



Portfolio Committee No. 1 – Premier and Finance
Inquiry into Artificial Intelligence (AI) in New South Wales

26 October 2023

SUBMISSION | NEW SOUTH WALES
BAR ASSOCIATION

Promoting the administration of justice

The NSW justice system is built on the principle that justice is best served when a fiercely independent Bar is available and accessible to everyone: to ensure all people can access independent advice and representation, and fearless specialist advocacy, regardless of popularity, belief, fear or favour.

NSW barristers owe their paramount duty to the administration of justice. Our members also owe duties to the Courts, clients, and colleagues.

The Association serves our members and the public by advocating to government, the Courts, the media and community to develop laws and policies that promote the Rule of Law, the public good, the administration of and access to justice.

The New South Wales Bar Association

The Association is a voluntary professional association comprised of more than 2,400 barristers who principally practise in NSW. We also include amongst our members judges, academics, and retired practitioners and judges.

Under our Constitution, the Association is committed to the administration of justice, making recommendations on legislation, law reform and the business and procedure of Courts, and ensuring the benefits of the administration of justice are reasonably and equally available to all members of the community.

If you would like any further information regarding this submission, please contact the Association's Department of Policy and Law Reform on 02 9232 4055.

Contents

A Executive Summary

B Recommendations

C Substantive Submissions

A. Executive Summary

1. The New South Wales Bar Association (the **Association**) thanks Portfolio Committee No. 1 – Premier and Finance for the opportunity to make submissions regarding the Inquiry into Artificial Intelligence (**AI**) in New South Wales (the **Inquiry**).
2. This submission addresses matters raised at (e), (g), (h), (i) and (k) of the Inquiry’s Terms of Reference.
3. The Association’s principal and overarching recommendation is that the NSW Government adopt a structured, evaluative approach to determine whether (and if so, what) legislative change is required to adapt to and regulate AI in NSW. The Association’s recommended approach draws upon aspects of the European Union’s proposed Artificial Intelligence Act, which is still within the legislative process.¹
4. Broadly, the recommended approach is as follows:
 - a. Step 1 – Identify any “prohibited practices”: This involves identifying whether there are any uses of AI that are so clearly undesirable to society that they ought to be prohibited by AI-specific legislation of general application, rather than attempting to address them within legislation that applies to different industries or sectors; and
 - b. Step 2 – Conduct a “regulatory gaps” analysis: After identifying any prohibited practices, consult with industry-, sector- or subject-matter experts to identify particular impacts that are foreseen, hoped-for or feared from the advent of AI, and then engage in a “regulatory gaps analysis” – i.e. a process of assessing whether existing legislation regulating the relevant area is sufficient to address the particular issues identified, and (if not) determine what changes to that legislation are required.
5. This submission makes additional recommendations as to particular AI practices that the Association considers should be prohibited under AI-specific legislation. Further, this submission urges the NSW Department of Communities and Justice to consider the potential for AI technologies to enhance access to justice for court users with disabilities.

¹ Further information about the European Union’s proposed Artificial Intelligence Act is available [here](#).

B. Recommendations

The Association makes the following recommendations.

Recommendation 1: The Overall Approach

That the NSW Government adopt a structured, evaluative approach to determine whether (and if so, what) legislative change is needed in respect of AI. The approach would:

- a. identify whether there are any AI practices that should be prohibited by legislation that addresses AI at a general level (as per Recommendation 2 below); and
- b. undertake a regulatory gap analysis (as per Recommendation 3 below) in consultation with industry-, sector- and subject matter-experts to identify the potential impacts and risks posed by AI, the adequacy of existing legislation to address those risks, and any necessary legislative changes required to address those particular risks.

Recommendation 2: Prohibited AI Practices

Should Recommendation 1 be supported, the Association considers that the following should be prohibited AI practices:

- a. the sale or use of an AI system that deploys subliminal techniques beyond a person's consciousness, or purposefully manipulative or deceptive techniques, in order to materially distort their behaviour in a manner that causes or is likely to cause significant harm to them or another person (other than approved therapeutic techniques used with the informed consent of the person or their legal guardian);
- b. the sale or use of an AI system that exploits any vulnerability of a specific group of persons, for example due to their age, disability, personality traits or social or economic situation, in a manner that causes or is likely to cause significant harm to them or another person;
- c. the sale or use of an AI system by or on behalf of a public authority for the social scoring evaluation or classification of the trustworthiness of natural persons over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics, that leads to detrimental or unfavourable treatment of persons or groups of persons:
 - i. in social contexts that are unrelated to the contexts in which the data was originally generated or collected; or
 - ii. that is unjustified or disproportionate to their social behaviour or its gravity;
- d. the sale or use of an AI system that employs or facilitates facial recognition and other biometric technology in decision-making that has a legal, or similarly significant, effect for individuals, or where there is a high risk to human rights, such as in policing and law enforcement;

- e. the use of an AI system to replace or supplement judicial discretion, including in sentencing, and other evaluative decisions affecting individual liberties and freedoms.

The NSW Government should consider whether the above practices ought to be addressed through AI-specific legislation of general application.

If this Recommendation is supported, there should be a moratorium on such practices until the proposed legislation is developed and enacted.

Recommendation 3: Regulatory Gap Analysis

Should Recommendation 1 be supported, and whether or not Recommendation 2 is supported:

- a. the NSW Government should conduct a regulatory gaps analysis, in consultation with industry-, sector- and subject matter-experts, to determine the application of existing legislation to AI; and
- b. where gaps are identified in specific pieces of legislation, that legislation should be reviewed and modernised to address AI.

Recommendation 4: Using AI to enhance access to justice for people with disabilities

The NSW Department of Communities and Justice should, in consultation with persons with disabilities:

- a. review the use of existing AI and related technologies in NSW courts, services and processes; and
- b. identify additional AI and related technologies that could be implemented to enhance access to justice for court users with disabilities.

C. Substantive submissions

Introduction

6. AI – like other technological developments such as telephones, computers, and the internet – has a potentially broad-reaching effect across many areas of society. A challenge for its regulation is that not only can AI facilitate formerly difficult practices in existing sectors and industries, but it is also likely to create entirely new practices within them, and even create entirely new sectors and industries. The changes that AI is likely to bring to different sectors range from incremental, to transformative, to disruptive.
7. History has shown that regulation often lags the introduction of new and disruptive technology, and that abuses may be committed in the “frontier” period between introduction of a new technology and recognition of the abuses by the public and regulators. Recent examples include intrusions into consumer privacy by new technological platforms over the last 20 years (e.g. collection of information by website operators via cookies, then by social networks and hosted apps such as those operated by Cambridge Analytica, and by software companies and mobile phone operators following the introduction of the smart phone).
8. Given that the details of how AI will affect specific sectors are likely to emerge relatively slowly, there is merit in identifying whether there are practices or outcomes that can be identified now and dealt with at a general level of principle to try to prevent abuse in the frontier period. This is the subject of the Association’s first and second recommendations and is addressed by way of example through case studies on:
 - a. the unregulated use of facial recognition technology;
 - b. the use of AI to replace or supplement judicial decision-making; and
 - c. the Robodebt Scheme.
9. The Association notes that numerous pieces of state and federal legislation will, or may, regulate and interact with the use of AI in specific circumstances. A recent Discussion Paper entitled “Safe and Responsible AI in Australia” issued by the Australian Government Department of Industry, Science and Resources noted that:

Australia’s current approach to date relies on a combination of:

- ***a broad set of general regulations*** that are mainly technology neutral (for example, consumer protection, online safety, privacy and criminal)
- ***sector-specific regulation*** (for example, therapeutic goods, financial services, food safety and motor vehicle safety)

- *voluntary or self-regulation initiatives such as ethical principles for AI that provide guidance to businesses and governments to responsibly design, develop and implement AI.*²

10. The Association’s recommendations reflect the first two bullet points above.
11. The Association’s recommendations are broadly consistent with the recommendations made by the Law Council of Australia (the **Law Council**)³ and the Australian Human Rights Commission (**AHRC**)⁴ in their responses to the Department of Industry, Science and Resources’ June 2023 Discussion Paper, *Safe and Responsible AI in Australia*.
12. Whether or not AI-specific “general regulations” are adopted, the Association considers that there is merit in conducting a “regulatory gaps analysis” to determine the application of existing legislation to AI. Where gaps are identified in specific pieces of legislation, that legislation should be reviewed and modernised. Where legislative gaps are identified and cannot be addressed by existing legislation, AI-specific legislation should be considered to address those gaps, particularly if the gaps appear in multiple, disparate areas.⁵
13. This process should involve consultation with industry-, sector- and subject-matter experts to identify what they currently see or foresee as issues with the use of AI. Examples of regulatory gaps analyses conducted by the Association in the areas of barristers’ professional obligations when considering the use of AI in legal practice and privacy of health records are provided below as case studies 4 and 5, respectively.
14. When considering the Association’s recommendations, a number of important points must be kept in mind.
15. First, the term AI does not have a fixed or defined meaning at a technological level. There are many different pieces of software (or integrated systems) that may be called AI, but they may operate very differently at the technological level. For example, the most popular, current forms of AI are “generative AI” in the form of chatbots (which are statistically-based language models commonly called “pre-trained generative transformers”) and image-generation models (commonly called “diffusion models”).
16. Those two forms of AI are technologically quite different from each other, and are both very different from, for example, AI technologies that are designed to play chess or autonomously drive motor vehicles.

² Australian Government, Department of Industry, Science and Resources Discussion Paper, *Safe and Responsible AI in Australia* (June 2023) at p. 26. Available [here](#).

³ Law Council of Australia, Submission to the Department of Industry, Science and Resources Discussion Paper, *Safe and Responsible AI in Australia*, dated 17 August 2023. Available [here](#).

⁴ Australian Human Rights Commission, Submission to the Department of Industry, Science and Resources June 2023 Discussion Paper, *Safe and Responsible AI in Australia*, dated 26 July 2023. Available [here](#).

⁵ This approach was suggested by the Australian Human Rights Commission in its submission to the Department of Industry, Science and Resources June 2023 Discussion Paper, *Safe and Responsible AI in Australia*, dated 26 July 2023. Available [here](#).

17. One challenge in regulating AI is to ensure that the definition is sufficiently broad to capture the wide variety of existing technologies and new technologies, while not inadvertently capturing software that would not be understood to be operating as an AI. The Association considers that the underlying regulatory approach should be based upon the impact of technology upon society.
18. Secondly, given that AI is not just computer-based, but will invariably be internet-based, the use, and impact, of AI is unlikely to be limited to a particular Australian jurisdiction. Therefore, a consistent response across the states and territories should be pursued where possible, noting also that it is likely that the Federal Parliament will have legislative power over the use of most, if not all, AI systems, and will be integral to the Australian legislative response to AI.⁶
19. Nonetheless, the NSW Parliament will (within constitutional limits) very likely have jurisdiction over AI, particularly at the level where it affects typically state concerns. As such, it is important that NSW has a principled legislative response.
20. Critically, the Association has long advocated for human rights legislation at both the state and federal level. Australia stands alone in the common law world as not having national protections for human rights in the form of a Human Rights Act or a Constitutional Charter of Rights. At present, Australians are particularly vulnerable to the risks and harms associated with AI, particularly during the “frontier” period.

Other relevant Inquiries, Reports and Discussion Papers

Australia

21. The Association acknowledges, and has had regard to, other relevant inquiries, reports and discussion papers produced by organisations that are currently considering policy development and law reform in the AI sphere.
22. In May 2021, the AHRC published a report on Human Rights and Technology.⁷ The AHRC Report made 38 recommendations aimed at:
 - a. promotion of responsible innovation and human rights through measures including regulation, investment and education;
 - b. protecting human rights where AI is used in decision-making and to provide effective accountability for such decisions;
 - c. creation of an AI Safety Commissioner to support regulators, policy makers, government and businesses in applying laws and other standards in respect of AI-informed decision making; and

⁶ See sections 51(v) and (xx) of the [Australian Constitution](#).

⁷ Australian Human Rights Commission, *Human Rights and Technology: Final Report 2021*. Available [here](#).

- d. improving the accessibility of goods, services and facilities that use Digital Communication Technology for people with disabilities.

23. As noted at paragraph [9] above, in June 2023, the Australian Government Department of Industry, Science and Resources released a Discussion Paper entitled “*Safe and Responsible AI in Australia*”. The Discussion Paper included an overview of the domestic and international landscape which will inform the development of governance mechanisms for AI in Australia.⁸ The Association submits that this Inquiry should have regard to any recommendations and responses that arise from the Australian Government’s Discussion Paper.
24. The Law Council made a submission to the Department of Industry, Science and Resources in response to the above Discussion Paper (**Law Council Response**).⁹ In the Law Council’s view, the significant risks posed by the use of AI justify a strengthened and precautionary approach to AI regulation where there is evidence that existing laws and regulations are insufficient to address the issues and harms arising. Further regulation should be multifaceted. It should include the expansion of current legislation and, where necessary, new targeted legislation, not merely a soft law approach of a voluntary code. However, the Law Council considers that the Australian Government should not seek to explicitly regulate AI via a comprehensive “AI Act”.¹⁰

Europe

25. As noted above, the proposed European Union Artificial Intelligence Act is currently within the legislative process. The European Commission first published a proposal for an AI regulation (**EU AI Regulation**) in April 2021.¹¹
26. In December 2022, the Council of the European Union adopted a common position on the legislation (**EU Council Position**).¹² Its approach would narrow the definition of an “AI system” to require a measure of autonomy, to try to ensure that simpler software systems were not inadvertently brought within the legislation’s reach.

⁸ Australian Government, Department of Industry, Science and Resources Discussion Paper, *Safe and Responsible AI in Australia* (June 2023). Available [here](#).

⁹ Law Council of Australia, Submission to the Department of Industry, Science and Resources Discussion Paper, *Safe and Responsible AI in Australia*, dated 17 August 2023, at [125]-[131]. Available [here](#).

¹⁰ Ibid at [3]. Available [here](#).

¹¹ 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, COM (2021) 206 (21 April 2021). Available [here](#).

¹² European Council, Press Release – “Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights” (6 December 2022). Available [here](#).

27. On 14 June 2023, the European Parliament voted to adopt its own negotiating position on the legislation.¹³ Pursuant to the legislative process, there will now be discussions to try to reconcile the three different positions to reach a final, agreed version of the legislation.
28. The European draft legislation was the first significant, proposed legislative approach to AI, and has served as a de facto reference point for discussion of legislation by other jurisdictions.

Other contributors

29. The Human Technology Institute (HTI) at the University of Technology Sydney (UTS) is currently undertaking a research project that explores the question: ‘What is the future of artificial intelligence (AI) regulation for Australia?’¹⁴ This project recognises the exponential growth and influence of AI across several spheres, with recent research conducted by the HTI revealing that two-thirds of Australian organisations are already using or actively planning to use AI systems.
30. The HTI stresses the importance of the law keeping pace with developments in AI, including by encouraging positive innovation while protecting Australians against AI-related harm. The Future of AI Regulation in Australia Project aims to address these challenges and will evaluate the gaps in Australian law and policy with a view to understanding how our regulatory system could be improved to better understand and respond to the risks and benefits posed by AI. The HTI will identify priority reforms and work with civil society, industry and government to identify best-practice regulatory approaches in Australia and abroad. The Association submits that this Inquiry should closely consider the HTI’s final report which will be published in due course.
31. Researchers from the Centre for the Study of Existential Risk at Cambridge are leading international academic consideration of the unusually broad and deep impact of AI on the world, which may prove as ‘transformative’ as the industrial revolution.¹⁵
32. The Centre’s work includes consideration of the underlying rationale for regulatory intervention of AI. According to the Centre’s Matthijs Maas, the question is not only how large particular sociotechnical changes are, but how, or whether, they touch on the general rationales for new regulatory interventions.¹⁶ Maas concludes that there may be certain AI innovations that do not create very large sociotechnical changes; technological change or improvements need not be qualitatively novel, dramatic, sudden, or cutting-edge for them to drive intense and meaningful change in balances

¹³ European Parliament, Press Release – “MEPs ready to negotiate first-ever rules for safe and transparent AI” (14 June 2023). Available [here](#).

¹⁴ More information about the HTI’s Future of AI Regulation in Australia Project can be found [here](#).

¹⁵ Maas, M, Centre for the Study of Existential Risk, ‘Aligning AI Regulation to Sociotechnical Change’, pre-print chapter draft (July 2021) at p. 2. In Bullock, J, Zhang, B, Chen, Y, Himmelreich, J, Young, M, Korinek, A and Hudson, V (eds.). *Oxford Handbook on AI Governance* (Oxford University Press, 2022 forthcoming). Available [here](#).

¹⁶ *Ibid* at p. 7.

of power, or in societal structures: “*We are not worried about technology; we are worried about its effects.*”¹⁷

33. As such, Maas argues that the primary regulatory concern should be with the ‘sociotechnical’ effects that occur, that is, the impact of AI or other sociotechnical changes which involve: (1) new possible market failures; (2) new risks to human health or safety, or the environment; (3) new risks to moral interests, rights, or values; (4) new threats to social solidarity; (5) new threats to democratic processes; (6) new threats directly to the coherence, efficacy or integrity of the existing regulatory ecosystem charged with mitigating the prior risks (1-5).¹⁸

Recommendation 1: Prohibited Practices and Possible AI-Specific Legislation

34. The EU is at the forefront of developments in the regulation of AI. Developments in the EU, including the risk-based approach reflected in the proposed Artificial Intelligence Act, offer helpful insights as to how Australia may address gaps in its own legislative framework. In its present form, the proposed EU Artificial Intelligence Act seeks to ensure that the development and use of AI is aligned with the fundamental rights of individuals.

35. The Association considers that the development and use of AI in Australia should be similarly aligned with the protection of the fundamental rights of individuals.

Prohibited AI Practices

36. A key aspect of the EU AI Regulation is its focus upon “Prohibited Artificial Intelligence Practices”, covered by Title II of the EU Proposal for a Regulation on AI.¹⁹ Those practices are – unsurprisingly – directly influenced by specific human rights provisions derived from European legislation.

37. Although NSW, and indeed the Commonwealth, does not yet have a Human Rights Act, the underlying interests protected by the EU AI Regulation are, in general, rights and interests that are understood to be worthy of protection. Many of these rights are protected in existing legislation (such as the *Privacy Act 1988* (Cth)) and by the *International Covenant on Civil and Political Rights*, which Australia has ratified. This approach is also supported by the Law Council’s submission to the Department of the Prime Minister and Cabinet Issues Paper, *Positioning Australia as a leader in digital economy regulation – Automated decision making and AI regulation*, which identified a range of human

¹⁷ Ibid at p. 8.

¹⁸ Ibid at p. 8.

¹⁹ Title II comprises all those AI systems whose use is considered unacceptable as contravening Union values, for instance by violating fundamental rights.

rights issues and proposed responses with respect to the use of AI,²⁰ and by the AHRC, which recommended the use of AI-specific legislation similar to the EU AI Regulation.²¹

38. It should be noted that the underlying justification for such laws also includes principles beyond strictly human rights norms. The CSIRO and the Gradient Institute²² have jointly developed a set of guidelines and principles to support the responsible use of AI. In broad terms, those principles and guidelines include considerations of: human, societal and environmental wellbeing; human-centred values; fairness; privacy protection and security; reliability and safety; transparency and “explainability”; and contestability and accountability.
39. The importance of generalised prohibitions is that they can ensure that practices that deleteriously affect individuals and society are addressed comprehensively, rather than being limited to particular sectors or industries. They form a core set of norms and rights that are to be protected, and can be expressed at a level of generality to best ensure they are effective. They can also be dealt with in relatively brief terms, allowing AI-specific regulation to be relatively short, and are located in one, central location rather than being addressed and likely repeated in many pieces of industry-, sector- or subject-specific legislation.
40. Finally, it is more likely that national harmonisation and uniformity could be reached by adopting a common piece of AI-specific legislation that operates at a thematic level, and imposes specific prohibitions without reference to particular parts of society. Amendments to state-, territory- or federal-specific legislation, which will reflect the existing content of that legislation, are likely to be more difficult to harmonise, and would need to be dealt with by each jurisdiction separately.

Recommendation 2: Prohibited AI Practices

41. The Association recommends that the following should be considered as prohibited AI practices:
 - a. the sale or use of an AI system that deploys subliminal techniques beyond a person’s consciousness, or purposefully manipulative or deceptive techniques, in order to materially distort their behaviour in a manner that causes or is likely to cause significant harm to them or another person (other than approved therapeutic techniques used with the informed consent of the person or their legal guardian);

²⁰ See Law Council of Australia, Submission to Digital Technology Taskforce, Department of the Prime Minister and Cabinet Issues Paper, *Positioning Australia as a leader in digital economy regulation – Automated decision making and AI regulation* (3 June 2022) at [3]-[7], [143]-[156]. Available [here](#).

²¹ Australian Human Rights Commission, Submission to the Department of Industry, Science and Resources June 2023 Discussion Paper, *Safe and Responsible AI in Australia*, dated 26 July 2023, at [152]-[157]. See especially Recommendations 21 and 22. Available [here](#).

²² Reid, A, O’Callaghan, S and Lu, Y, *Implementing Australia’s AI Ethics Principles: A selection of Responsible AI practices and resources* (2023). Gradient Institute and CSIRO, at p. 42.

- b. the sale or use of an AI system that exploits any vulnerability of a specific group of persons, for example due to their age, disability, personality traits or social or economic situation, in a manner that causes or is likely to cause significant harm to them or another person;
- c. the sale or use of an AI system by or on behalf of a public authority for the social scoring evaluation or classification of the trustworthiness of natural persons over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics, that leads to detrimental or unfavourable treatment of persons or groups of persons:
 - i. in social contexts that are unrelated to the contexts in which the data was originally generated or collected; or
 - ii. that is unjustified or disproportionate to their social behaviour or its gravity;
- d. the sale or use of an AI system that employs or facilitates facial recognition and other biometric technology in decision making that has a legal, or similarly significant, effect for individuals, or where there is a high risk to human rights, such as in policing and law enforcement;
- e. the use of an AI system to replace or supplement judicial discretion, including in sentencing, and other evaluative decisions affecting individual liberties and freedoms.

42. The first four proposed prohibited practices have been adapted from Article 5 of the EU AI Regulation, taking into account the EU Council Position. The fifth proposed prohibited practice is one that has been identified by the Association.

43. The first practice appears to be a natural complement to, and development of, existing protections afforded to consumers for conduct in trade or commerce, as well as participants in markets for financial services. Given that an AI system is not likely to be a “person” under current legislative provisions,²³ and that there may be factual complexity as to whether a person selling or using a system is acting in trade or commerce, the Association considers there is good reason to ensure that an AI technology cannot be lawfully used to deploy subliminal or purposefully manipulative or deceptive techniques designed to produce harmful behavior.

44. This practice was noted in the Department of Industry, Science and Resources June 2023 Discussion Paper, *Safe and Responsible AI in Australia*, as a particular challenge posed by the increased application of AI technologies.²⁴ It includes, but is not limited to, “deepfakes”, which use a form of AI called “deep learning” to generate convincing but entirely fictional images, audio and video content, which can be used for insidious and harmful purposes, and may undermine public trust.²⁵

²³ Compare, for example, section 4 of the *Crimes Act 1900* (NSW) and section 124(1) of the *Corporations Act 2001* (Cth).

²⁴ Australian Government, Department of Industry, Science and Resources Discussion Paper, *Safe and Responsible AI in Australia* (June 2023) at p. 7. Available [here](#).

²⁵ Sample, I, *The Guardian*, ‘What are deepfakes – and how can you spot them?’ (13 January 2020). Available [here](#).

45. The second practice is anchored in a concern for the welfare of vulnerable people. Rather than focusing upon particular industries, contexts or occasions, a blanket prohibition on exploitative conduct should be imposed as an a priori norm of behaviour while any sector- or technology-specific aspects are determined and developed.
46. The Association notes that AI presents particular risks to vulnerable people, including but not limited to:
- a. algorithmic bias and discrimination, which occurs when automated systems contribute to unjustified different treatment or impacts on people based for example, on their race, colour, descent or national or ethnic origin, sexual orientation, gender identity, disability or any other classification protected by law;
 - b. automation bias, which describes the tendency of people to “*disregard or not search for contradictory information in light of a computer-generated solution that is accepted as correct*”. This tendency can be exacerbated in time-critical domains;²⁶ and
 - c. misinformation and disinformation.
47. The Association notes that another reason to enact an a priori prohibition of the practice referred to at paragraph [41](b) is that it is difficult to ascertain the extent to which algorithmic bias and discrimination are already occurring, such that adopting a reactive rather than a proactive approach may result in the practice occurring undetected for a period of time.
48. The third practice, of using AI to engage in or facilitate “social scoring”, is currently a likely candidate for development of AI systems. AI is likely to enable new uses of large datasets (including historical data that is currently infeasible to use) that will produce deleterious results impacting on large sections of society without their knowledge or control.
49. The fourth practice identified as paragraph [41](d) above, has very significant potential impacts. Recommendations 19 and 20 of the AHRC’s *Human Rights and Technology Final Report* urge a moratorium pending the development of legislation to regulate the use of these technologies.²⁷ The HTI notes that such moratoria have been put in place prohibiting the use of facial recognition technology in circumstances where human rights are most at risk in other jurisdictions.²⁸ The Association considers that that this recommendation should extend to all five areas, identified by the Association above, until legislation is developed.

²⁶ Cummings, M, “Automation Bias in Intelligent Time Critical Decision Support Systems” (2004), American Institute of Aeronautics and Astronautics, AIIA 1st Intelligent Systems Technical Conference, 20-22 September 2004, Chicago Illinois.

²⁷ Australian Human Rights Commission, *Human Rights and Technology: Final Report 2021*. Available [here](#).

²⁸ Davis, N, Perry L and Santow, E, *Facial Recognition Technology: Towards a Model Law* (2022) at p. 35, Human Technology Institute, University of Technology Sydney. The HTI also recognises that Australian law does not provide the legal guardrails necessary to ensure that facial recognition technology is developed and deployed in ways that uphold basic human rights, and proposes a Model Law for facial recognition technology. Its report is referred to in the Law Council’s submission to the Department of Industry, Science and Resources’ Discussion Paper, *Safe and Responsible AI in Australia*, at [65]-[67].

50. The fifth practice concerns the potential use of AI systems to supplement, or conceivably replace, the evaluative judgement of a court in determining facts and future outcomes (including, for example, the risk of re-offending when undertaking sentencing), based on an individual’s past behaviour. It is the Association’s view that discretions and evaluative decisions that have been reposed in a judicial officer must be exercised by that officer. Just as they cannot be delegated to another person, the Association considers that they should not be able to be delegated—even *de facto*—to software.²⁹ The Association endorses the Law Council’s position that use of automated decision-making and AI processes in Australian courtrooms or the justice system more broadly should not be supported, particularly where it would limit judicial discretion. The Law Council has observed that automated decision-making processes “*which rely on factors derived from historic data to predict future outcomes raise issues of the entitlement to procedural fairness and for a judgment to be based on the circumstances of the case*”.³⁰ The Association considers that automated decision-making processes are inappropriate for use in relation to decisions which could have a significant effect on individual liberties and freedoms, or substantive rights.
51. The Association recognises that, in the process of developing legislation for the regulation of AI technology, it may become apparent that some qualifications or exceptions to the above prohibited AI practices are considered appropriate. Developments in technology, and in the use of technology, may also require reassessment of the prohibited AI practices from time to time. However, until that wider legislative response can be undertaken, the Association recommends the introduction of the five prohibited AI practices stated above.
52. The Association notes that other possible prohibited practices may be included in AI-specific legislation, and the Association is not suggesting that the five identified practices are the only practices that should be prohibited. Two areas are of note.
53. Firstly, the Association notes that the electronic collection, retention and use of personal information presents key challenges for the protection of privacy of individuals. Those challenges are likely to be amplified by the use of AI. AI may be used to collect such information and/or use it to derive “social credit” scores, which are then used for differential treatment. Other than tightly-controlled and strictly justified situations (for example, to detect victims of child trafficking, people smuggling, or modern slavery; identify missing persons; or locate known criminals or terrorist suspects), use of AI in relation to personal information, or biometric data (including facial recognition) is likely to infringe upon individuals’ right to privacy.
54. There is a partial overlap with this consideration and the third prohibited practice, concerning social scoring. Beyond the use of AI in social scoring, questions arise as to the ability and desirability of

²⁹ The High Court has emphasised the importance of a judge exercising the discretion reposed in him or her in *Wong v The Queen* (2001) 207 CLR 584 at [31], [75]-[78], [85], [139] and [168].

³⁰ Law Council of Australia, Submission to the Department of Industry, Science and Resources Discussion Paper, Safe and Responsible AI in Australia, dated 17 August 2023 at [229].

regulating privacy in AI-specific legislation rather than in existing Commonwealth privacy legislation, and it is beyond the scope of this submission to address that area.

55. Additionally, the Association notes that Title III of the EU AI Regulation addresses “High Risk AI Systems”. These definitions are lengthy and involve evaluative criteria that cannot be usefully addressed in this submission. If Recommendation 1 of this submission is accepted, the Association respectfully submits that the NSW Parliament should consider whether to include “high risk” AI systems in any AI-specific legislation, and specify what those high-risk AI systems might be.

Case study #1: Unregulated Use of Facial Recognition Technology

56. An example of the second AI practice recommended for prohibition at paragraph [41](d) above is the current use of facial recognition technology (FRT) in policing.

57. FRT is a form of AI that extracts spatial and geometric distribution of facial features and converts this information into a digital template. FRT has become increasingly common due to the prevalence of CCTV cameras, the ability to access millions of open-source facial images and the development of technology to utilise this data.

58. However, current FRT is not always accurate or reliable and has the potential to disproportionately impact on vulnerable and minority groups, including already over-policed First Nations people, due to racial biases within algorithms and training data sets.

59. The use of FRT is largely unregulated. While the covert use of listening, data and optical surveillances devices is governed by the *Surveillance Devices Act 2007* (NSW) and the use of telecommunications surveillance is governed by the *Telecommunications (Interception and Access) Act 1979* (Cth), neither of these Acts are sufficient to deal with AI and FRT.

60. Despite the well-known risks of FRT, the Law Council has noted that law enforcement agencies in Australia have reportedly embraced FRT in recent years. An example of this is the use of Clearview AI’s social media-derived database by the Australian Federal Police, as well as police in NSW, Queensland and Victoria.³¹ FRT was reportedly used by NSW Police at the 2023 Sydney Gay and Lesbian Mardi Gras to track participants’ moods as a form of crowd management.³²

61. Without an adequate regulatory framework in place, the risks of FRT are unmitigated. Some of these risks include:

³¹ Law Council of Australia, Submission to the Department of Industry, Science and Resources Discussion Paper, *Safe and Responsible AI in Australia*, dated 17 August 2023 at [63]. Available [here](#).

³² Grubb, B, ‘How your phone and mood will be tracked at Mardi Gras’, *The Sydney Morning Herald* (online, 24 February 2023). Available [here](#).

- a. the disproportionate impact of FRT upon minority groups (especially First Nations communities) due to data set limitations and racial biases within algorithms. The training data sets used to provide the basis for probabilistic reasoning upon algorithmic decision making can result in either an in-built bias or increased error rate with respect to minority groups. Currently, FRT is generally less accurate when identifying women and people from minority racial groups, as compared with other people.³³ The over-representation of First Nations people in the underlying data on interactions with police will result in disproportionate and inaccurate identifications and exacerbate inequality of treatment;
- b. the disproportionate impact of FRT on other vulnerable groups, such as people experiencing homelessness who are more likely to be in public spaces and therefore subject to greater surveillance and more frequent interactions with police than the general population;
- c. unreliable identification by FRT during an investigation may ultimately hinder a successful prosecution. A study into the accuracy of Scotland Yard's FRT found that the technology was verifiably accurate in only 19% of cases. This high error rate can focus the investigation and scarce resources upon the wrong target, whilst the real perpetrator is excluded from the investigation. It can give investigators false confidence in the identity of their suspect, similar to the 'white coat effect' upon jurors. Indeed, the High Court of Australia in *Honeysett v R* [2014] HCA 29 at [45] rejected the purported expert evidence of 'forensic identification' by means of comparison of footage from the crime scene with images known to be of the accused;
- d. the high error rates and inability of FRT to positively identify a person severely diminishes the usefulness of the technology. FRT does not positively identify a person, rather it provides a list of possible matches based on the algorithm and it will be for the human expert to determine a match or otherwise. In circumstances where the High Court of Australia has firmly rejected this type of identification evidence, as nothing more than subjective belief or unsupported speculation,³⁴ it appears that neither the technology nor the human interface with the technology is sufficiently reliable to justify its use as evidence in a criminal trial; and
- e. the potential for FRT to be combined with other data driven technologies can pose a serious risk to fundamental social conditions that make the protection of individual rights

³³ See, for example, Buolamwini, J and Guru, T, 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification' (2018) 81 *Proceedings of Machine Learning Research* 1 (available [here](#)); Krishnapriya, K, Vangara, K, King, M, Albiero, V and Bowyer, K, 'Characterizing the Variability in Face Recognition Accuracy Relative to Race' (Conference Paper, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2019) (available [here](#)); Raji, I and Buolamwini, J, 'Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products' (Conference on Artificial Intelligence, Ethics, and Society, 2019) (available [here](#)).

³⁴ *Honeysett v R* [2014] HCA 29 [23].

possible. The combination of FRT with relatively innocuous data sets such as financial records from a shop, traffic cameras and vehicle registration, to name but a few, can quickly reveal detailed and highly personal information about an individual. This is known as the ‘mosaic effect’ and is prevalent in China.³⁵

62. The AHRC’s *Human Rights and Technology Final Report* recommended a moratorium on the use of biometric technologies, including facial recognition, in high-risk areas of decision making.³⁶ The Association supports that recommendation. The significant risks to individuals posed by FRT when used as part of law enforcement, including in discretionary policing practices, outweighs the potential perceived benefits of such technology for criminal investigation purposes. The Association considers that the use of FRT for criminal investigation and law enforcement purposes should be prohibited until a legislated framework is developed to ensure its use is accurate, transparent, effective and proportionate.

Case study #2: Use of AI to replace or supplement judicial decision-making

63. An example of the fifth practice recommended for prohibition at paragraph [41](e) and Recommendation 2(e) above is the potential use of AI and related technology in criminal and quasi-criminal proceedings, including the use of actuarial risk assessment tools in sentence proceedings and proceedings for post sentence detention or supervision in NSW.

64. A significant aspect of judicial decision-making, particularly in a criminal context, is the determination of factual disputes. The determination of these factual disputes, in accordance with legal principles and reasoning, can have a significant bearing on the ultimate judicial determination. Less than full disclosure of the judicial reasoning process is likely to damage public confidence in the administration of justice.³⁷

65. Judicial officers hearing criminal matters are also called upon to engage in ‘risk assessment’ on a daily basis. Traditionally this assessment is applied in sentencing and bail determinations, where the risk of an individual’s offending is one factor amongst many to be taken into consideration.

³⁵ See, for example, Pozen, D, ‘The Mosaic Theory, National Security, and the Freedom of Information Act’ (2005) 115 *Yale Law Journal* 628 (available [here](#)); Australian Human Rights Commission, Department of Industry, Science and Resources Discussion Paper, *Safe and Responsible AI in Australia* (26 July 2023) 44 (available [here](#)); Human Rights Watch, ‘China’s Algorithms of Repression: Reverse Engineering a Xingjian Police Mass Surveillance App’ (Report, 1 May 2019) (available [here](#)).

³⁶ Australian Human Rights Commission, Submission to the Department of Industry, Science and Resources Discussion Paper, *Safe and Responsible AI in Australia*, dated 26 July 2023. Recommendation 40 (available [here](#)).

³⁷ See the Hon Sir Anthony Mason AC KBE, ‘The Nature of the Judicial Process and Judicial Decision Making’ in Sheard, R (ed), *A Matter of Judgment: Judicial Decision-making and Judgment Writing* (Judicial Commission of NSW, 2003) at pp. 1-2 (available [here](#)).

66. The Association considers that the use of AI and related technology in prediction of risk of future offending behavior should be prohibited given the significant impact on fundamental human rights.³⁸
67. Where AI and related technology is used as part of evidence relied on in sentence or post sentence order proceedings, it is often unclear what input has been delivered into the AI system and what role an individual factor might have had in the overall determination. This is particularly problematic as, unless a judge is controlling the inputs into a particular AI tool (which creates its own problems), the tool will always be reflective of the input of the user. For example, if an offender is said to have made a pledge to a terrorist organisation, that information will clearly have a significant role to play in an overall assessment of risk for the purpose of sentence or post sentence order determinations. However, if the evidence of this pledge only comes from remote hearsay evidence, a judicial officer properly applying their task will give it little weight. An AI tool may not take the same approach.
68. Further, reliance on AI in judicial decision-making risks importing a discretion into the decision-making process which is not that of the court, but instead a combination of the discretion of the user of the AI tool and the maker of the AI tool. For example, many AI tools can require the user to ‘score’ an offender from low-medium-high on the presence of a particular factor. However, what one user might assess as medium might clearly be different to another and may altogether be different to the weight attributed by the court to the specific factor. Inherent in any output of the AI is the user’s characterisation of certain input factors.
69. Additionally, there are serious concerns about the presence of bias within AI risk assessment tools. In 2016, ProPublica conducted a study of risk assessment tools and underlying algorithms used in certain states in the United States of America and found that the algorithms made mistakes with black and white defendants at different rates. The formula was particularly likely to falsely flag black defendants as future criminals; wrongly labelling them this way at almost twice the rate as white defendants. White defendants were mislabeled as low-risk more often than black defendants.³⁹
70. Finally, AI tools often lack transparency by relying on a ‘black box’ system in which the internal workings are invisible and impenetrable – meaning the reasoning, coding and weighting of different factors are completely unknown. This is antithetical to transparent judicial decision-

³⁸ The use of AI systems in the justice system has been determined in the European Union model of regulation to be “high-risk” and carries the most stringent level of regulation of permissible AI: 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, Art 6(2) Annex III [7] (available [here](#)). The Law Council has expressed similar concern that, if used, automatic decision-making (ADM) processes which rely on factors derived from historic data to predict future outcomes would raise serious issues for the entitlement to due process and compromise judgments predicated on the overall circumstances of the case. The Association agrees that ADM processes are inappropriate for use in relation to decisions which could have a significant effect on individual liberties and freedoms: Law Council of Australia, Submission to the Department of Industry, Science and Resources, *Safe and Responsible AI in Australia*, dated 17 August 2023 at p. [168] (available [here](#)).

³⁹ Angwin, J et al ‘Machine Bias’, *ProPublica* (online, 23 May 2016) (available [here](#)).

making and open justice principles whereby courts are expected to clearly demonstrate their reasoning process.

71. The above concerns have been highlighted in recent criticism by the Independent National Security Legislation Monitor of the Violent Extremism Risk Assessment Version 2 Revision (VERA-2R), a risk assessment tool explicitly provided for by the Terrorism (High Risk Offenders) Regulation 2018 (NSW). Under the *Terrorism (High Risk Offenders) Act 2017* (NSW) (THRO Act), the court must consider its results. VERA-2R is a tool applied by psychologists and psychiatrists engaged by the State in Risk Assessment Reports under the THRO Act. It is a form of algorithmic risk assessment tools to make determinations of risk in post-sentence order proceedings in relation to ‘high risk’ sex, violence and terrorism offenders.⁴⁰
72. The reliability, validity and equity of VERA-2R has been called into question by ANU academics Dr Emily Corner and Dr Helen Taylor, who have found that it lacks a strong theoretical and empirical foundation, has poor inter-rater reliability, and poor predictive validity no better than chance.⁴¹ These findings highlight the risks of courts and governments placing an over-reliance on algorithmic tools without a proper understanding, analysis and testing of the tools. The use of AI in such processes would magnify these risks due to the difficulty of identifying what an AI system does or does not consider, whether its decision is biased due to skewed underlying data, and the lack of transparency and accountability in such decisions.

Case study #3: The Robodebt Scheme

73. The notorious “Robodebt Scheme” is a recent reminder of the potential consequences of unregulated and unchecked use of AI-powered algorithmic and automated decision-making processes in the administrative context. The Robodebt Scheme provides strong support for the Association’s recommendation that certain AI practices should be prohibited through principled-based, AI-specific legislation.
74. The Robodebt Scheme was an automated debt assessment and recovery scheme employed by the Australian Government between 2015 and 2019 as part of its Centrelink payment

⁴⁰ These ‘tools’ are a form of actuarial risk assessment instruments whereby groups of offenders are evaluated on objective characteristics (e.g. age, marital status, history of offending, type of victim, relationship to victim); the files of offenders who have been released and whose reconviction status is known are evaluated. This information is subject to statistical analyses (e.g. logistic regression), and then risk groups characterised by different rates of reconviction (e.g. low, medium, high and very high) are constructed. This information is then used to generate a prediction about a new individual – the subject of concern. It is argued that their risk will mimic the average risk of the group they fall within: see Donaldson, G, *Review into Division 105A (and related provisions) of the Criminal Code Act 1995 (Cth)* (Independent National Security Legislation Monitor Report, 4th ed, 3 March 2023) at [251] (available [here](#)).

⁴¹ Corner, E and Taylor, H, ‘Testing the Reliability, Validity, and Equity of Terrorism Risk Assessment Instruments’ (Centre for Social Research and Methods, Australian National University) (available [here](#)). The use of VERA-2R in Commonwealth Post Sentence Order Proceedings is outlined in Donaldson, G, *Review into Division 105A (and related provisions) of the Criminal Code Act 1995 (Cth)* (Independent National Security Legislation Monitor Report, 4th ed, 3 March 2023) Ch 7 (available [here](#)).

compliance program. The Scheme aimed to replace the formerly manual system of calculating overpayments and the discretionary issuing of debt notices to welfare recipients with an automated data-matching process that compared Centrelink records with income data from the Australian Taxation Office which was averaged using an algorithm. However, flaws in the algorithm used for the income averaging process resulted in hundreds of thousands of welfare recipients being wrongly accused of owing money to Centrelink and being forced to enter into repayment arrangements.

75. In its Report,⁴² the Royal Commission into the Robodebt Scheme described it as “*a crude and cruel mechanism, neither fair nor legal*” and “*a costly failure of public administration, in both human and economic terms.*”⁴³ The Royal Commission made a number of recommendations relating to the automated decision-making process used in the Scheme, including Recommendation 17.1, which recommended that the Commonwealth consider legislative reform to introduce a consistent legal framework in which automation in government services can operate and which enshrines principles of transparency of automated decision-making processes, availability of information and a right of review for those affected, and scope for independent expert scrutiny.
76. The Association strongly endorses this recommendation, which broadly reflects the OECD AI Principles that AI actors should respect the rule of law, human rights and democratic values throughout the AI system lifecycle and commit to transparency and responsible disclosure regarding AI systems.⁴⁴ More specifically, Recommendation 17.1 and indeed many of the Royal Commission’s findings support the Association’s case for overarching AI-specific legislation. In particular, had a principled-based, AI-specific Act prohibiting certain AI practices existed prior to the introduction of the Robodebt Scheme, it is likely that the AI methods and technologies employed throughout the Scheme would have been outlawed for failure to comply with fundamental requirements relating to transparency, privacy and fairness, and the disastrous consequences of the Scheme could have been avoided.

Definition of AI system

77. As noted above, the definition of an AI system is central to the operation of any AI-specific legislation, but also involves significant technical and technological expertise. Although the Association recognises that any drafting process would require significant consideration and consultation, it is nonetheless useful to note the ambit of the definition in the EU AI Regulation, as modified by the EU Council Position, so as to inform an understanding of the scope of operation of the prohibited practices in Recommendation 2.
78. That definition is:

⁴² *Report: Royal Commission into the Robodebt Scheme*, handed down 7 July 2023. Available [here](#).

⁴³ *Ibid* Vol 1, p. xxix.

⁴⁴ OECD AI Principles, Principles 1.2 and 1.3. Available [here](#).

‘artificial intelligence system’ (AI system) means a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic and knowledge based approaches, and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts.

79. The Association is of the view that consideration should be given to terms used by providers, operators or vendors of systems, such that if they describe, promote or market a system as using AI (or a cognate term), this would prima facie bring that system within the definition of an “AI system”.
80. This would tend to ensure that if such a person considered it accurate to describe their system as an AI system (particularly for the purposes of obtaining any commercial benefit), and it would not otherwise fall within the definition, then that system would be deemed to fall within the definition and thereby regulated. This would avoid the possibility that a technological system which the market saw fit to describe as utilising AI would not be caught if the legislative definition were too narrow. It would also allow the definition to be expanded (but not contracted) by the market, to catch technological developments that were not anticipated, or adequately captured, when the legislation was drafted.

Recommendation 3: Regulatory Gap Analysis

81. If Recommendation 1 above is supported, the Association recommends that the NSW Government conduct a regulatory gap analysis. This would involve: (a) determining the application of existing legislation to AI; and (b) where gaps are identified in specific pieces of legislation, reviewing and modernising that legislation to address AI. This approach was recommended by the AHRC in both its 2020 Report and its recent submission to the Department of Industry, Science and Resources in response to the Discussion Paper, *Safe and Responsible AI in Australia*.⁴⁵
82. The following two case studies are examples of preliminary regulatory gap analyses conducted by the Association.

Case study #4: Barristers’ professional obligations when considering the use of AI in their legal practice

83. The Association has already conducted a regulatory gap analysis for the use of AI by NSW barristers. This process resulted in the publication on 12 July 2023 of Guidelines on “Issues

⁴⁵ Australian Human Rights Commission, “Australia Needs AI Regulation” (16 August 2023). Available [here](#). This method was also discussed in the HTI’s submission to the Department of Industry, Science and Resources Discussion Paper, *Safe and Responsible AI in Australia* (June 2023) referred to above, which makes six recommendations including: Recommendation 3: *The Australian Government should undertake a legal gap analysis, focused on areas where AI presents an especially significant risk of harm. The Australian Government should prioritise reform that addresses those risks;* and Recommendation 6: *Australia should adopt framework legislation for AI (an Australian AI Act)...*

Arising from the Use of AI Language Models (including ChatGPT) in Legal Practice”,⁴⁶ and provides a useful case study.

84. The first step was to identify the existing applicable laws and regulations. The obligations upon NSW barristers are likely to be the same as barristers in other Uniform Law jurisdictions.⁴⁷ The legislation deals with disciplinary matters, including definitions for professional misconduct (section 297) and unsatisfactory professional conduct (section 296). Rules specifically governing barristers are made as delegated legislation, and contravention of those rules can constitute professional misconduct or unsatisfactory professional conduct.⁴⁸ In NSW, those rules are the *Legal Profession Uniform Conduct (Barristers) Rules 2015* (NSW) (**Barristers Rules**).
85. The Association is a “local regulatory authority” for NSW barristers, and is able to issue guidelines to barristers to assist them in complying with their ethical obligations, as well as other requirements under general law. These guidelines can be taken into account when considering a complaint against a barrister and thus assist in identifying minimum standards expected of barristers, as well as proscribed conduct.
86. The next step was to identify any potential gaps in the regulatory framework, particularly in relation to known limitations of AI language models (including ChatGPT) which may have serious and significant implications in the context of legal professional practice. These included but were not limited to:
- a. overestimation and incompleteness of machine-generated answers to research queries due to the finite body of information on which GPTs are trained;
 - b. currency of information;
 - c. ‘hallucination’, being the fabrication of facts or sources by GPTs when they do not have access to sufficient data to produce an answer;
 - d. reproduction of any bias contained within the data set on which the GPT was trained; and
 - e. “unexplainability” and lack of verifiability.
87. The NSW Bar Council considered that, whilst no legislative changes were necessary at this time, the identified risks should be dealt with through the development and publication of guidelines to the Association’s membership. Amongst other things, the “AI and GPT Guidelines for Barristers” advised members of the limitations identified above and the risks of reliance on machine-generated answers and, most importantly, reminded members of their professional conduct and ethical obligations – such as those relating to the maintenance of independence and integrity, the provision of competent and diligent representation, and the preservation of client

⁴⁶ Available here: [NSW Bar Association AI and GPT Guidelines for Barristers \(12 July 2023\)](#).

⁴⁷ The *Legal Profession Uniform Law 2014* (NSW) is given statutory force by the *Legal Profession Uniform Law Application Act 2014* (NSW).

⁴⁸ Office of the NSW Legal Services Commissioner, Webpage: ‘Types of complaints’, updated 2022. Available [here](#).

confidentiality – which must be considered before, and not breached in the event of, use of AI technologies in legal practice.

Case study #5: Privacy of Health Records

88. The Association has conducted a preliminary regulatory gap analysis of the legislative framework for managing privacy concerns in relation to individuals' health information. The Association acknowledges the benefits of increased access to efficient, cost-effective, quality medical care enabled by advances in AI in recent years.⁴⁹ In that context, the Association asked whether the existing privacy framework in NSW appropriately responded to the heightened risks to the privacy and security of individuals' health information posed by the expansion and increased use of AI technologies in the healthcare sector.
89. The first step was to identify existing applicable laws and regulations. In NSW, the storage of health records is regulated through the *Health Records and Information Privacy Act 2002* (NSW) (**HRIP Act**). The HRIP Act imposes, on organisations which provide health services or collect, hold or use health information, an obligation to comply with the Health Privacy Principles set out in Schedule 1 of the Act.⁵⁰
90. The second step was to identify potential regulatory gaps. A preliminary analysis conducted by the Association revealed a number of potential regulatory gaps in relation to privacy.⁵¹ These included, but were not limited to:
- a. gaps in the regulation of organisations that use medical devices that involve AI for therapeutic purposes such as diagnosis, disease prevention, monitoring, prognosis and treatment, but who do not provide a 'health service' as defined;
 - b. gaps in the regulation of information sharing between third party organisations who provide therapeutic services which involve AI technologies and 'health service providers';⁵² and
 - c. gaps in the regulation of the use of health information for a secondary purpose (such as training or research) in circumstances where de-identifying practices may not appropriately or adequately address the heightened risk of linkage of de-identified data across registries posed by the growth in sophistication of AI algorithms and the prevalence of genetic therapies.

⁴⁹ For example, machine learning models have been developed to identify suspicious skin lesions, detect markers for Alzheimer's disease, and automate perfusion mapping to predict coronary artery disease from imaging. For more information, see [here](#).

⁵⁰ Refer to the definitions in section 4 of the Act and the HPPs in Schedule 1.

⁵¹ Potential gaps were identified in the definition of 'health service' and in HPP 4, HPP 10 and HPP 14.

⁵² For example, HPP 4, which requires an individual to be made aware of certain matters, could exclude from its application third-party sub-contractors who use AI to provide therapeutic health services but do not 'collect' (as distinct from 'hold' or 'use') health information.

91. The third and final step was to identify what, if any, legislative changes were necessary. From its preliminary analysis, the Association concluded that the existing regulatory framework was likely insufficient to address identified concerns regarding privacy protection of individuals' health information emerging as a result of AI technologies, and required modernisation and amendment.

92. This is a preliminary analysis only. The Association acknowledges the need for consultative input from privacy and health sector experts in order to conduct a proper and fulsome analysis to inform any legislative reform.

93. We note that the above case studies are only two examples of regulatory gaps analyses that the Association recommends. There are, of course, many other legislative fields which are ripe for similar analyses. This includes, but is not limited to, the impact of AI on copyright laws and intellectual property.⁵³ The Association acknowledges that impacts upon intellectual property are likely to be dealt with at a federal level.

Recommendation 4: Using AI to enhance access to justice for people with disabilities

94. The case studies set out above support the Association's overall position in favour of a cautious approach and careful consideration of the risks posed by AI technologies. Nonetheless, the potential benefits of AI to enable the equal enjoyment of rights by persons with disabilities are significant. In particular, the appropriate use of AI can enhance access to and effective participation in the justice system by people with disabilities, in fulfillment of their rights under the *International Covenant on Civil and Political Rights* and the *Convention on the Rights of Persons with Disabilities*.⁵⁴

95. The use of AI has the potential to drive inclusive equality by promoting greater independence and dignity, making justice more affordable and accessible, as well as better function for people with disabilities.⁵⁵ In the courtroom context, for example, AI technologies can have significant advantages in terms of improving inclusion, equity and accessibility for court participants and increasing access to justice for vulnerable groups, including people with disabilities.

⁵³ There is extensive 'scraping' of copyright material from online sources without permission. While this input of unlicensed material is regulated by copyright legislation, in practice it is difficult to identify and prove infringement and to protect and enforce copyright rights. In addition, copyright material is being used to generate new works without acknowledgement of or compensation to the owners of the works used, including on a commercial basis. the *Copyright Act 1968* (Cth) has technology neutral provisions which do not adequately deal with AI.

⁵⁴ ICCPR, Article 26 (right to equality before the law, and equal protection of the law, protecting against discrimination). Article 9 of the CRPD requires States Parties to take appropriate measures to identify and eliminate barriers to people with disability accessing technology, including by implementing minimum accessibility standards, and promoting the design and development of accessible technology. Article 3(c) refers to the principle of '[f]ull and effective participation and inclusion in society' of people with disabilities.

⁵⁵ CSIRO, *Framework for Artificial Intelligence-enabled Assistive Technology as Supports under the National Disability Insurance Scheme: Final Report* (May 2022). Available [here](#).

96. As a matter of principle, all those involved in court proceedings (whether as legal practitioners, litigants, witnesses, etc.) ought to enjoy equal access to, and use of, court facilities, services and processes. Accessibility needs may arise in respect of physical, mental, neurodevelopmental, sensory or other differences between individuals.
97. In its recently published Final Report of the Royal Commission into Violence, Abuse, Neglect and Exploitation of People with Disability (**the Disability Royal Commission**), the Disability Royal Commission made a number of recommendations aimed at increasing the autonomy and participation of people with disabilities in decisions that affect them,⁵⁶ as well as recommending reforms to better enable the autonomy of, and access for, people with disabilities. This includes promoting accessible information and communications. The Disability Royal Commission also recommended that a new Disability Rights Act include a positive duty on Government to promote disability equality and inclusion, to address barriers that disadvantage people with disability, including barriers compounded by a person with disability's combination of attributes and experiences, and to promote accessibility and universal design, and appropriate remedial action to existing infrastructure (Recommendation 4.12).
98. While the Disability Royal Commission Final Report does not directly address the use of AI in the legal system, the Association considers that the development of technologies that integrate AI can have real benefits in terms of improving accessibility for persons who interact and engage with that system. By way of example:
- a. Real-time captioning/translation programs and advancements in lip-reading recognition programs can assist those with impaired hearing to participate more effectively in virtual court hearings conducted via AVL and web conferences (e.g. via Zoom or Microsoft Teams);
 - b. AI-based image/facial recognition technologies can assist those with visual impairments to read documents, identify people and surroundings, detect facial expressions etc.⁵⁷ Similarly, AI programs which automatically convert images into text alternatives or provide more sophisticated text-to-audio services can assist those with visual impairments to view images or read text; and
 - c. AI technologies such as ChatGPT can summarise large amounts of text virtually instantaneously to assist those with cognitive differences to comprehend and process written information.⁵⁸

⁵⁶ The Commission recommended that the Australian Government and state and territory governments should develop and agree on an Associated Plan in connection with *Australia's Disability Strategy 2021–2031* to improve the accessibility of information and communications for people with disabilities. The Associated Plan should be co-designed with people with disabilities and their representative organisations. It should be finalised by the end of 2024 (Recommendation 6.1).

⁵⁷ See recommendation of the Victorian Law Reform Commission in its recently published report, *Inclusive Juries: Access for People Who are Deaf, Hard of Hearing, Blind or Have Low Vision*, available [here](#). Implementing accessible technology into juries allows people with disabilities to fully participate with all aspects of life (such as civic duty), as protected under Article 9 the CRPD.

⁵⁸ Examples of how this technology can assist people with disabilities is available [here](#).

99. There are also possibilities for the extension of AI-based programs into live court proceedings. In particular, the Association submits that the ability to generate reasonably accurate live transcripts of proceedings using software would enhance the ability of court participants with diverse accessibility needs to participate in hearings. Such a technological approach may also be more efficient, and less expensive than reliance on human operators. It would require consideration being given to the relevant provisions of the *Court Security Act 2005* (NSW) dealing with the recording of proceedings, and how any such transcript system would be implemented (e.g. whether it would be provided by the court or a question of a practitioner or participant using an authorised program). This would similarly be the case with any new technology that interpreted audio or visual recordings into a new format.
100. More generally, AI-assisted technologies may have the potential to advance or streamline aspects of legal work, such as legal research. It is important, as with any new technology, that the adoption of new approaches take into account the fact that accessibility needs may arise with their use. The use of computers is a good example. Computers and text display have been the workhorse of the legal profession for decades, but minimum accessibility standards have only recently been introduced and enforced. Apart from the ethical considerations around the use of new technologies identified in paragraph [86] above, the potential use of AI by participants in the legal system should be developed and introduced with accessibility in mind.⁵⁹
101. Building on the recommendations of both the AHRC and the Disability Royal Commission, there is a significant opportunity for the NSW Government to develop AI assistive technology to better facilitate access to courts and tribunal processes for people with disabilities. This would be consistent with Australia's AI Ethics Principles, which aim for AI systems to respect human rights, diversity and the autonomy of individuals,⁶⁰ as well as broader findings and recommendations of the Disability Royal Commission and international human rights obligations referred to above.
102. There is limited information publicly available as to the current use and availability of AI and related technology in NSW courts. For this reason, the Association recommends that, as a starting point, the NSW Government should scope the availability and potential future use of AI assistive technologies to assist people with disabilities to participate fully and effectively in court proceedings in NSW in fulfillment of their right to equality before the law.⁶¹ This scoping exercise should be undertaken in consultation with persons with disabilities, including legal practitioners with disabilities, and should entail: (i) a review the use of existing AI and related technologies in NSW court facilities, services and processes; and (ii) identification of additional AI and related technologies that could enhance access to justice for people with disabilities.

⁵⁹ See generally, Australian Human Rights Commission, *Human Rights and Technology: Final Report 2021*, 'Part D: Accessible Technology'. Available [here](#).

⁶⁰ To view Australia's AI Ethics Principles, please see [here](#).

⁶¹ See Castan, M, Olivares Jones, A and Walker, S, *Artificial Intelligence and the Rights of Persons with Disabilities – Submission to the Special Rapporteur on the rights of persons with disabilities* (November 2021) at p. 26.

103. The Association agrees with the Law Council that AI used to automate transcription services can likely be used in a courtroom, but reiterates its submission that AI should not be used where it would limit or replace judicial discretion.⁶² The Association otherwise urges caution and consultation before other AI processes are used in Australian courtrooms or the justice system more broadly, even though it seems clear that there may be potential benefits identified above.

⁶² Law Council of Australia, Submission to the Department of Industry, Science and Resources Discussion Paper, *Safe and Responsible AI in Australia*, dated 17 August 2023 at [168]. Available [here](#).