# Artificial Intelligence: Australia’s Ethics Framework

Department of Industry, Innovation and Science

28 June 2019

*Telephone* +61 2 6246 3788 • *Fax* +61 2 6248 0639
*Email* mail@lawcouncil.asn.au
GPO Box 1989, Canberra ACT 2601, DX 5719 Canberra
19 Torrens St Braddon ACT 2612
Law Council of Australia Limited ABN 85 005 260 622
*www.lawcouncil.asn.au*

# Table of Contents

**About the Law Council of Australia ... 4**

**Acknowledgement ... 5**

**Executive Summary ... 6**

Introduction ... 7

Question 1: are the principles put forward in the discussion paper the right ones? Is anything missing? ... 7

Question 2: Do the principles put forward in the discussion paper sufficiently reflect the values of the Australian public? ... 7

Question 5: What other tools or support mechanisms would you need to be able to implement principles for ethical AI? ... 7

Question 7: Are there additional ethical issues related to AI that have not been raised in the discussion paper? What are they and why are they important? ... 7

General observations: ... 7

**Introduction ... 9**

**Question 1: are the principles put forward in the discussion paper the right ones?  Is anything missing? ... 11**

Principle 1: generates net-benefits ... 13

Principle 2: doing no harm ... 14

Principle 3: regulatory and legal compliance. ... 16

Importance of enforcement and potential models ... 16

Compliance mechanisms ... 17

Principle 4: privacy protection ... 19

Legislative framework ... 20

Consistent terminology ... 20

Consent process is not fundamental to protecting privacy ... 22

Right to be forgotten ... 23

Protecting privacy extends beyond confidentiality and security ... 23

Eligible data breach ... 25

De-identification ... 25

Other legislative initiatives ... 25

Importance and benefit of a charter of rights ... 26

Principle 5: fairness ... 27

Principle 6: transparency and explainability ... 27

Principle 7: contestability ... 27

Principle 8: accountability ... 28

**Question 2: Do the principles put forward in the discussion paper sufficiently reflect the values of the Australian public? ... 29**

Values-based vs rights-based approach to AI ethics ... 29

Workers protections ... 30

**Question 5: What other tools or support mechanisms would you need to be able to implement principles for ethical AI? ....30**

Additional tools....30

Responsible Innovation Organisation ....31

**Question 6: Are there already best-practice models that you know of in related fields that can serve as a template to follow in the practical application of ethical AI?....32**

**Question 7: Are there additional ethical issues related to AI that have not been raised in the discussion paper? What are they and why are they important? ....33**

Administrative law principles ....33

Resolving tensions between the Principles....34

**General observations ....34**

Role of ethicists and other specialists in developing an AI ethics framework ....34

Significant and insignificant risk....35

International initiatives to build a global framework....35

Data Governance ....35

# About the Law Council of Australia

The Law Council of Australia exists to represent the legal profession at the national level, to speak on behalf of its Constituent Bodies on national issues, and to promote the administration of justice, access to justice and general improvement of the law.

The Law Council advises governments, courts and federal agencies on ways in which the law and the justice system can be improved for the benefit of the community. The Law Council also represents the Australian legal profession overseas, and maintains close relationships with legal professional bodies throughout the world.

The Law Council was established in 1933, and represents 16 Australian State and Territory law societies and bar associations and the Law Firms Australia, which are known collectively as the Council's Constituent Bodies. The Law Council's Constituent Bodies are:

- Australian Capital Territory Bar Association
- Australian Capital Territory Law Society
- Bar Association of Queensland Inc
- Law Institute of Victoria
- Law Society of New South Wales
- Law Society of South Australia
- Law Society of Tasmania
- Law Society Northern Territory
- Law Society of Western Australia
- New South Wales Bar Association
- Northern Territory Bar Association
- Queensland Law Society
- South Australian Bar Association
- Tasmanian Bar
- Law Firms Australia
- The Victorian Bar Inc
- Western Australian Bar Association

Through this representation, the Law Council effectively acts on behalf of more than 60,000 lawyers across Australia.

The Law Council is governed by a board of 23 Directors – one from each of the constituent bodies and six elected Executive members. The Directors meet quarterly to set objectives, policy and priorities for the Law Council. Between the meetings of Directors, policies and governance responsibility for the Law Council is exercised by the elected Executive members, led by the President who normally serves a 12-month term. The Council's six Executive members are nominated and elected by the board of Directors.

Members of the 2019 Executive as at 28 June 2019 are:

- Mr Arthur Moses SC, President
- President-elect (vacant)
- Ms Pauline Wright, Treasurer
- Mr Tass Liveris, Executive Member
- Dr Jacoba Brasch QC, Executive Member
- Mr Ross Drinnan, Executive Member

The Secretariat serves the Law Council nationally and is based in Canberra.

# Acknowledgement

The Law Council is grateful to the Law Institute of Victoria (**LIV**), the Law Society of New South Wales (**LSNSW**) and the Law Society of Western Australia (**LSWA**) for their assistance with the preparation of this submission, as well as input from its Administrative Law Committee, Business and Human Rights Committee, Migration Law Committee, National Human Rights Committee, Privacy Law Committee and Professional Ethics Committee.

# Executive Summary

1. The Law Council of Australia welcomes the consultation of the Department of Industry, Innovation and Science (**DIIS**) into Artificial Intelligence: Australia's Ethics Framework (**Consultation**), and it is pleased to make a submission in response to the Data61 CSIRO discussion paper *Artificial Intelligence: Australia's Ethics Framework*[1] (**Discussion Paper**).

2. The Law Council believes that new and evolving technologies, including artificial intelligence (**AI**), machine learning and other forms of automated decision-making offer important benefits, including the potential to contribute to strengthening the economy, increasing the cohesion and inclusiveness of society, supporting sustainability and the efficient use of resources, and increasing human wellbeing. However a number of significant risks and challenges are also present, which it is necessary, and timely, to discuss.

3. The Discussion Paper explores a number of these issues, and proposes an ethical framework comprising eight core principles for AI, as well as a toolkit of processes, safeguards and resources to support the implementation of that framework. Key issues raised by the Law Council in responding to the Discussion Paper include:

   (a) the rapid development of AI and related technologies is in many respects outpacing the legal and regulatory frameworks necessary to guide and govern them, giving rise to risks that the privacy and other fundamental rights of individuals may be placed at risk;

   (b) an ethics framework for AI should be rights-based and strongly grounded in overarching principles, including those drawn from international human rights law, and subject to principles of the rule of law and procedural fairness;

   (c) to be effective and consistently applied, an ethics framework must be enforceable. Further, careful discussion is required regarding potential models of an enforcement mechanism, which may vary between contexts;

   (d) establishment of a regulatory body should be carefully considered to provide oversight of AI systems and their development and use;

   (e) accountability and liability associated with the development, implementation and use of AI systems should be determined so as to provide remedies in cases where damage is caused, and to allow for appropriate scrutiny by the Courts; and

   (f) greater distinction (and further discussion) is needed between public and private sector applications of AI, noting the different principles of administrative law which apply particularly to decision-making by government organisations and other public sector entities.

4. This submission puts forward a number of recommendations which, in summary, address the following issues:

[1] D Dawson et al, 'Artificial Intelligence: Australia's Ethics Framework' (Discussion Paper, Data61 CSIRO, 2019) 3 <https://consult.industry.gov.au/strategic-policy/artificial-intelligence-ethics-framework/supporting_documents/ArtificialIntelligenceethicsframeworkdiscussionpaper.pdf>.

*Introduction*

(a) the need for a flexible and inclusive definition of AI [17];

*Question 1: are the principles put forward in the discussion paper the right ones? Is anything missing?*

(b) inclusion of a principle of 'respect for human rights and human autonomy' [22];

(c) further consideration of what constitutes 'net-benefits' [30] and their subjection to principles of the rule of law and equality before the law [32]; and a requirement that AI systems disclose their benefits and detriments [33];

(d) expanding the principle of 'doing no harm' to address system design, integrity and vulnerabilities over time [36]; and restriction of the use of AI for 'scoring' of citizens, in line with a rights-based framework [38];

(e) formal implementation of compliance measures to provide oversight, enforcement and redress [42]; as well as a requirement for registration with [48] and periodic audit by [49] an independent regulator;

(f) further refinement of the principle of 'privacy protection' [57]; consistent use of privacy-related terminology [66]; further discussion of lawful, fair and transparent data-handling and necessary restrictions arising from privacy law [72]; limiting the use of personal information for secondary purposes [80]; addressing data quality to reduce bias [81]; and consideration of a federal Charter of Rights as an ethical foundation for AI systems [89];

*Question 2: Do the principles put forward in the discussion paper sufficiently reflect the values of the Australian public?*

(g) importance of inclusive public consultation with people likely to be affected by AI [102];

(h) an ethical framework for AI to be based on human rights principles, informed by good precedents [105] and giving further consideration to inclusion of employment protections [106];

*Question 5: What other tools or support mechanisms would you need to be able to implement principles for ethical AI?*

(i) capability of regulatory bodies and key civil society actors to be increased, complementary to the role of an AI regulator [108];

*Question 7: Are there additional ethical issues related to AI that have not been raised in the discussion paper? What are they and why are they important?*

(j) further consideration to be given to the application of administrative law principles to AI, including where used by government and public sector decision-makers [114];

*General observations:*

(k) caution recommended with regard to linking levels of risk to numbers of persons affected [121];

(l) further consideration to be given to data governance in AI [125]; and

(m) a potential onus on AI systems to demonstrate accuracy to be considered.

5. The Law Council welcomes the opportunity to provide this submission and would be happy to elaborate further on any of the points addressed.

# Introduction

6. During 2018, the Data61 Insight Team of CSIRO undertook Commonwealth-funded research on the ethical framework associated with the development and application of AI systems in non-military contexts in Australia and globally. The Discussion Paper documents that research and now forms the basis of the Consultation.

7. For the purpose of the Consultation, the Discussion Paper defines AI as: 'A collection of interrelated technologies used to solve problems autonomously and perform tasks to achieve defined objectives without explicit guidance from a human being'.[2]

8. Within that definition, the Discussion Paper notes that technologies benefiting from AI have the potential to strengthen the economy, increase the cohesion and inclusiveness of society, support sustainability and efficient use of resources, and increase wellbeing. However, it also acknowledges that the realisation of these benefits, and securing the trust and acceptance of the Australian community for AI systems, is substantially contingent upon the design, application and governance of AI systems being consistent with accepted ethical standards and values.

9. While the Discussion Paper notes that there are no simple answers to complex issues, it proposes eight 'core principles' for AI[3] (**Principles**), and a 'toolkit for ethical AI'[4] containing nine elements. The <u>Principles</u> are:

   (a) generation of net benefits;

   (b) doing no harm;

   (c) regulatory and legal compliance;

   (d) privacy protection;

   (e) fairness;

   (f) transparency and explainability;

   (g) contestability; and

   (h) accountability.

10. The components of the <u>toolkit</u> are:

   (a) impact assessments;

   (b) review processes;

   (c) risk assessments;

   (d) best practice guidelines;

   (e) education, training and standards;

   (f) business and academic collaboration;

[2] Ibid, 14.
[3] Ibid, 57.
[4] Ibid, 8.

(g) mechanisms for monitoring and improvement;

(h) recourse mechanisms; and

(i) consultation.

11. The Law Council supports the steps the Australian Government is taking to ensure the emerging ethical issues associated with AI are properly considered. While generally supportive of the approach taken in the Discussion Paper, the Law Council considers that several issues require careful further consideration and development. These issues are set out below in response to questions 1, 2, 5, 6 and 7, as posed by the Discussion Paper. The Law Council then makes some additional general observations.

12. In making its submission, the Law Council is mindful of the continuing and rapid development of AI and associated technologies, the difficulty of predicting how such technologies may interact with or impact on people in the future, and the difficulty faced by the law in keeping pace with technological developments. The Law Council has previously stated:

> *Technological developments are advancing at an exponential rate that outpaces the ability of our current legal and regulatory frameworks to keep pace. This uneven development poses a very real threat to fundamental human rights and freedoms, particularly the right to privacy, the right to a fair trial, the right to non-discrimination and equal protection of the law and the right to liberty and security of the person*.[5]

13. Australia is a party to a number of key human rights treaties and other instruments, and it is obliged to implement the provisions of those instruments within its territory. A rights-based approach to the regulation of AI would place a positive obligation on the developer of any AI system to ensure that any impact on rights flowing from the operation of that system is necessary, reasonable, proportionate and in pursuit of a legitimate objective. As a matter of overarching policy, the Law Council supports the introduction of a federal charter of rights in Australia. Such a charter would provide domestic certainty and consistency over time with respect to the specific rights and freedoms protected and how tensions between these rights should be balanced.

14. Regarding the regulation of AI, the Law Council has stated that:

> *…fairness, transparency, non-discrimination and accountability should be the central focus of regulation in the area of AI so as to prevent inequality from becoming further entrenched within social, governmental and economic systems. The Law Council supports the establishment of appropriately regulated AI-informed decision-making processes which will allow for the benefits of AI to be provided to society while protecting fundamental rights, including the rights to privacy and non-discrimination*.[6]

15. With regard to defining AI, the Law Council accepts the distinction set out in the Discussion Paper between 'narrow' and 'general' AI,[7] where the former is focused on the performance of specific functions and not intended—as in the case of the latter—

[5] Law Council of Australia, 'Artificial Intelligence: Governance and Leadership' (18 March 2019), 2, <https://www.lawcouncil.asn.au/resources/submissions/artificial-intelligence-governance-and-leadership>.
[6] Law Council of Australia, 'Human Rights and Technology' (25 October 2018), [83], <https://www.lawcouncil.asn.au/resources/submissions/human-rights-and-technology>.
[7] D Dawson et al, 'Artificial Intelligence: Australia's Ethics Framework' (Discussion Paper, Data61 CSIRO, 2019) 14.

to replicate (or necessarily be capable of replicating) human intelligence and autonomy across multiple fields. While the Law Council does not dispute that 'general' AI is unlikely to be achieved prior to 2030, it nevertheless recommends the use of a definition of AI which is broad enough to ensure uniform and consistent regulation of varying and evolving forms of AI systems, as well as other forms of machine learning or automated decision-making.

**Recommendation:**

- **For the purpose developing an ethics framework for AI, a definition should be used which is flexible enough to encompass future developments, and inclusive enough to cover the full range of related technologies, including AI, machine learning and automated decision-making systems.**

16. The Law Council also considers it important that DIIS and CSIRO distinguish between public and private sector decision-making and give further consideration to the different ethical frameworks which may apply to each. In particular, the Law Council considers that expert advice is needed on the application of administrative law principles to government decision-making and similar public sector uses of AI. These issues are addressed in more detail below.

## Question 1: are the principles put forward in the discussion paper the right ones? Is anything missing?

17. The Law Council in general supports the Principles and risk assessment as proposed by the Discussion Paper,[8] however it makes a number of comments for the further consideration of DIIS.

18. The Law Council notes that at least 63 public-private initiatives have produced statements describing high-level principles, values, and other tenets to guide the ethical development, deployment and governance of AI.[9] All statements include similar principles on transparency, equality/non-discrimination, accountability and safety. Some statements include additional principles, such as the requirement for AI to be socially beneficial and to protect human rights.

19. While the Law Council considers the Principles broadly reflect these listings, and form a suitable basis upon which to start a discussion about when and how they should be applied, we suggest that the principles be tested for comprehensiveness. To that end, the Law Council recommends DIIS and CSIRO consider the inclusion of additional principles to ensure a robust ethics framework.

20. The Law Council suggests that the principle of 'respect for human rights and human autonomy' should be included as an independent principle.

[8] Ibid, 6, 57, 64.

[9] For a non-exhaustive recent listing, see AlgorithmWatch, *AI Ethics Guidelines Global Inventory*, <https://www.rri-tools.eu/-/ai-ethics-guidelines-global-inventory>.

**Recommendation:**

- **Principles for ethical AI should include 'respect for human rights and human autonomy'.**

21. Other principles that warrant further exploration are:

    (a) human agency (sufficient information, training and education to make informed choices);

    (b) human oversight (there are many possibilities for human oversight and intervention during the life cycle which require discussion);

    (c) accuracy, reliability and robustness of inputs, processes and outputs (including good data governance);

    (d) technical robustness (level of imperviousness to malware and cyberattack);

    (e) safety and wellbeing (what to do in the event of a problem); and

    (f) environmental impact and conservation of natural resources.

22. These principles are not specific to AI: many considerations equally apply to other advanced data analytics services that may lead to significant adverse effects on humans, whether or not these services involve decision-making with humans in-the-loop or automated decision-making. AI may introduce more intractable issues as to explainability and autonomy but, generally, the same principles should be applied in the evaluation of all applications of advanced data analytics systems, including those incorporating machine learning and other forms of automated decision-making. For example, there are a number of statistical, data driven tools used in criminal procedure to predict future reoffending and assess 'unacceptable risk' that are not strictly AI but are, rather, 'actuarial' or 'algorithmic' instruments. These tools have been critiqued for their use of group data instead of individual data, thus challenging individualised justice.[10]

23. The Law Council notes that the Principles reflect existing requirements under specific laws, the operation of human rights protections in Australian law, and non-mandatory ethical principles. It recommends that the distinction between fundamental rights, requiring comprehensive legal protection, and ethical considerations forming the basis of guiding principles should be considered in more detail. The Law Council recommends that DIIS and CSIRO specifically note this interplay so that compliance with law and regulation, in addition to ethical considerations, is properly addressed.

24. The Law Council further notes the significant overlap between concerns about nurturing of citizen and consumer digital trust, or 'social licence', and the fact that

[10] See, for example, Rajan Darjee et al, 'Risk of Sexual Violence Protocol (RSVP): A real world study of the reliability, validity and utility of a structured professional judgement instrument in the assessment and management of sexual offenders in South East Scotland' (NHS Lothian Sex Offender Liaison Service, January 2016) <https://www.researchgate.net/publication/294718597_Risk_of_Sexual_Violence_Protocol_RSVP_A_real_world_study_of_the_reliability_validity_and_utility_of_a_structured_professional_judgement_instrument_in_the_assessment_and_management_of_sexual_offenders_>.

often, good ethical decisions about applications of digital data reflect sound business judgement as to sustainability of business models. For example, there is now growing public focus on the uses of facial recognition and other automated surveillance technologies, and the secondary uses of health data, biometric[11] and of geo-location data. As public concerns around uses of data evolve, business models that have been built upon such activities may be rendered redundant. Ethical review should include consideration of the effect of a particular application of AI upon digital trust of consumers, citizens and users.

25. The Law Council submits that any statement of principles is of little or no benefit unless the principles are in practice consistently, reliably and verifiably applied whenever they should be applied.[12] This requires a clear statement of the threshold condition for determining when the principles should be applied, including a risk assessment methodology to determine at what point risks are such that a comprehensive impact assessment should be undertaken. The Law Council considers that an ethics framework should include an analysis that will assist organisations to understand when, and in what contexts, an initial risk assessment is required, and the point at which a comprehensive impact assessment should be undertaken.

26. The Law Council also notes that risk assessment, and risk mitigation in the design and deployment of AI applications, requires appropriate risk management in AI-related decision-making, robust internal governance systems and measures, accountability mechanisms, and appropriate relationship management of users and consumers. It is particularly important that an accountability mechanism (particularly within organisations employing AI) clearly attributes responsibility to specific individuals for ensuring that principles have been appropriately applied. The basis for relevant decisions made by those individuals, and the extent of executive oversight of those decisions, must also be covered.

## Principle 1: generates net-benefits

27. The Law Council considers that the principle of 'doing no harm' (principle 2) is key to the development of AI. As such, it should be the first listed principle, followed by consideration of net-benefits. The discussion below also considers how the applicable principles may be appropriately balanced with respect to a human rights framework analysis when tensions arise.

28. The Law Council recommends that further consideration be given, or further clarification provided, regarding the first principle. In particular, it is not clear from the Discussion Paper what would constitute net benefits and who are the people who should benefit from them. It should be clarified whether the 'people' to whom principle 1 refers must include any or all of the user (or users), the proprietor, operator, developer, or possibly investors associated with a given AI system.

[11] *Jeremy Lee v Superior Wood Pty Ltd* [2019] FWCFB 2946.

[12] See further, Luciano Floridi, *Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical*, (2019) <https://doi.org/10.1007/s13347-019-00354-x>; Personal Data Protection Commission Singapore, *A Proposed Model Artificial Intelligence Governance Framework,* (January 2019); UK Information Commissioner's Office*, An Overview of the Auditing Framework for Artificial Intelligence and its Core Components,* (26 March 2019) < https://ai-auditingframework.blogspot.com/2019/03/an-overview-of-auditing-framework-for_26.html>.

**Recommendation:**

- **The first principle ('generates net-benefits') should be considered further, including to clarify what would constitute net benefits and who should benefit from them.**

29. Additionally, net benefits to a community or to society as a whole should be considered. It is possible that an AI system may offer potential benefit to the community but be detrimental to the individual or vice versa. The Discussion Paper considers this issue in a limited fashion in the context of automated vehicles,[13] however an important set of ethical questions are raised, which the Law Council considers require further attention.

30. The Law Council submits that this principle should be made expressly subject to the legal principles of the rule of law and equality before the law. It should not be permitted for certain users of AI to pay for priority treatment in respect to fundamental legal or human rights.

**Recommendation:**

- **Principle 1 ('generates net-benefits') should be expressly subject to the legal principles of the rule of law and equality before the law.**

31. The Law Council also recommends that an organisation using or proposing to deploy an AI system should, potentially as part of a prior assessment process, specify comprehensively both what the benefits will be (and who will receive them), and also what the detriments will be (and who may experience them).

**Recommendation:**

- **An organisation using or proposing to deploy an AI system should be required to specify what the benefits and detriments will be, and who, respectively, will receive or experience them.**

## Principle 2: doing no harm

32. The Law Council considers that this principle may be too narrow in its current framing. At present, the principle requires that AI systems must not be designed to harm or deceive people and should be implemented in a way that minimises any negative outcomes.

33. The principle does not, however, address harm caused by negligent or reckless design, and it is not specific in regard to what constitutes a negative outcome. The

[13] D Dawson et al, 'Artificial Intelligence: Australia's Ethics Framework' (Discussion Paper, Data61 CSIRO, 2019) chapter 6.1.3.

potential for a system, even where otherwise designed with sufficient integrity, to be tricked or spoofed also requires additional emphasis.

34. The Law Council recommends that the principle be extended to cover negligence and recklessness in the design of AI, as well as reasonable resistance to hacking, tricking or spoofing. Additionally, the Law Council recommends an ongoing requirement to address swiftly data security concerns or other vulnerabilities that may become apparent at a later stage, throughout the operational life of an AI system.

**Recommendation:**

- **Principle 2 ('doing no harm') should be extended to address, in addition, negligent or reckless design, resistance to hacking and spoofing, and a requirement to rectify flaws or vulnerabilities as they are identified throughout the operational life of an AI system.**

35. The Law Council suggests that a clear prohibition on the use of AI systems for inappropriate or discriminatory 'scoring' of citizens (such as, for example, occurs within the Chinese 'social credit' system)[14] be included.

36. The Law Council recommends that guidance be taken from the EC *Ethics Guidelines for Trustworthy AI*, which state:

> *Societies should strive to protect the freedom and autonomy of all citizens. Any form of citizen scoring can lead to the loss of this autonomy and endanger the principle of non-discrimination. Scoring should only be used if there is a clear justification, and where measures are proportionate and fair. Normative citizen scoring (general assessment of "moral personality" or "ethical integrity") in all aspects and on a large scale by public authorities or private actors endangers these values, especially when used not in accordance with fundamental rights, and when used disproportionately and without a delineated and communicated legitimate purpose*.[15]

**Recommendation:**

- **The use of AI systems for 'scoring' of citizens should be restricted, to avoid the undermining of human rights, in accordance with a rights-based framework incorporating principles of legitimate purpose, necessary justification, reasonableness and proportionality. The terms of the EC *Guidelines for Trustworthy AI* should be used as a guiding framework.**

37. Applications with the potential to reduce the access of certain individuals to essential services (for example, where higher paying AI users receive access to faster and

[14] The calculation of 'social credit' in China may result in a person being deemed 'untrustworthy' and, as a consequence, being denied access to certain flights or train lines, being denied credit or home ownership, or potentially being subjected to pre-emptive arrest and 're-education'. See further: Nicole Kobie, 'The Complicated Truth about China's social credit system', *Wired* (7 June 2019), <https://www.wired.co.uk/article/china-social-credit-system-explained>.

[15] *Ethics Guidelines for Trustworthy AI*, Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission (8 April 2019), 34, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

higher quality healthcare, causing delay and reduction in resources available to free or lower paying users) may also require careful consideration and regulation or restriction.

38. Additionally, the risk of heightened disadvantage, or of harm resulting from a lack of access to services, must be considered in circumstances where individuals or groups within the community are unable to connect with AI systems. Among other groups, people living in remote locations, those with little digital literacy, people with disability or certain health conditions, those of culturally and linguistically diverse backgrounds, and people living in poverty may find access to the internet or the use of digital systems very difficult. The regulation and increasing deployment of such systems should be undertaken with care to avoid a ‘digital divide’, and to offer alternative, ‘analogue’ means of access to services for those who need it.

## Principle 3: regulatory and legal compliance

### Importance of enforcement and potential models

39. Principle 3 provides: 'The AI system must comply with all relevant international, Australian Local, State/Territory and Federal government obligations, regulations and laws’. The Law Council considers that this principle should be clarified to address whether AI systems and practices comply with existing law and regulation, to identify any regulatory gaps, and to provide for necessary law reform and creation of new regulations to address these gaps. These should extend to the data and outcomes produced as part of the operation of AI.

40. By way of example, in the context of government decision-making relating to immigration, section 495A of the *Migration Act 1958* (Cth) allows the Minister to:

> (1) …arrange for the use, under the Minister's control, of computer programs for any purposes for which the Minister may, or must, under the designated migration law:
>
> (a) make a decision; or
>
> (b) exercise any power, or comply with any obligation; or
>
> (c) do anything else related to making a decision, exercising a power, or complying with an obligation.

41. The range of functions for which ‘computer programs’ may be used is broad and includes decisions which require the Minister to be satisfied that applicable visa criteria have been met.[16] Although the use of a computer program is authorised, it is unclear whether the delegation of assessments or judgment-based decisions to an AI system (as opposed to the use of a computer program for efficiency or programme management purposes) would otherwise satisfy requirements for the exercise of a discretion. It would also be difficult to show whether regard had been had (by the AI system) to prescribed factors in reaching that decision.

42. The Law Council notes that the Discussion Paper does not directly address questions of enforcement of the proposed ethics framework. The Law Council believes that an effective oversight, enforcement and redress regime is essential to the proper regulation of AI, particularly considering the serious consequences for people which

[16] See: *Migration Act 1958* (Cth), s 65(1)(a).

may flow from its improper use. Accordingly, it recommends the formal implementation of legal and regulatory compliance to ensure enforceability.

**Recommendation:**

- **Legal and regulatory compliance measures should be formally implemented to ensure an appropriate and effective oversight, enforcement and redress regime for AI systems.**

43. The Law Council also supports a requirement that ethics and rule of law principles be included 'by design', similarly to the requirement for privacy by design. The Law Council recommends adoption of the European Commission's (**EC**) *Ethics Guidelines for Trustworthy Artificial Intelligence* in this respect.[17]

44. While the Law Council believes it is essential that the ethical framework applied to AI complements its legal regulation and is enforceable, the mechanism used to ensure compliance requires careful further consideration, including with regard to the range of contexts in which AI will operate and the different responsibilities to which public and private sector operators of AI are subject. One approach may be to ensure full regulation of AI systems and/or developers and operators. This approach could, however, prove be very bureaucratic, time consuming and resource intensive. A different approach may be to ensure that accountability and liability are clearly attributed to specified parties (addressing the obligations of users of AI systems as well as those of developers and owners), and thereby place reliance on market reputation and the risk of litigation acting as incentive and deterrent (respectively) to enforce the ethical framework.

45. The Law Council has received support for the development of a regulatory framework reflecting the characteristics set out below, however further analysis and discussion of these issues is needed, mindful also that different frameworks may be appropriate for different scenarios. The Law Council notes that the Discussion Paper engages with these issues at a high level and would be pleased to comment further as the discussion progresses.

## Compliance mechanisms

46. With regard to mechanisms to ensure compliance, the Law Council has made previous submissions, as discussed under question 5 below, regarding the possible establishment of a 'responsible innovation organisation' as an independent AI regulator which could contribute substantially to these functions. Regardless, any regulator that is established should oversee the implementation of, compliance with, and enforcement of an ethical framework and principles for AI.

47. For example, the framework should regulate how an organisation will develop, review and audit AI-enabled products and services so that the organisation's development process is independent of any conflicting interests. This may require an organisation to have an approved policy on the ethical use of AI (with its own modelled framework and procedures), data security and systems integrity controls, training requirements, and an independent review committee responsible for approval prior to the launch of

[17] *Ethics Guidelines for Trustworthy AI*, Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission (8 April 2019), 21, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

proposed AI products into the market. Projects classified as large or high risk would require separate assessment by the independent regulator.

48. The Law Council recommends that organisations which develop and/or implement AI-related products and services should be registered with a regulator to ensure full oversight by the regulator, and to ensure that the products and services are developed in full compliance with the ethical framework. In circumstances where a sufficiently serious system bug or security breach is identified, the regulator may be required to issue a public notice to alert consumers to the associated risks until such time as the issue has been resolved.

**Recommendation:**

- **Organisations which develop or implement AI should be registered with a regulator responsible for oversight and compliance.**

49. The Law Council recommends that, in addition to compliance with existing security and cyber-security standards, organisations which develop and/or implement AI-related produces and services be subject to periodic external audit by the regulator. Responsibility to remedy any compliance gaps would rest with the developing/implementing organisation, as would liability for any damage caused.

**Recommendation:**

- **Periodic external audit by the regulator should complement ongoing oversight of AI development and implementation.**

50. This approach is considered important by the Law Council, as it will address uncertainty regarding responsibility at law for impacts flowing from the use of AI systems, and it will encourage the early development of risk assessment and mitigation strategies. An organisation's investment into creating robust and ethical AI systems will strengthen the integrity of the resulting products and services, and also provide reassurance to consumers regarding the safety and fitness for purpose of those products and services.

51. Chapter 5.5 of the Discussion Paper addresses the topic of medical predictions, commenting that 'AI systems used in health care require close management and gold standard research before implementation.'[18] The Law Council agrees, but understands that the implementation of AI is already proceeding in health care and is, at present, subject only to limited ethical and regulatory guidance. The Royal Australian and New Zealand College of Radiologists (**RANZCR**) has announced its own draft ethical principles for AI in medicine[19] and the Therapeutic Goods Administration (**TGA**) has some regulatory oversight of AI in the medical field. However, this only applies where an AI system fits the definition of a 'medical device'

[18] D Dawson et al, 'Artificial Intelligence: Australia's Ethics Framework' (Discussion Paper, Data61 CSIRO, 2019) 44.

[19] See: RANZCR, 'Media Release – RANSCR unveils unique Artificial Intelligence Guidelines for Healthcare' (21 February 2019), <https://www.ranzcr.com/whats-on/news-media/308-media-release-ranzcr-unveils-unique-artificial-intelligence-guidelines-for-healthcare>.

per the *Therapeutic Goods Act 1989* (Cth).[20] Any application of AI which is not directly involved in diagnosis, prevention and monitoring of disease will likely fall outside this definition, meaning that broader health or lifestyle applications of AI are unregulated by the TGA.

52. The Law Council believes that the necessity of new laws or regulation be recognised as a key element to be taken into account in the development of an ethical framework for AI.

## Principle 4: privacy protection

53. Principle 4 is described by the Discussion Paper as follows: 'Any system, including AI systems, must ensure people's private data is protected and kept private and confidential plus prevent data breaches which could cause reputational, psychological, financial, professional or other types of harm.'[21]

54. The Discussion Paper recognises that issues related to AI ethics are closely intertwined with those that relate to data sharing, typically on a large scale involving some data matching and forms of monitoring. It also recognises that privacy is crucial in any discussion related to AI ethics. The Discussion Paper seeks to set out the key requirements of privacy and data sharing laws in Australia in sections 2.1.3 and 2.1.4 as well as chapter 3.

55. The Law Council is concerned that the Discussion Paper is not an accurate summary of privacy law in Australia and considers that a more robust understanding of privacy law, both in theory and in practice, is essential to developing an AI ethics framework that firmly aligns with the legal framework. The explanation of privacy law in Australia and its application to AI is incomplete and, as a result, the Law Council considers that Principle 4 risks placing undue importance on confidentiality and security of 'private data' (further detail is set out below).

56. In developing an AI ethics framework, the Law Council submits that the key focus for privacy and AI should be on privacy protection and adequate data governance. The EC *Ethics Guidelines for Trustworthy AI*[22] provides relevant guidance in that respect and covers:

    (a) privacy and data protection: ensuring the lawfulness of the initial collection of personal data, and of the collection or generation of data about the individual over the course of their interaction with a given system; and also ensuring that the data collected will not be used unlawfully or unfairly to discriminate against the data subject; and

    (b) adequate data governance: ensuring the quality and integrity of the data used, its relevance and proportionality to the context and purpose for which it is collected, as well as ensuring the use of access protocols and secure and privacy-sensitive data handling processes.

[20] *Therapeutic Goods Act 1989* (Cth), s 41BD.

[21] D Dawson et al, 'Artificial Intelligence: Australia's Ethics Framework' (Discussion Paper, Data61 CSIRO, 2019) 6.

[22] *Ethics Guidelines for Trustworthy AI*, Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission, (8 April 2019) <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> 17.

57. The Law Council recommends that DIIS and CSIRO engage with privacy regulators, and practitioners with specialist expertise in privacy law and practice, to assist in refining Principle 4. The rationale for this recommendation is as follows.

**Recommendation:**

- **DIIS and CSIRO should engage with specialist privacy regulators and practitioners to refine principle 4 ('privacy protection').**

## Legislative framework

58. The Law Council considers that the technical complexity of privacy law in Australia and privacy principles needs to be dealt with in greater depth than it is in the Discussion Paper. It recommends a more robust and technically nuanced discussion about rights and obligations under privacy law in Australia to properly address privacy as a central issue of the Discussion Paper.

59. The Discussion Paper does not currently address the development of lawful and ethical AI within the framework of Commonwealth, state and territory privacy legislation. The *Privacy Act (1988)* (Cth) (**Privacy Act**) and the Australian Privacy Principles (**APP**) set out in the *Privacy Act*[23] are not the only privacy laws in Australia governing the use of AI. Privacy principles established at the state and territory level also regulate access to personal information, including providing certain exemptions in the case of authorised law enforcement activities.[24] The interaction of AI with these laws and processes must be considered, along with implications for the access and use of meta-data including geo-location information generated by communications services and data generated by use of apps provided by heath service providers, universities and other publicly funded bodies when employing AI systems.

60. The Law Council also notes that the European Union's (**EU**) General Data Protection Regulation (**GDPR**)[25] may apply to organisations in Australia employing AI systems that operate in the EU, or offer their goods or services to, or monitor the behaviour of, people in the EU.[26]

## Consistent terminology

61. Principle 4 uses the term 'private data'. While the terms 'personal information' and 'personal data' are sometimes used interchangeably, the term 'private data' does not reflect privacy law in Australia or overseas. The use of this terminology, which the Discussion Paper does not define, is of concern to the Law Council for two reasons.

62. The first is that the Law Council believes it risks implying that not all data owned by an individual, or relating to an individual, is private by default. Personal information does not need to be 'private' for it to be protected under privacy laws. Additionally, references to 'private data' (of or relating to an individual) could create uncertainty

[23] *Privacy Act 1988* (Cth) s 14 and Schedule 1.

[24] See, for instance, the *Privacy and Personal Information Protection Act 1998* (NSW), Part 2 'Information protection principles'; the *Information Privacy Act 2009 2009* (Qld), Chapter 2 'Privacy principles'; and the *Personal Information Protection Act 2004* (Tas), Part 3 'Personal Information protection principles'.

[25] *Parliament and Council Regulation EU/2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)* [2016] OJ L 119/1.

[26] Ibid, art 3.

about exactly what constitutes private data, and the circumstances in which it becomes, remains or ceases to be private. Discussion of private data could also be taken to imply that other data, by default, is not private and not subject to the same protections.

63. The second reason is that this terminology is out of step with other key regulatory instruments, which draw a distinction between personal and non-personal data, rather than public and private data. Within Australian law, the Privacy Act provides protections for ‘personal information’, which it defines as:

> … information or an opinion about an identified individual, or an individual who is reasonably identifiable:
>
> (i) whether the information or opinion is true or not; and
>
> (ii) whether the information or opinion is recorded in a material form or not.[27]

64. Internationally, the GDPR also engages with ‘personal data’[28] as does the Office of the United Nations High Commissioner for Refugees in its Data Protection Policy.[29]

65. The Law Council submits that ‘the fundamental starting point must be that “your data is always your data”.’[30] The Law Council also observes that this approach is consistent with recent recommendations of the Productivity Commission, which have called for giving individuals and consumers (including some businesses) more control over their data.[31]

66. The Law Council recommends that the terminology of 'personal information' be used, to ensure organisations employing AI systems have certainty about their compliance requirements.

**Recommendation:**

- **The use of terminology, particularly concerning privacy, should be reviewed for consistency with key terminology in existing law and regulations. The phrase ‘personal information’ should be used in place of ‘private data’.**

67. The description of personal information as 'sensitive' should also be used in consistency with the *Privacy Act*, which establishes 'sensitive information' as a subset of personal information.[32] Higher standards apply under the Privacy Act when sensitive information is collected, used or disclosed.

[27] *Privacy Act 1998* (Cth) s 6.

[28] Ibid, art 4(1).

[29] UNHCR, *Policy on the Protection of Personal Data of Persons of Concern to UNHCR* (May 2015), 1.4, <https://www.refworld.org/docid/55643c1d4.html>.

[30] Law Institute of Victoria, ‘Consultation on Artificial Intelligence: Australia’s Ethics Framework’ (24 May 2019), 1.

[31] Productivity Commission, *Data Availability and Use* (Report No 82, 31 March 2017), 14-15.

[32] *Privacy Act 1988* (Cth) s 6.

## Consent process is not fundamental to protecting privacy

68. Section 3.1 of the Discussion Paper states that ‘protecting the consent process is fundamental to protecting privacy’.[33] The Discussion Paper does not discuss the rationale for this conclusion and the Law Council considers that the risk assessment framework in section 7.2 of the Discussion Paper consequently places an over-reliance on 'consent'.

69. There is often a misconception that Australian law requires organisations to obtain consent from the individual to collect, use and disclose their personal information. This misconception is compounded by the fact that many organisations require users to ‘agree’ to their privacy policy as part of a registration process.

70. Although there may be some overlap in their contents, the requirements for privacy policies pursuant to APP 1 and collection notices pursuant to APP 5 are two separate requirements under the Privacy Act. Neither APP 1 nor APP 5 require a user to consent to a privacy policy or collection notice.

71. Consent is not the sole mechanism by which the collection, use or disclosure of personal information may lawfully be authorised under the *Privacy Act*. In fact, consent is often the exception for collection, use or disclosure of personal information under the *Privacy Act*. The Privacy Act currently requires that individuals provide consent when their personal information is collected in only limited circumstances, including:

    (a) the use or disclosure of personal information for a secondary purpose (APP 6.1(a));

    (b) the collection of sensitive information (APP 3.3(a));

    (c) the collection of personal information by an agency from someone other than the individual (i.e. an individual must consent for an agency to disclose their personal information to another agency) (APP 3.6(a)(i));

    (d) the use or disclosure of personal information or sensitive information for direct marketing purposes (APP 7.3(b) and APP 7.4); and

    (e) the disclosure of personal information to an overseas recipient (APP 8.2(b)).

72. The Law Council therefore recommends that DIIS and CSIRO emphasise the requirement for the lawful, fair and transparent collection, use and disclosure of personal information. This should include discussion of the restrictions or limitations on collection, use and disclosure of personal information as set out in the relevant privacy laws. Future consideration of this topic should also focus on how these restrictions and limitations apply to personal information collected initially from the user, as well as data generated about the user over the course of their interaction with the given AI system.

[33] D Dawson et al, ‘Artificial Intelligence: Australia’s Ethics Framework’ (Discussion Paper, Data61 CSIRO, 2019) 28.

**Recommendation:**

- **The lawful, fair and transparent collection, use and disclosure of personal information should be emphasised by DIIS and CSIRO, with further discussion of the restrictions and limitations established by privacy laws.**

## Right to be forgotten

73. The Discussion Paper links the currency of consent with the absence of a 'right to be forgotten'.[34] The 'right to be forgotten' or 'right to erasure' in the GDPR[35] is in part related to the issue of consent,[36] but is connected with the right of access and correction. The 'right to be forgotten' is also not an absolute right under the GDPR and will not apply if the organisation has collected or is using the personal information lawfully and fairly and still has a current need for that information. There is no 'right to be forgotten' in Australian privacy law.

74. APP 11.2 does adopt a similar approach, requiring an APP entity to destroy or ensure that personal information is de-identified if the entity no longer needs the information for any purpose for which the information may be used or disclosed, irrespective of whether the individual has requested that action or not.

## Protecting privacy extends beyond confidentiality and security

75. Privacy protection presents broader requirements, extending beyond ensuring that the security and confidentiality of personal information is maintained. The Law Council considers that privacy should not be conflated with confidentiality. Securing personal information does not necessarily mean privacy has not been breached.

76. Invasions of privacy could arise from unlawful collection of personal information to develop an AI system, for example by using personal information collected for an unauthorised purpose or a secondary purpose beyond the expectation of the data subject upon initially providing the information. Invasions of privacy have the potential to cause harm to the individual, even in circumstances where there was no data breach.

77. AI systems must be capable of respecting specific protections owed to certain categories of data. For example, information which is subject to legal professional privilege should not be accessed, transmitted or used in any way contrary to that status.

78. The Law Council considers that specific protection for employment related data should be discussed, noting that many AI applications can operate in the employment relationship and rely on data collected in the course of the employment relationship, including recruitment and compliance related activities. Discussion of this issue

[34] Ibid.

[35] *Parliament and Council Regulation EU/2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)* [2016] OJ L 119/1, art 17.

[36] Ibid.

should take account of the current employee record exception under the *Privacy Act*[37] and its limits and complexities (including definitional arguments as to its scope).[38]

79. The Law Council considers that a significant challenge to the protection of privacy, arising from the development or application of AI, is the use and disclosure of personal information for secondary purposes. For example, personal information that was collected for a different primary purpose could be re-used for a secondary purpose of development of an AI system, or be drawn upon by an AI system operating beyond the scope of the original primary purpose. The Law Council has previously stated related concerns, including that

> *…with technological developments in data analytics, consumers are increasingly at risk when information provided at one point in time when consent was given could be used in the future in ways the consumer had not envisaged when they gave their consent.*[39]

80. The use of personal information for secondary purposes is not permitted under the Privacy Act unless an exception applies under APP 6. Consent is one of those exceptions,[40] but obtaining consent may be considered impracticable, for example in situations where a large cohort of individuals' personal information is used, and where the requirement for consent to be informed and specific means that the data subject must know specifically how their information will be used in the AI system.

81. The Law Council recommends that DIIS and CSIRO address the restrictions and limitations on the use and disclosure of personal information for secondary purposes and the requirements for obtaining consent in the AI context.

**Recommendation:**

- **Restrictions and limitations on the use and disclosure of personal information for secondary purposes should be addressed, along with requirements for obtaining consent in the context of AI.**

82. The Law Council also recommends that future consideration of this issue address the importance of the quality of data and access to data in the privacy and data protection context. Transparency of algorithms can mitigate harmful privacy risks and visibility of the data used in automated decision-making can prevent skewed data input and thus, the generation of biased datasets.

**Recommendation:**

- **Data quality and data access should be addressed in the privacy context, including to reduce risks of bias.**

[37] *Privacy Act 1988* (Cth), s 7B(3).

[38] See, for example: *Jeremy Lee v Superior Wood Pty Ltd* [2019] FWCFB 2946.

[39] Law Council of Australia, 'Digital Platforms Inquiry – Preliminary Report' (15 February 2019), [24] <https://www.lawcouncil.asn.au/resources/submissions/digital-platforms-inquiry-preliminary-report>.

[40] *Privacy Act 1998* (Cth), Schedule 1, subsection 6.1(a).

83. Privacy impact assessments may be a means by which the potential privacy impacts associated with specific AI applications can be assessed and risk mitigation strategies developed prior to the collection and use of personal information. Safeguards such as data minimisation and purpose specification should be implemented to prevent the unauthorised collection, use and disclosure of personal information. Data subjects should also be able to access their personal information that has been collected, used or generated in the AI system and seek redress if they have been affected by a decision made by the AI system.

## Eligible data breach

84. The Discussion Paper in section 3.2 raises a notification requirement in the event of a data breach: 'if personal data is accessed or disclosed in any unauthorised way that may cause harm, all affected individuals must be notified'.[41] The Law Council considers that this description could benefit from review.

85. The Notifiable Data Breach scheme requirements[42] are not limited to unauthorised access or disclosure; they also include loss of the information. The Notifiable Data Breach scheme applies to all personal information regulated under the Privacy Act, including consumer credit reporting information and Tax File Numbers.[43] A data breach giving rise to the obligation to notify is one where an individual is likely to suffer 'serious harm', not just harm, and the obligation is to the Office of the Australian Information Commissioner (**OAIC**) as well as affected individuals.

## De-identification

86. The Law Council considers that the discussion of de-identification and re-identification in section 3.3 of the Discussion Paper does not at this stage reflect the full complexity associated with AI systems. De-identification—which is sometimes able to be reversed—does not in and of itself discharge the obligations to comply with the privacy laws discussed above, particularly in regard to the lawful collection and use of personal information. De-identification should be used as a privacy enhancing tool, as discussed in the OAIC guide *De-identification and the Privacy Act*.[44]

## Other legislative initiatives

87. The Discussion Paper outlines some Government initiatives relevant to data, including the Consumer Data Right.[45] The Law Council cautions that the introduction of the Consumer Data Right is unlikely to change the Australian privacy landscape or provide effective privacy protections and safeguards in the AI context.[46] This is because consumers already have a right to access their information under APP 12. The privacy obligations under the Consumer Data Right largely replicate existing obligations under the *Privacy Act*. The effect of the Consumer Data Right is to provide

[41] D Dawson et al, 'Artificial Intelligence: Australia's Ethics Framework' (Discussion Paper, Data61 CSIRO, 2019) 29.
[42] *Privacy Act 1988* (Cth) Part IIIC.
[43] Ibid, s 26WE(1).
[44] Office of the Australian Information Commissioner, *De-identification and the Privacy Act* (March 2018) <https://www.oaic.gov.au/agencies-and-organisations/guides/de-identification-and-the-privacy-act>.
[45] D Dawson et al, 'Artificial Intelligence: Australia's Ethics Framework' (Discussion Paper, Data61 CSIRO, 2019) 30.
[46] Further detail of the Law Council's position regarding the Consumer Data Right is available in: Law Council of Australia, 'Treasury Laws Amendment (Consumer Data Right) Bill 2018' (27 February 2019), <https://www.lawcouncil.asn.au/resources/submissions/treasury-laws-amendment-consumer-data-right-bill-2018>.

another mechanism for businesses to share customer information and use it for a range of purposes, including in AI systems.

88. However, there have been numerous other recent Government initiatives that aim to enhance privacy protections, and which the Law Council believes will have an impact in the AI context. In particular, the Law Council notes the privacy-related recommendation of the Australian Competition and Consumer Commission's preliminary report on the Digital Platforms Inquiry[47] and the Attorney General's proposed Privacy Act amendments in anticipation of those recommendations.[48]

## Importance and benefit of a charter of rights

89. The Law Council notes that despite numerous recommendations at both the federal[49] and state[50] levels, there is currently no 'right to privacy' in Australia. Mindful that the right to freedom from arbitrary or unlawful interference in privacy is a universal human right,[51] the Law Council takes this opportunity to restate its established policy position that

> *every treaty to which Australia is party is binding upon it, and must be performed by it in good faith; …and that Australia is bound to comply with their provisions and to implement them domestically.*[52]

90. The Law Council supports the development of a charter or bill of rights at the federal level,[53] encompassing the right to privacy as a fundamental right and a range of other rights which may be impacted by the application of AI. This position was clearly stated in the Law Council's 'Call to Parties 2019'.[54] While the discussion in regard to a potential charter of rights is part of a much broader topic that falls outside the scope of this submission, the Law Council notes that a charter of rights would provide a beneficial ethical guide for the development of AI systems, and against which to assess proposed or actual applications.

**Recommendation:**

- **Careful consideration be given to the potential for a federal Charter of Rights to provide an ethical foundation for the development and implementation of AI systems.**

[47] The Australian Competition and Consumer Commission, *Digital Platforms Inquiry*, Preliminary Report (2018).

[48] Attorney-General of Australia, 'Tougher Penalties to Keep Australians Safe Online' (Media Release, 24 March 2019).

[49] Australian Law Reform Commission, *For Your Information: Australian Privacy Law and Practice,* Report No 108 (2008); Australian Law Reform Commission, *Serious Invasions of Privacy in the Digital Era Final Report*, Report No 123 (2014).

[50] See for example, Parliament of NSW Standing Committee on Law and Justice, *Remedies for the Serious Invasion of Privacy in New South Wales,* Report No 57 (2016).

[51] *International Covenant on Civil and Political Rights*, opened for signature 16 December 1966, 999 UNTS 171, art 17.

[52] Law Council of Australia, 'Policy Statement on Human Rights and the Legal Profession' (May 2017), <https://www.lawcouncil.asn.au/docs/bdc8f95f-eca0-e811-93fc-005056be13b5/Human%20Rights%20Policy.pdf>.

[53] Law Council of Australia, 'Policy Statement: A Charter protecting the rights of all Australians' (29 November 2008), 2, <https://www.lawcouncil.asn.au/docs/369b4912-cd39-e711-93fb-005056be13b5/081129-Policy-Statement-Bill-of-Rights.pdf>.

[54] Law Council of Australia, '2019 Federal Election Call to Parties' (May 2019), <https://www.lawcouncil.asn.au/resources/publications/call-to-parties-2019>.

## Principle 5: fairness

91. The Law Council considers that this principle would benefit from the provision of additional detail, and suggests that the EC *Ethics Guidelines for Trustworthy Artificial Intelligence* provide a useful model.[55] In particular, the Law Council suggests that a more thorough explanation of the implementation of fairness, including with respect to discrimination and unfair bias, be included.

92. The Law Council notes that trained AI systems have the potential to be manipulated through algorithmic bias, if the limitations and possibilities of AI, machine learning and related systems are poorly applied or understood. As has been noted by the LIV:

> *The lack of diversity and inclusivity in the design of AI systems is a key concern, as such systems may reinforce discrimination and prejudices while having an appearance of objectivity*.[56]

93. The Law Council suggests that fairness principles be defined as protecting against discrimination and unfair bias, including against vulnerable persons and groups. This should encompass women, LGBTQI people, people with disability, culturally and linguistically diverse minorities, children, workers, consumers, and any other groups at risk of marginalisation or whose rights may be compromised.

## Principle 6: transparency and explainability

94. The Law Council notes that there is a lack of understanding about the outcomes of AI, machine learning and related systems. These outcomes are not transparent or understandable, in some cases even by people with advanced training in the area. In many instances, not only the highly technical nature of the systems, but also the protection of proprietary algorithms and processes ensures that this continues to be the case.

95. The Law Council therefore considers that a balance must be struck between the protection of intellectual property and other commercially sensitive proprietary information, so as not to stifle development and innovation, and the requirement for sufficient transparency in decision-making processes to allow for review. Where possible, the data subject should be provided with enough information to understand the key parameters and factors taken into account in a decision affecting them. Where this level of transparency is not possible, the Law Council proposes that a confidential oversight relationship with a regulator be established, allowing for independent review of AI-enabled decisions without requiring the public release of commercial data. The Law Council notes that some precedents exist for a relationship of this sort, including certain functions of the Inspector-General of Intelligence and Security and the Australian Securities and Investments Commission.

## Principle 7: contestability

96. The Law Council is concerned that there can be no proper form of contestability without enforceability. Accordingly, as submitted in relation to Principle 3 above, the

[55] *Ethics Guidelines for Trustworthy AI*, Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission (8 April 2019), 12, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

[56] Law Institute of Victoria, 'Consultation on Artificial Intelligence: Australia's Ethics Framework' (24 May 2019), 5.

Law Council recommends that an ethical and regulatory framework be implemented formally so as to provide for enforceability.

97. The Law Council suggests a clear process flow be established to contest decisions or results of the operation of AI systems. This process, in the case of private sector organisations, could provide an internal dispute resolution mechanism at first instance, following which referral would be available to an independent regulator empowered to enforce the relevant ethical principles and regulations. The role of the courts should not be excluded by the use of privative or arbitration clauses. In the case of government and public organisations, internal review could precede referral to a regulator such as, for example, an AI Ombudsman. Judicial review should be available. The Administrative Review Council (**ARC**) has proposed 'Best-practice principles for automated assistance in administrative decision making',[57] which contain practical steps to implement appropriate governance and review processes.

98. The Law Council also suggests that training be provided to legal practitioners on how to challenge AI processes and decisions, including education on identification of discriminatory, unfair, incidental and unconscious bias in AI programming. There should also be increased interaction between programmers and lawyers during construction and renovation of algorithmic systems.

## Principle 8: accountability

99. This submission expresses support for the implementation of a regulatory body appropriately empowered to ensure accountability. However, the Law Council reiterates that clarity must be provided on where accountability and liability will rest. The question should be addressed of whether (similar to the legal personality accorded to a corporate entity) accountability will be permitted to rest with AI systems themselves, or, instead, with their proprietors or the individuals responsible for their design, production, implementation, maintenance or oversight. It should also be considered what responsibility will rest with relevant government authorities.

100. Further, the Law Council suggests additional provisions be created to ensure contractual acknowledgements, exclusions of liability and limitations of liability do not shield otherwise responsible individuals from liability for causing negligent or intentional harm flowing from AI products or services.

101. The Law Council has previously given consideration to accountability in the context of automated cars[58] and it is possible that guidance can be drawn from such analyses of specific applications. Careful consideration is required, however, to ensure that the full spectrum of contexts in which AI may be applied are adequately addressed.

[57] Administrative Review Council, 'Automated Assistance in Administrative Decision Making' (Report no. 46, 12 November 2004), vii - xi.

[58] Law Council of Australia, 'Motor Accident Injury Insurance & Automated Vehicles' (12 December 2018), <https://www.lawcouncil.asn.au/resources/submissions/motor-accident-injury-insurance-automated-vehicles>.

# Question 2: Do the principles put forward in the discussion paper sufficiently reflect the values of the Australian public?

## Values-based vs rights-based approach to AI ethics

102. The Discussion Paper refers to the 'values' of the Australian public in several instances but does not discuss how those values should be identified or how—given the diversity of the Australian population, within which different groups may place weight on different values—discrepancies or contradictions could be addressed. The Law Council considers that, if a 'values-based' approach were to be adopted to guide the development of AI, a consensus decision about the values of the Australian public would need to be underpinned by current empirical research.

103. The Law Council recommends that public consultation be undertaken in a way that would engage the members of the public who are likely to be affected by AI tools. This should take particular account of those groups likely to be affected more significantly by government AI decision-making activities, including economically and socially disadvantaged groups, the disabled and chronically ill, culturally and linguistically diverse people, people in regional, rural and remote locations, and older people.

**Recommendation:**

- **Public consultation should be undertaken with all groups, including vulnerable groups and minorities, who are likely to be affected by AI.**

104. Instead of a values-based approach to framing AI ethics, the Law Council refers again to its submission above[59] and recommends consideration of an objective, rights-based approach to framing AI ethics based on defined international human rights law obligations.

105. The EU offers a model in this respect and is progressing an approach that aligns AI ethics with fundamental rights. There is a growing body of scholarly research and case law addressing fundamental human rights, from Europe, Canada and Australia. The EC *Ethics Guidelines for Trustworthy Artificial Intelligence* identifies ethical principles and correlated values using an approach founded on fundamental rights.[60]

106. Until such time as Australia enacts a human rights act or charter able to inform a rights-based approach to the ethical framework to underpin AI, the Law Council recommends that a number of sources be used to inform the approach. Those sources should include the international human rights treaties and other instruments to which Australia is a party, other relevant provisions of international human rights

[59] See under 'Importance and benefit of a charter of rights'.

[60] *Ethics Guidelines for Trustworthy AI*, Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission (8 April 2019), 2, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

law, as well as other sources such as the *Charter of Fundamental Rights of the EU*,[61] from which guidance may be drawn.

**Recommendation:**

- **An ethical framework for AI should be rights-based and informed by provisions of international human rights law and precedents available in the EU and elsewhere.**

## Workers protections

107. The Law Council proposes, given the extensive and increasing impacts AI is likely to have on the labour market and its potential to lead to restructuring of workforces and associated redundancies, a further principle should be included to require consideration of workers' rights. The Law Council notes that the right to work is protected by the *International Covenant on Economic, Social and Cultural Rights*.[62] Policy development related to AI should incorporate consideration of employment protections—including access to retraining and reskilling for affected workers—as key issues deserving further attention. This will need to be considered in the context of the operation of the employee record exception under the Privacy Act.[63]

**Recommendation:**

- **The development of policy and principles concerning AI should incorporate consideration of employment protections, including access to retraining and reskilling for affected workers, as key issues.**

# Question 5: What other tools or support mechanisms would you need to be able to implement principles for ethical AI?

## Additional tools

108. The Law Council recommends that consideration be given to the following, in addition to the nine tools proposed by the Discussion Paper,[64] with the purpose being to incentivise, encourage and enforce adherence to the required standards:

    (a) legislative and regulatory requirements;[65]

[61] *Charter of Fundamental Rights of the European Union* [26 October 2012] OJ C 326/02.

[62] *International Covenant on Economic, Social and Cultural Rights*, opened for signature 16 December 1966, 993 UNTS 3, art 6.

[63] *Privacy Act 1988* (Cth), s 7B(3).

[64] D Dawson et al, 'Artificial Intelligence: Australia's Ethics Framework' (Discussion Paper, Data61 CSIRO, 2019) 8.

[65] Katie Miller, 'The Application of Administrative Law Principles to Technology-assisted Decision-Making' (2016) 86 *AIAL Forum* 2, 22.

(b) standards, such as those prepared by the Joint Technical Committee of the International Organisation for Standardisation (**ISO**) and International Electrotechnical Commission (**IEC**);[66]

(c) codes of conduct and codes of ethics;

(d) accreditation (subject to necessary ongoing education, training and certification to provide services, including for those providing independent advice, assessment and expertise);

(e) governance frameworks (additional organisations responsible for the development of ethical AI; independent ethics committees and panels to review for example, data used to train AI); and

(f) independent experts.[67]

## Responsible Innovation Organisation

109. In its submission to the Australian Human Rights Commission consultation on Artificial Intelligence: Governance and Leadership,[68] the Law Council expressed qualified support for the creation of a 'responsible innovation organisation' capable of providing guidance, leadership and mitigating risks, whilst also supporting capacity building and promoting innovation.[69] While the Law Council supports that proposal, it voices caution that the creation of a purpose-specified organisation may create a false sense of security or complacency.[70] Accordingly, it recommends that existing regulatory bodies, as well as key civil society actors, also be enabled to address issues presenting 'here and now', by developing 'the capability to recognise, address, manage and mitigate risks and harms arising through data-driven decision-making and technological innovation, including advances in AI.'[71]

**Recommendation:**

- **The capability of existing regulatory bodies and key civil society actors should be increased to complement the role of an AI regulator.**

110. A responsible innovation organisation may be well placed to assist with the practical implementation and support mechanisms needed to implement principles for ethical AI. An example is the UK's Centre for Data Ethics and Innovation (**CDEI**). The CDEI's first Work Programme[72] and Strategy[73] state that it will analyse and anticipate the

[66] See: ISO/IEC JTC 1/SC 42 'Artificial Intelligence' < https://www.iso.org/committee/6794475.html>.
[67] See for example: D Hogan-Doran SC 'Computer says 'no': automation, algorithms and artificial intelligence in Government decision-making' (2017) *The Judicial Review* 1.
[68] See: <https://www.humanrights.gov.au/our-work/rights-and-freedoms/publications/artificial-intelligence-governance-and-leadership>.
[69] Law Council of Australia, 'Artificial Intelligence: Governance and Leadership' (18 March 2019), 3, <https://www.lawcouncil.asn.au/resources/submissions/artificial-intelligence-governance-and-leadership>.
[70] Ibid, 2.
[71] Ibid, 3.
[72] Centre for Data Ethics and Innovation, '2019/20 Work Programme', <https://www.gov.uk/government/publications/the-centre-for-data-ethics-and-innovation-cdei-2019-20-work-programme>.
[73] Centre for Data Ethics and Innovation, '2 Year Strategy', <https://www.gov.uk/government/publications/the-centre-for-data-ethics-and-innovation-cdei-2-year-strategy/centre-for-data-ethics-cdei-2-year-strategy>.

opportunities and risks posed by data-driven technology and will put forward practical and evidence-based advice to address them. It states its functions as follows.

(a) Analyse and Anticipate: will convene communities and expertise to provide an overview and insight of opportunities and risks, and review existing regulatory and governance frameworks to identify gaps. It will also carry out thematic projects to enable CDEI to explore live or urgent issues, drawing together lessons from existing work and setting out how they should be taken forward.

(b) Reviews: will identify and articulate best practice for the responsible use of data driven technology within specific sectors or for specific applications of technology. They will consider any gaps in governance and make recommendations to the government, as well as advice to regulators, creators and users of data-driven technology as to how those gaps should be addressed.[74]

## Question 6: Are there already best-practice models that you know of in related fields that can serve as a template to follow in the practical application of ethical AI?

111. The Law Council would, in particular, recommend the EU Ethics Guidelines for Trustworthy AI as a measured and practical resource, including the useful 'Pilot Version AI Assessment List'.[75]

112. In addition, the Law Council recommends that consideration be given to the following:

(a) the GDPR, which aims to strengthen and unify data protections for individuals within the EU. Specific protections and references to profiling of personal information are included, some of which refer to automated decision-making, and demand increased levels of protection for data subjects in recognition of the inherent risks associated with those systems;[76]

(b) the 'Toronto Declaration: Protecting the right to equality and non-discrimination in machine learning systems'[77] also seeks to address the human rights implications of AI and associated technologies; and

(c) the international IEEE Guidelines and Papers, which provide further important guidance on the development of a human rights-based approach to AI.[78]

[74]<https://www.gov.uk/government/groups/centre-for-data-ethics-and-innovation-cdei>.

[75] *Ethics Guidelines for Trustworthy AI*, Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission, (8 April 2019), chapter 3, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

[76] *Parliament and Council Regulation EU/2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)* [2016] OJ L 119/1, article 22.

[77] Access Now and Amnesty International, 'The Toronto Declaration: Protecting the right to equality and non-discrimination in machine learning systems' (2018), <https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf>.

[78] IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, *Ethically Aligned Design: a Vision for Prioritising Human Wellbeing with Autonomous and Intelligent Systems*, (First Edition, December 2017), <https://ethicsinaction.ieee.org>.

# Question 7: Are there additional ethical issues related to AI that have not been raised in the discussion paper? What are they and why are they important?

## Administrative law principles

113. The Law Council is concerned that the Discussion Paper conflates public sector and private sector AI decision-making issues, and insufficiently distinguishes some of the different ethical frameworks which apply. For instance, while the Discussion Paper notes that key government agencies are reportedly increasingly making AI-assisted decisions (including with respect to social services, health, education and training, immigration and border protection, agriculture and water resources and veterans' affairs), there is little analysis in the Discussion Paper of the scope or appropriateness of existing processes, having regard to key values underpinning administrative law system concerning lawfulness, fairness, rationality, openness and transparency and rationality.

114. As the ARC noted in 2004, 'these values are coincident with concepts of administrative justice, which are seen to include four basic requirements for just decision making in a society governed by the rule of law—lawfulness, fairness, rationality and intelligibility.'[79] Its report identified a range of best-practice principles to ensure that decision-making done by or with the assistance of automated systems is consistent with these administrative law values. There are few references to either administrative law or the ARC's report in the Discussion Paper. While it does refer to a best practice guide developed in 2003 and updated in 2007 by the Department of Finance Working Group regarding Automated Assistance in Administrative Decision Making, the extent to which this guide remains fit for purpose, is aligned with the ARC's proposals or is followed by Commonwealth agencies, is unclear.

115. The Law Council considers that the lack of engagement by the Discussion Paper with key administrative law principles and government decision-making processes is concerning—given that the kinds of decisions which are encompassed by the above portfolios are fundamentally important to many Australians and the scale of decision-making potentially very broad. It recommends that further specific work should be undertaken, drawing on the advice of independent administrative law experts. This should focus on how existing administrative decision-making utilises AI, the extent to which this upholds key administrative law principles having regard to the ARC's recommendations, and whether further guidance or frameworks are necessary to guide the adoption of AI by government agencies into the future. Should the Commonwealth wish to lead the discussion on ethical frameworks concerning AI—and in the Law Council's view, it is vital that it does so—it is critical that it consider the extent to which its own AI-assisted decision-making processes are best practice.

**Recommendation:**

- **Further specific work should be undertaken on questions relating to the interaction of administrative law principles with AI in public sector decision-making processes. Consideration should be given to whether further**

[79] Administrative Review Council, 'Automated Assistance in Administrative Decision Making' (Report no. 46, 12 November 2004), 3.

**guidance or frameworks are necessary to guide adoption of AI by government agencies.**

## Resolving tensions between the Principles

116. Although the Discussion Paper does not expressly place the Principles in order of priority, the fact that they appear in a numbered list may suggest relative priority. The LSWA has submitted that simultaneous implementation of all Principles 'may lead to some incompatibility in the future',[80] and that guidance will therefore be necessary in any event to establish an effective hierarchy of principles in such instances.

117. The Law Council suggests that the international human rights framework may provide some assistance in this respect. While all human rights are universal, only certain rights are absolute, while others are subject to limitations[81] and principles of legitimate purpose, necessity, reasonableness and proportionality are used to address tensions between rights.

118. It is inevitable that tensions will arise between principles applied to AI. Examples might include situations where the law appears to permit harm between the privacy and contestability principles, or where an entity responsible for AI refuses to discover evidence key to a successful contest on the basis of inadvertent disclosure of other users' personal data or the entity's commercial data. Another example could be seen in the context of 'preventative policing', where increasingly restrictive actions are applied to given individuals on the basis of AI-calculated assessments of risk of offending.

119. The Law Council considers that it is important to avoid a utilitarian response to such tensions. As the LIV has submitted, this is necessary because 'there are certain human rights which cannot be rationalised away on a "generates net-benefit" utilitarian basis (eg do not kill).'[82] This submission recommends[83] that the current first Principle, 'generation of net benefits', should be subordinate to the second principle, 'doing no harm'.

# General observations

120. The Law Council offers the following general comments in response to the Discussion Paper.

## Role of ethicists and other specialists in developing an AI ethics framework

121. The Law Council considers that convening a multi-disciplinary group including ethicists, social scientists, data scientists, privacy specialists and lawyers, would

[80] Law Society of Western Australia, 'Consultation on the ethics of artificial intelligence: Department of Industry, Innovation and Science' (15 May 2019), 1.

[81] See, for example, limitations on the right to freedom of expression, *International Covenant on Civil and Political Rights*, opened for signature 16 December 1966, 999 UNTS 171, art 19.

[82] Law Institute of Victoria, 'Consultation on Artificial Intelligence: Australia's Ethics Framework' (24 May 2019), 2.

[83] See under 'Principle 1: generates net-benefits'.

provide depth to specialist areas and would greatly benefit the development of an AI ethics framework.

## Significant and insignificant risk

122. The Law Council recommends caution in linking 'numbers of people affected' to the assessment of risk, as per the table on page 64 of the Discussion Paper. While the rationale is understood for considering how broadly the impacts of a particular occurrence could be felt, the assessment of risk should nonetheless recognise that that an impact felt by only one person could regardless have 'major' or 'critical' consequences, potentially including fundamental rights breaches experienced by that person. An insignificant or minor risk could only be said to arise if its consequences would also be minor.

**Recommendation:**

- **Caution should be exercised in linking levels of risk merely to numbers of persons affected, having regard to the potential for significant impacts on relatively few people.**

## International initiatives to build a global framework

123. The Law Council notes that work is being undertaken by numerous actors within Australia to consider the development and regulation of AI, including the Australian Human Rights Commission and the Australian Government Digital Transformation Agency, as well as DIIS, CSIRO and others. The Law Council welcomes and supports collaboration and cooperative approaches to the development of an AI ethics framework.

124. Internationally, the Law Council observes that AI will have far-reaching, supranational impacts, and that neither the use nor regulation of AI systems is likely to be limited to a single national jurisdiction. The role of convening and thought leadership bodies, such as the World Economic Forum (**WEF**),[84] and of international standard setting bodies, such as the Joint Technical Committee 1 of the ISO and IEC, [85] is considered key, and the Law Council encourages Australia's active support.

125. Given the range of international initiatives, the Law Council considers that, once consensus is reached on an approach in Australia, Australia should take a more active role and engage with like-minded international initiatives and organisations to develop a collaborative approach to AI development and the embedding of shared principles, safeguards and ethical standards.

## Data Governance

126. While the central role of data in AI is acknowledged in the Discussion Paper, the Law Council recommends that it be more fully explored in developing an AI ethics

[84] See: World Economic Forum, *AI Governance: A Holistic Approach to Implement Ethics into AI* (White Paper, January 2019), <https://weforum.my.salesforce.com/sfc/p/#b0000000GycE/a/0X000000cPl1/i.8ZWL2HIR_kAnvckyqVA.nVVgrWIS4LCM1ueGy.gBc>.

[85] See: <https://www.iso.org/isoiec-jtc-1.html>.

framework. Effective data governance requires a bottom-up review of processes, systems and methodologies. This is to ensure that principles are properly embedded in all aspects of given systems, and that their implementation is consistent, robust (taking due account of the range of contexts and circumstances in which AI is being deployed and the variety of decisions which may be taken or influenced by it), replicable and appropriately transparent. Appropriate reporting, oversight and feedback mechanisms should be integrated.

**Recommendation:**

- **Further consideration should be given to the role of data governance in AI.**

127. The Law Council considers that additional attention should be given to the nature and quality of the data used to develop and train AI and the resulting risk of bias. This risk is famously illustrated by the flawed operation of an AI system, deployed in the United States to assess reoffending risks and inform decision-making by criminal courts about bail and sentences, which was found to be biased systematically against black people.[86] The bias demonstrated that system was found to have been caused by the AI's adoption and amplification of biases in the data it received.

128. The Law Council considers that standards and training need to be consistently applied, to ensure the maintenance of data quality and to monitor for, and correct, processes, data sources, and other variables which may lead to incorrect outcomes.

129. The Law Council recommends that an onus be placed on AI systems to demonstrate the accuracy of their decisions and other outputs, rather than allowing an assumption to develop of the correctness or infallibility of AI-derived or supported decisions.

**Recommendation:**

- **In the development of legislative and regulatory frameworks concerning AI, consideration should be given to placing an onus on AI systems to demonstrate accuracy.**

130. Future consideration of this issue should address the importance of assessing the data sets being captured and analysed to determine:

(a) whether there is a potential breach of an individual's privacy (for example, if an anonymised data set is so small that individuals are identifiable);

(b) if biases are likely to be entrenched by using the data sets; and

(c) whether AI is the right tool to analyse the data.

[86] Julia Angwin et al, 'Machine Bias', *Pro Publica* (23 May 2016), <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.