rights issues can be ignored, but **an understanding of the technology is necessary to ensure that the solutions proposed are implementable and will realistically contribute to human rights realisation**.

Nathalie Smuha has argued that it is time ‘to move beyond the cry for a human rights-based approach’ to ‘start securing the essential constituents that such framework implores’.⁵⁰ We agree that such work to bridge AI/ML/ADM and disability human rights⁵¹ is necessary, and this must include:

- **close attention to both the contributions and potential limits of adopting a human rights lens;**
- **a cross-sector, interdisciplinary approach** that is informed by human rights law, norms and principles, a detailed understanding of the technology (including its capabilities and limitations), and the experiences, needs and opinions of persons with disabilities. This means drawing on the expertise of persons with disabilities and their representative organisations, experts in disciplines including law and computer science, policy-makers and industry;⁵² and
- **improving public sector accountability, including adequately resourcing relevant institutions**, which will be vital to addressing the risks of automation and other data-driven technologies in the context of disability. This includes developing empowered state-sponsored regulatory frameworks as well as resources for affected people, disabled people’s organisations and other civil society organisations to proactively contribute to enforcement.

We also recommend close attention to the adequacy of self-regulation and co-regulation by developers and users of AI, ML and ADM and the role of producers and users of AI in defining the terms of regulation, in meeting human rights requirements relating to persons with disabilities. Here, we note that most stakeholders who made submissions to the Australian Human Rights Commission’s ‘Human Rights and Technology’ consultation agreed that self-regulation should not be the only form of human rights protection of the tech sector.⁵³

At a minimum, governance would seemingly require **a legal framework that effectively prohibits systems that, by their very nature, will be used to infringe human rights in the disability context**. This could include:

- mandatory, publicly accessible and contestable public interest or human rights impact assessments for forms of automation to determine the appropriate safeguards, including the potential for prohibiting uses that infringe on fundamental rights;

⁴⁹ Although see Walsh and colleagues’ discussion of research and opportunities to include data outliers and ‘change the shape of the bell curve’: Toby Walsh et al, *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve Our Wellbeing* (ACOLA, 2019) 117; on the limits of training data see, e.g., Sushant Kafle et al, ‘Artificial Intelligence Fairness in the Context of Accessibility Research on Intelligent Systems for People Who Are Deaf or Hard of Hearing’ [2020] (125) *ACM SIGACCESS Accessibility and Computing* 4:1; Guo et al (n 6).

⁵⁰ Smuha (n 44) 10.

⁵¹ Latonero (n 45).

⁵² See Marzin (n 43).

⁵³ Australian Human Rights Commission (n 41) 87.

- proportionality testing of any engagement of people’s rights by algorithmic and data-driven practices to ensure those engagements are proportionate to the objective; and
- stronger enforcement of existing data protection and other fundamental rights laws.

The EU’s proposed AI Act offers **an example of a framework that regulates not just the outputs, but also the design, of AI systems**. It focuses on risk, including risks to fundamental rights, and requires *ex ante* compliance checks and ongoing monitoring of AI systems, among other things.⁵⁴

6. Please provide information on the extent to which persons with disabilities, civil society and organizations of persons with disabilities (OPDs) are consulted and participate in the development of policy and practices in relation to the matters outlined in questions 1 to 2 and the extent to which they are included on the same basis as others.

Despite some good examples, there remain serious concerns that **much activity in academia, the market and government have not adopted the standard of active involvement of persons with disabilities** that is becoming more common to many policies and practices concerning disability around the world.⁵⁵

For example, in many public documents celebrating the positive potential of algorithmic and data-driven technologies in mental healthcare, there is a concerning lack of engagement with persons with psychosocial disabilities and their representative organisations.⁵⁶ In 2021 Piers Gooding and Timothy Kariotis conducted a scoping review of all applied studies that used algorithmic and data-driven technologies in ‘online mental health interventions’.⁵⁷ Of the 132 papers that met their selection criteria, only four studies (or three per cent of the field) appeared to involve people who have experienced mental health conditions or psychosocial disability, in the design, evaluation or implementation of the proposals in any substantive way. The studies demonstrated ‘a near-complete exclusion of service users in the conceptualization or development of algorithmic and data-driven technologies and their application to mental health initiatives’.⁵⁸

Exclusion does not define all initiatives in the field, and indeed **there are good examples of technological responses that have been designed with a high level of active input** by people who are viewed as end-users, or those most affected by the technology.⁵⁹ Indeed, there are several data-driven technologies initiated and *led by* persons with disabilities,⁶⁰ including a growing range of ‘digital peer support measures’ around the world.⁶¹ This includes informal initiatives that may not make it into the public spotlight in the same way governments, health practitioners, large NGOs, and private sector actors do. These grassroots initiatives may warrant resources or further research to determine how and why they are working, if indeed they are, and how they can be supported.

⁵⁴ European Commission, *Proposal for a Regulation of the European Parliament and of the Council: Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts* (2021/0106 (COD)) <<https://artificialintelligenceact.eu/the-act/>>.

⁵⁵ Sarah Carr, ‘“AI Gone Mental”: Engagement and Ethics in Data-Driven Technology for Mental Health’ (2020) 0(0) *Journal of Mental Health* 1; Til Wykes, ‘Racing towards a Digital Paradise or a Digital Hell?’ (2019) 28(1) *Journal of Mental Health* 1; Piers Gooding and Timothy Kariotis, ‘A Scoping Review of Algorithmic and Data-Driven Technology in Online Mental Healthcare: What Is Underway and What Place for Ethics and Law?’ 8(6) *JMIR Mental Health* e24668; Aimi Hamraie and Kelly Fritsch, ‘Crip Technoscience Manifesto’ (2019) 5(1) *Catalyst: Feminism, Theory, Technoscience* 1.

⁵⁶ Carr (n 55); Wykes (n 55)

⁵⁷ Gooding and Kariotis (n 55).

⁵⁸ *Ibid.*

⁵⁹ John Torous et al, ‘Creating a Digital Health Smartphone App and Digital Phenotyping Platform for Mental Health and Diverse Healthcare Needs: An Interdisciplinary and Collaborative Approach’ (2019) 4(2) *Journal of Technology in Behavioral Science* 73.

⁶⁰ Hamraie and Fritsch (n 55).

⁶¹ Karen L Fortuna et al, ‘Digital Peer Support Mental Health Interventions for People With a Lived Experience of a Serious Mental Illness: Systematic Review’ (2020) 7(4) *JMIR Mental Health* e16460.

Attention is also required to the **accessibility and quality of consultation of persons with disabilities**. For example, governance in the Australian disability sector in recent years has trended towards promises of ‘more consultation’ and calls for ‘more evidence’ in the process of designing or reforming policy and practice.⁶² However, to give one example, persons with disabilities, their representative organisations and advocates have been providing documented evidence of their experiences, perspectives, concerns, needs, and priorities in relation to the NDIS to government since 2013, with limited policy traction. Over that timeframe, in published responses to inquiries and Royal Commissions, evidence in public hearings, in legal challenges, in the media, in commissioned reviews and public consultations, and in research, there is ample and consistent testimony about issues in the design and implementation of the NDIS from the perspective of potential and current participants, families, carers, advocacy organisations, and other key stakeholders, that has not been reflected in changes to policy and practice. Current public consultation processes throw up barriers to engagement ranging from dense information to rolling calls for responses within imposed parameters in short timelines, and these barriers are compounded where disability intersects with other measures of socio-economic disadvantage. Avenues for influencing law, policy and practice that are open to persons with disabilities can be intimidating, time and resource intensive, and extract a heavy psychological and financial toll with low odds of success. While contractors employed to carry out these consultations are well paid for their time, persons with disabilities, their families and representative organisations rarely are. Representative organisations in particular are poorly resourced, limiting their capacity to engage. These issues are only likely to be exacerbated – and to require targeted action – in relation to AI, ML and ADM given their technical and regulatory complexity.

⁶² Department of Social Services, *engage.dss.gov.au* (Web Page, 2021) <<https://engage.dss.gov.au/>>.