THE LAW SOCIETY OF NEW SOUTH WALES
young LAWYERS

Submission on the Human Rights Commission's Discussion Paper - *Human Rights and Technology*

14 April 2020

Australian Human Rights Commission
tech@humanrights.gov.au

Contact: **David Edney**
President, NSW Young Lawyers

Ashleigh Fehrenbach
Chair, NSW Young Lawyers Communications, Entertainment and Technology Committee

Managing Editor: Olivia Irvine

Contributors: Onur Saygin, Sophia Urlich, Ravi Nayyar

THE LAW SOCIETY OF NEW SOUTH WALES
young LAWYERS

The NSW Young Lawyers Communications, Entertainment and Technology Committee (Committee) makes the following submission in response to the *Human Rights and Technology* Discussion Paper

NSW Young Lawyers

NSW Young Lawyers is a division of The Law Society of New South Wales. NSW Young Lawyers supports practitioners in their professional and career development in numerous ways, including by encouraging active participation in its 15 separate committees, each dedicated to particular areas of practice. Membership is automatic for all NSW lawyers (solicitors and barristers) under 36 years and/or in their first five years of practice, as well as law students. NSW Young Lawyers currently has over 15,000 members.

The Communications, Entertainment and Technology Law Committee of NSW Young Lawyers aims to serve the interests of lawyers, law students and other members of the community concerned with areas of law relating to information and communication technology (including technology affecting legal practice), intellectual property, advertising and consumer protection, confidential information and privacy, entertainment, and the media. As innovation inevitably challenges custom, the CET Committee promotes forward thinking, particularly about the shape of the law and the legal profession.

Summary of Recommendations

The NSW Young Lawyers Communications, Entertainment and Technology Law Committee (**the CET Committee**) welcomes the opportunity to comment on *Human Rights and Technology* Discussion Paper (**Discussion Paper**) on behalf of NSW Young Lawyers.

The Committee has responded to the selected questions outlined below and have otherwise not made submissions on the remaining questions. The Committee has outlined considerations that it recommends the Australian Human Rights Commission (**the Commission**) take into account when reviewing these issues. The Committee hopes that these considerations provide helpful guidance to the Committee in conducting this review.

1. Proposal 4: 'The Australian Government should introduce a statutory cause of action for serious invasion of privacy'.

    i. The CET Committee submits that private sector decision-making should be subject to such an action, and

    ii. That such an action much be formulated with consideration to:

        a. the difficulty of retrieving information that has been breached,

b. the emotional and social harm to the purported victim in the situation, to be assessed on the balance of probabilities, and

c. the potential longevity, or delayed crystallisation of potential harm, should be accounted for in the formulation.

2. The CET Committee next considers Proposal 11: 'The Australian Government should introduce a legal moratorium on the use of facial recognition technology in decision-making that has a legal, or similarly significant, effect for individuals, until an appropriate legal framework has been put in place.'

   i. The CET Committee submits the uniquity of facial recognition and other biometric data requires careful legal treatment, and

   ii. The CET Committee supports the moratorium on the technology.

3. Question C: 'Does Australian law need to be reformed to make it easier to assess the lawfulness of an AI-informed decision-making system, by providing better access to technical information used in AI-informed decision-making systems such as algorithms?'

   i. The CET Committee submits that reform will be required in an administrative law context, and

   ii. That accessibility required to determine lawfulness will include technical information, institutional processes, and data sources, particularly where a system relies on personal or private information.

4. Proposal 7: The Australian Government should introduce legislation regarding the explainability of AI-informed decision-making. This legislation should make clear that, if an individual would have been entitled to an explanation of the decision were it not made using AI, the individual should be able to demand: (a) a non-technical explanation of the AI-informed decision, which would be comprehensible by a lay person, and (b) a technical explanation of the AI-informed decision that can be assessed and validated by a person with relevant technical expertise…'

   i. The CET Committee submits that it supports the introduction of a right of explainability;

   ii. That some proposed changes are necessary to existing administrative review to ensure the efficacy of such a right;

   iii. The importance and correlated expectation of human oversight, and

   iv. Administrative law protections should be favoured over the protection of trade secrecy.

5. Question B: 'Where a person is responsible for an AI-informed decision and the person does not provide a reasonable explanation for that decision, should Australian law impose a rebuttable presumption that the decision was not lawfully made?'

   i. The CET Committee supports the introduction of rebuttable presumption due to the significance and normative consistency within administrative law.

6. Question D: 'How should Australian law require or encourage the intervention by human decision-makers in the process of AI-informed decision-making?'

    i. The CET Committee submits that accountability and explainability are key principles to successful AI, and require at a minimum, the capacity for human intervention;

    ii. That a proposed list of 'Requirements' should be used to evaluate the accountability and explainability of a decision-making system;

    iii. That any introduction of legislation in this area would need to be technology and sector neutral, and

    iv. That minimum requirements are necessary to ensure ongoing commitment to transparency beyond the developer phase.

7. Question F: 'What should be the key features of a regulatory sandbox to test AI-informed decision-making systems for compliance with Human Rights? In particular…'

    i. The CET Committee submits that a regulatory sandbox may be a significant measure to respond to the rapid development of the sector;

    ii. That technological neutrality, or a multi-sector specific system would be necessary to ensure universal Human Rights;

    iii. That in addition to Human Rights, key areas will include privacy, data governance, and cyber security.

    iv. Regarding criteria of entry, the CET Committee submits that requirements of cyber security expertise and solvency will be necessary to preserve efficacy;

    v. That a multi-disciplinary, multi-stakeholder approach is necessary;

    vi. That the aims of the sandbox, on balance, should supersede other legal rights such as trade secrets, and

    vii. That the sandbox should be evaluated by reference to the Commission's proposed 'Requirements' at minimum;

8. Proposal 10: 'The Australian Government should introduce legislation that creates a rebuttable presumption that the legal person who deploys an AI-informed decision-making system is legally liable for the use of the system.'

    i. The CET Committee supports such a right as the logical and practical extension of accountability.

1. Privacy

1. The impact on Human Rights of inadequate controls on the storage, transmission and use of sensitive personal information about individuals is not only of concern in the domain of government decision-makers. It is also the concern of decision-makers in private entities (ie companies), whose adverse decisions can equally cause detriment to individuals.[1]

2. Where private entities use AI to process data about an individual in a way which discriminates against that individual, for example by informing whether to provide goods or services to the individual, those entities should be accountable to explain the source of the data,[2] what the data is and provide a basic explanation of the process by which the decision was made.[3]

3. As more sophisticated AI and models of data analysis are developed, unless adequately regulated, private entities may increasingly be able to covertly use data collected without the individual's knowledge (or with their knowledge but without their full understanding)[4] to discriminate against that individual and in doing so, impact an individual's ability to obtain essentials such as employment,[5] housing,[6] loans[7] and healthcare.[8] Additionally, an individual's ability to obtain insurance and membership to various organisations could similarly be impacted.

4. Furthermore, if unchecked, there may be flow on effects and an entrenchment of biases adverse to an individual. For example, if a data source that has been consulted to make a decision is erroneous and as a consequence an adverse decision is made, the fact that this decision has been made may then itself constitute a piece of information which other private entities may rely on in their decision-making. This would create a cascading series of adverse decisions.[9]

---

[1] For example, an individual's creditworthiness: Andrew Selbst and Solon Barocas, 'The Intuitive Appeal of Explainable Machines' (2018) 87(3) *Fordham Law Review* 1085, 1102-1104; Christian Ernst, 'Artificial Intelligence and Autonomy: Self Determination in the Age of Automated Systems' in Thomas Wischmeyer and Timo Rademacher (ed), *Regulating Artificial Intelligence* (Springer, 2020) 53, 55 [4]-[5].
[2] Moritz Hennemann, 'Artificial Intelligence and Competition Law' in Thomas Wischmeyer and Timo Rademacher (ed), *Regulating Artificial Intelligence* (Springer, 2020) 361, 368 [16], 369 [18].
[3] Selbst (n 1) 1102-1104.
[4] Eg. Gabriele Bucholtz, 'Artificial Intelligence and Legal Tech: Challenges to the Rule of Law' in Thomas Wischmeyer and Timo Rademacher (ed), *Regulating Artificial Intelligence* (Springer, 2020) 175, 178[5]; Steven Seybold, 'Somebody's Watching Me: Civilian Oversight of Data-Collection Technologies' (2015) 93(4) *Texas Law Review* 1029, 1040; Selbst (n 1) 1101.
[5] Alexander Tischbirek, 'Artificial Intelligence and Discrimination: Discriminating Against Discriminatory Systems' in Thomas Wischmeyer and Timo Rademacher (ed), *Regulating Artificial Intelligence* (Springer, 2020) 103, 112 -113 [25]-[26].
[6] Bryan Casey, Ashkon Farhangi, and Roland Vogt, 'Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise' (2019) 34(1) *Berkeley Technology Law Journal* 143, 147.
[7] Selbst (n 1) 1102-1104; Ernst (n 1) 55 [4]-[5].
[8] Fruzsina Molnar-Gabor, 'Artificial Intelligence in Healthcare: Doctors, Patients and Liabilities' in Thomas Wischmeyer and Timo Rademacher (ed), *Regulating Artificial Intelligence* (Springer, 2020) 337, 337.
[9] For example, flow on effects of a creditworthiness decision: Selbst (n 1) 1102-1104.

5. Another important consideration in the context of Human Rights and technology is use of data collected (and potentially processed by AI) about an individual by a private entity to then set a unique price point for that individual.[10] For example, the IP address combined with behavioural purchase information about an individual may allow for a website that aggregates flights to set a price and present it in a way which maximises its profit against a specific individual. Presently, such a business is not accountable to explain the technology they use to tailor the price point to their customer. Indeed, to do so is likely to render the technology less effective, however, the impact on the individual may be a slight, but important loss of autonomy.[11]

6. Where a private entity uses a dataset which contains sensitive data about an individual and makes a decision to discriminate in the goods or services they will provide to the individual, the individual should at least have the right to know what specific information about them was consulted in making the decision. They should also be able to evaluate that information themselves (either directly or by providing it to a third party for analysis).

Proposal 4

7. **Summary**: The CET Committee supports Proposal 4 and submits that private sector decision-making should be subject to such an action, and that such an action must be formulated with consideration to:

    i. the difficulty of retrieving information that has been breached,

    ii. the emotional and social harm to the purported victim in the situation, to be assessed on the balance of probabilities, and

    iii. the potential longevity, or delayed crystallisation of potential harm, should be accounted for in the formulation.

8. The CET Committee suggests that some of the key considerations in formulating any legislation in this area should be that:

    i. Private information, once exposed because of a serious invasion of privacy, is difficult to retrieve or erase and often permanently remains in the public domain. The burden of proving that adequate mitigative steps were taken should lie with the entity responsible for the serious invasion of privacy. This could be realised through inclusion of such a rebuttable presumption as part of any legislation.

    ii. Harm that flows from a serious invasion of privacy may not only be financial but also emotional or reputational. In acknowledgement of this, remedies should include the ability to order a party that has committed a serious invasion of privacy to take steps to mitigate emotional or reputational damage to the extent this is possible, for example, to destroy the offending material and any copies. Furthermore, the information in question should be made accessible to the individual whose privacy has been breached.

---

[10] Ernst (n 1) 66[37].
[11] Ibid, 66 [37].

iii. In many cases there are difficulties associated with showing a connection between a serious invasion of privacy and harm to the relevant individual. For example, harm to an individual whose sensitive personal information is collected without their consent and by no fault of their own may not immediately suffer harm. However, they may suffer harm many years later when a public or private decision-maker consults a dataset informed by that sensitive personal information.[12]

9. Any cause of action created should extend liability to downstream transactions by third parties[13] who **know or ought to reasonably know** that the information they are dealing with has come into their possession because of a serious invasion of privacy and do not respond accordingly. For example, we can compare this to the responsibility of intermediary platforms to remove content which is cited to be a breach of intellectual property, or to a more directly problematic behaviour, such as selling or purchasing the misappropriated private information. The exact nature of the liability for third parties should reflect other areas of law, such as torts, with degrees of liability and damages in proportion to culpability.

10. The CET Committee proposes further that serious consideration should be given to whether there should be a standalone punitive element to any such cause of action. This would ensure that, notwithstanding a lack of harm to the individual whose privacy has been breached, if the event is of the requisite seriousness, a financial or practical consequence should flow to the person or entity responsible for the serious invasion of privacy.

Proposal 11

11. **Summary**: The CET Committee supports Proposal 11 and submits the uniquity of facial recognition and other biometric data requires careful legal treatment and supports the moratorium on the technology.

12. Ubiquitous collection of, and reliance on, any biometric information for the purpose of making decisions that affect the rights of citizens, absent adequate consequences for misuse of the power, creates space for the significant infringement of Human Rights.

13. Biometric data should be considered “sensitive information” per the meaning under The General Data Protection Regulation 2016/679 and should be presumed to be excluded from most automated decisions.[14] A lack of transparency around this kind of ‘sensitive’ data may allow for both knowing and inadvertent discrimination.

14. This is in part because collection of this kind will often be partly or fully covert: if the individual whose data is collected does not have access to information about the kind or quality of data that is being collected about them, they will be unable to inform their actions on the basis of that information.[15] This means that the individual is less likely to “censor their own activities”[16] which may give commercial or government bodies aggregate data based on their everyday behaviours without their knowledge. This can allow for

[12] Hennemann (n 2).
[13] Hennemann (n 2) 380-381 [41]-[42].
[14] Buchholtz (n 4) 189 [28]-[29].
[15] Seybold (n 4) 1038.
[16] Ibid 1038.

inferences on everything from buying habits to political and religious beliefs.[17] Without oversight of how these inferences impact decisions, any discrimination which may occur may not be evaluated.

15. Where an individual *is* informed of this collection, they may not be aware of *how* this information impacted the decision that was made (most problematic where an adverse decision results). This may mean that rather than failing to censor behaviour, they may refrain from protected activities, such as political protest, for fear of how this information may be used.[18]

16. This will be particularly problematic where the use of the technology relates to intelligence or police investigation, which, by its nature, is excluded from most existing transparency requirements. Such lack in transparency is of substantial importance where that technology is potentially discriminatory and where reliance on such technology impacts the rights of communities and individuals.[19] For example, the NSW Police Force's use of Suspect Target Management Plan is a non-transparent algorithm which enables police to track persons of interest, or persons related to subjects. An analysis of those targeted by the system found disproportionate and overwhelming focus on young, Aboriginal Australian subjects,[20] who, as a result of the targeting, may have experienced a greatly increased number of police encounters which may not have been justified by the legal standard of reasonable suspicion.[21] The need for transparency here also needs to be balanced with the protection of the public at large, and the practicality of law enforcement bodies being able to complete their duties in a pragmatic manner.

17. AI systems thrive on data, meaning there is a strong incentive for information to be stored in an indiscriminate manner, particularly as technology advances allow us to store more data for longer.[22] The static nature of facial information, which is similar in nature to a person's fingerprints, dental evidence, iris scans or DNA, warrants extra protections. The enormous specificity of this data makes it uniquely valuable in a financial sense to any government or commercial organisation that may use this information.

18. This raises important issues of the security of that information, as, in the hands of criminals, biometric information, and detailed logs of a person's whereabouts and behavioural patterns exposes them to enhanced risks relating to identity theft and scamming. Careful considerations about how this information can be stored and transmitted and which organisations should be allowed to process and share this information are a necessary step as this trend evolves.[23] Public organisations which use biometric information or derivative data to inform decisions should not be able to defer the responsibility of protecting that information to private service providers.[24]

---

[17] Ibid 1038.
[18] Ibid 1039.
[19] Seybold (n 4) 1039: The author notes the example of a New York Police Department secret program targeting Muslim communities.
[20] Vicki Sentas and Camilla Pandolfini, *Policing Young People in NSW: A Study of the Suspect Target Management Plan* (Report, Youth Justice Coalition NSW 2017).
[21] Timo Rademacher, 'Artificial Intelligence and Law Enforcement' in Thomas Wischmeyer and Timo Rademacher (ed), *Regulating Artificial Intelligence* (Springer, 2020) 225, 45[33].
[22] Seybold (n 4) 1036.
[23] Seybold (n 4) 1037; Hoffman Riem, 10[29]
[24] Wolfgang Hoffman-Riem, 'Artificial Intelligence as a Challenge for Law and Regulation' in Thomas Wischmeyer and Timo Rademacher (ed), *Regulating Artificial Intelligence* (Springer, 2020) 1, 8 [25].

19. The simplicity with which personal information can be collected, transmitted, stored and used as a direct result of new technologies should not be the factor which determines whether such technologies should be allowed to go unchecked within society. The sensitivity of this kind of data should require greater accountability, not less. Pending an appropriate legal framework, a moratorium should be in place.

2. Decision-Making and Public Access

20. The transparency, accountability and accessibility of an AI or algorithmic system is essential to the effective legal and normative evaluation of AI technologies and the decisions they produce.

21. Besides its use in entertainment, employment, banking and social media, AI-informed or algorithmically-supported decision-making is already a key part of government decision processes.[25] As reliance on emerging technologies by government organisations expands, we must consider how decisions made with the support of, or independently by, AI or algorithms are, and should be, regulated by administrative law.

22. Individuals subject to these decisions will have questions about the conclusions reached by AI-informed decision-making systems. These questions may get to the basis of a decision, or the mechanics of the decision-making process, and are likely to demand meaningful explanations for the decisions reached.[26]

23. AI-informed decision-making systems come with the risk of bearing 'wrong' (normatively or factually) decisions, which may be difficult to detect and explain. These decisions may be based on inaccurate, biased or incomplete information and data,[27] and, importantly, may adversely affect or infringe the Human Rights of the individuals or groups to which they apply.[28] This may cause or contribute to real-world harms.[29]

24. The CET Committee submits that a guiding principle in the quest to regulate AI-informed and algorithmic decision-making systems, and the decisions they produce, should be to ensure that governments and the legal persons behind them are properly accountable for any resulting negative consequences[30] through regulation and liability.

---

[25] *Pintarch v Deputy Commissioner of Taxation* [2018] FCAFC 79 [47].
[26] Selbst (n 1) 1118.
[27] Yoan Hermstrüwer, 'Artificial Intelligence and Administrative Decisions Under Uncertainty' in Thomas Wischmeyer and Timo Rademacher (ed), *Regulating Artificial Intelligence* (Springer, 2020) 199, 213 [44]; Tischberek (n 5) 104 [4].
[28] Tischberek (n 5) 105 [5].
[29] Hermstrüwer (n 26) 213 [47] balancing false negatives against false positives, 216 [56], [60] ("gaming").
[30] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (ed), *Regulating Artificial Intelligence* (Springer, 2020) 75, 78; Hoffman-Riem (n 23) 10 [29]-[30].

25. The CET Committee further submits that appropriate human oversight of, responsibility for, and intervention in AI-informed decision-making is extremely important. This is especially the case where such decision-making is fully automated.[31]

26. Australian law will need to determine who will be legally liable for AI-informed decisions, fully automated or otherwise. There will also be an increased need for individuals in the Australian legal system who are properly equipped to explain and defend AI-informed decision-making systems and the decisions they produce.[32]

Question C

27. **Summary**: The CET Committee submits that:

    i. reform will be required in an administrative law context; and

    ii. that accessibility required to determine lawfulness will include technical information, institutional processes, and data sources, particularly where a system relies on personal or private information.

28. The CET Committee submits that law reform will be necessary to appropriately assess the lawfulness of AI-informed decision-making systems and the decisions they produce, by providing access to:

    - Technical information used in the development of the system, (such as the model, "values and constraints that shape…conceptualization", how these values shaped the machine learning, and "how outputs…inform final decisions");[33]

    - Information on "institutional process" around system outputs, including how they are used, and "what role discretion play[s];"[34] and,

    - The personal/private data upon which such decisions are based.[35]

29. The CET Committee notes that the appropriateness of access will be factually dependent and will not require public disclosure in all cases. For example, disclosure in a civil action versus a public reporting requirement such as publication of an impact statement.

---

[31] Hermstrüwer (n 26) 219-220, [70]-[71]; Selbst (n 1) 1139.
[32] Wsichmeyer (n 29) 95 [43].
[33] Selbst (n 1) 1130.
[34] Ibid 1132.
[35]Tischberek (n 5) 117; Nadja Braun Binder, 'Artificial Intelligence and Taxation: Risk Management in Fully Automated Taxation Procedures' in Thomas Wischmeyer and Timo Rademacher (ed), *Regulating Artificial Intelligence* (Springer, 2020) 295, 298: the author discusses two examples from the interpretation of the GDPR; Selbst (n 1) 1134.

THE LAW SOCIETY OF NEW SOUTH WALES
young LAWYERS

Proposal 7

30. **Summary**: The CET Committee submits that it supports the introduction of a right of explainability and that:

    i. some proposed changes are necessary to existing administrative review to ensure the efficacy of such a right;

    ii. The importance and correlated expectation of human oversight, and

    iii. Administrative law protections should be favoured over the protection of trade secrecy.

31. The CET Committee agrees that the Australian Government should introduce legal requirements regarding the "explainability" of AI-informed decision-making[36] to set the bar regarding the reasonableness of explanations. This is in keeping with current requirements in administrative law in which reasons for decisions are provided and would also help promote Human Rights.

32. Matters relevant to the content and quality of explanations may include the purpose of the explanation (ie to allow for judicial or administrative accountability),[37] and the level of detail and technicality in an explanation.[38]

33. The Committee agrees that the legislation enacted should make it clear that, if an individual would have had an administrative right to an explanation of the decision if made without AI support, then the individual should be able to request:

    - A non-technical explanation of the AI-informed decision, which would be "interpretable"[39] by a lay person and with the aim of providing individuals and courts with the "knowledge…to initiate or conduct a judicial or administrative review of a decision:"[40] and

    - A technical explanation of the AI-informed decision that could be assessed and validated by a person with relevant technical expertise.[41]

34. As a point of comparison, s13(1) of the *Administrative Decisions (Judicial Review) Act 1977* (Cth) (**ADJR Act**) presently provides that where a person aggrieved by a relevant decision can apply to a court to have it reviewed, that person is entitled to request the person who made the decision to provide a statement of reasons, being '*…a statement in writing setting out the findings on material questions of fact, referring to the evidence or other material on which those findings were based and giving the reasons for the decision*'.

---

[36] Wischmeyer (n 29) 87 [26].
[37] Ibid 78 [6].
[38] Ibid 77[4].
[39] Selbst (n 1) 1110; Wischemeyer (n 29) 87 [25].
[40] Wischemeyer (n 29) 78[6].
[41] Selbst (n 1) 1093-1094.

35. Similarly, s49(1) of the *Administrative Decisions Review Act 1997* (NSW) provides that, if an administrator makes an administratively-reviewable decision, an interested person may make a written request to the administrator for the reasons for the decision. Section 49(3) provides that the statement of reasons provided in response to such a request is to set out the findings on **material questions of fact** (referring to the evidence or other material on which those findings were based); the **administrator's understanding of the applicable law**; and the **reasoning processes that led the administrator to the conclusions** made by the administrator.

36. On this basis, the Committee submits that, in relation to explanations for administratively-reviewable AI-informed decisions, a statement of reasons will need to extend to the methodology used in reaching an AI-informed decision,[42] the quality of the data or information relied upon in the decision,[43] and the source of that data or information.[44]

37. Regarding the statement of reasons for administrative decisions, the Committee suggests that:

- The definition of the term 'writing' may need to broaden so that it extends to 'codification', since it may not be possible to set out the results of an AI-informed decision in writing;

- An administrator's understanding of the role and methodology of an applicable AI-technology or system used should be included in a statement of reasons;[45]

- The model information (the relationship between normative aims, system values, and outputs) of the AI-informed decision-making system should be included in a statement of reasons;[46]

- That the technical specifics of model information may be excepted where this would reveal trade secrets or fail to explain why the decision was made; and,

- The source and quality of the data or information relied on should be explained.[47]

Grounds of review

38. The Committee further submits that in order to be applicable to AI-informed decisions, grounds of review of administrative decisions in civil matters will need to be revised, updated or broadened, so that existing grounds of review encompass issues such as algorithm bias.

Access to information

[42] Selbst (n 1) 1130.
[43] Tony Boobier, *Advanced Analytics and AI: Impact, Implementation, and the Future of Work* (John Wiley & Sons, 2018), 167.
[44] Tischbirek (n 5) 105-106.
[45] Selbst (n 1) 1132.
[46] Wischmeyer (n 29) 95 [43].
[47] Tischbirek (n 5) 105-106.

39. The Committee also notes that AI-informed decision-making systems and the decisions they produce may raise novel issues regarding public access to government information.

40. For example, in NSW, s9(1) of the *Government Information (Public Access) Act 2009* presently provides that a person who makes an application for access to government information '*has a legally enforceable right to be provided with access to the information… unless there is an overriding public interest against disclosure of the information'*.

41. The Committee submits that Australian law should be reformed in order to anticipate applications from individual members of the public to information or data used by AI-informed decision-making systems, with a strong public interest in favour of disclosure such as considered in this paper.

Human oversight and intervention in AI-informed decision-making

42. The Committee submits that it is essential that human oversight and intervention in AI-informed decision-making be maintained as far as possible in order to bolster the legality of AI-informed decision-making. This would also ensure that accountability for AI-informed decision-making can be properly achieved.[48]

Trade Secrets

43. The Committee acknowledges that opacity surrounding certain aspects of an algorithm or AI program may be necessary for legitimate protections of commerciality or trade secrets. In the context of government decision-making there are three key distinctions to consider against the protection of trade secrets in the event of conflict: actual "competitive advantage," [49] the kind of disclosure necessary to achieve transparency, [50] and the kind of decision being made.

44. With respect to 'competitive advantage,' whilst this kind of transparency may create commercial costs to entering this area of development, this cost would apply to any group which sought to enter this field, diminishing the competitive disadvantage for any single developer.

45. Secondly, it is important to distinguish the difference between a "global" explanation of a model, and the explanation of a single decision ("local").[51] A 'holistic' explanation of a model will be both more challenging, and more likely to engage with material which might constitute a trade secret. Whilst there may be occasions where holistic examination of a model is in the interests of the rights of those affected, the established rights of administrative law are focused on single decisions.

46. Finally, and most importantly, the legal rights of persons impacted by administrative decisions should be understood to outweigh the commercial considerations of companies developing such tools. Administrative

---

[48] Selbst (n 1) 1132, 1139; Hermstrüwer (n 26) 204-205 [16]-[19]: example in the EU context.
[49] Selbst (n 1) 1093.
[50] Ibid 1130.
[51] Lisa Käde and Stephanie von Maltzan, 'Towards a Demystificaiton of the Black Box – Explainable AI and Legal Ramifications' (2019) 23(3) *Journal of Internet Law* 3, 5.

decisions have specific rights of review and transparency that should not be diminished by the inclusion of AI or algorithmic support.

Question B

47. **Summary**: The CET Committee supports the introduction of rebuttable presumption due to the significance and normative consistency within administrative law.

48. The Committee submits that, in respect to administrative law, a human decision-maker or government body must be responsible for an AI-informed decision, and that if a reasonable explanation for that decision cannot be provided, a rebuttable presumption that the decision was not made lawfully should apply.

49. Beyond the general standard of 'reasonableness', the Committee submits that administrative to actively ensure developers build-in transparency by design. Standards for transparency need to be developed with the input of the administrative bodies proposing to rely on the system,[52] to ensure that model and normative aims are aligned, [53] and that the 'transparency' outcomes are conducive to judicial and administrative review.[54]

3. Models of Regulation

Question D

50. **Summary**: The CET Committee submits that:

    i. accountability and explainability are key principles to successful AI, and require at a minimum, the capacity for human intervention;

    ii. a proposed list of 'Requirements' should be used to evaluate the accountability and explainability of a decision-making system;

    iii. any introduction of legislation in this area would need to be technology and sector neutral; and

    iv. minimum requirements are necessary to ensure ongoing commitment to transparency beyond the developer phase.

51. The Committee submits that Australian law should require the intervention by human decision-makers. This requirement is a practical manifestation of *accountability* and *explainability* when it comes to AI-

---

[52] Australian Human Rights Commission, *Human Rights and Technology: Discussion Paper* (December 2019),190.
[53] Selbst (n 1) 1130.
[54] Wischemeyer (n 29) 78[6].

informed decision-making, without which, there is potential for harm to the Human Rights of the subjects of decisions.[55] Such interventions should account for the following (henceforth '***the Requirements'***):

- Consider suitable parameters for decisions in light of the potential effect they can have on the Human Rights of their subjects;
- Provide for Human Rights and ethics risk assessments;
- Provide mechanisms for monitoring compliance of the decision-making with the set parameters, and
- Provide mechanisms for ensuring reasonably transparent accountability of appropriate actors for any harm to Human Rights of the subjects of the decision-making, and compensation for the harm if restoration to their status prior to the decision-making is not feasible or is unreasonable.

52. In order to 'preserve a human-centric society' by 'protecting ethical values as defined in fundamental rights and basic constitutional principles',[56] human intervention is vital. This is particularly given the high potential harms of normative mismatch.[57] Such mismatch being a distinct possibility as AI systems may lack concepts of causality,[58] a key feature of most ethical systems. The Committee views humans as necessary to ensure AI systems, which are becoming both ubiquitous and essential,[59] are 'trustworthy'.[60]

53. The Committee submits that the inherent requirement for human intervention must not distinguish between AI-informed decision-making occurring in any sector of the economy. Just as existing laws covering subjects like discrimination and consumer protection are technology-neutral,[61] the inherent requirement must be sector-neutral as a minimum. The precise nature of this requirement, and any unique elements of it, would have to be calibrated by the nature and magnitude of the infringement upon the Human Rights of the subjects of the decision in question.

54. The requirement for intervention of human decision-makers (across sectors), at minimum should include the following (formatted here in an example legislative framework):

> Human beings who are actively involved in the development and/or deployment of an AI-informed system with the express intention that that AI-informed system is to be used to make a decision with a substantive effect on the Human Rights of the subject(s) of that decision, or who are otherwise

---

[55] See eg Australian Human Rights Commission, *Human Rights and Technology: Discussion Paper* (December 2019) 75, 81-2; Select Committee on Artificial Intelligence, *AI in the UK: Ready, Willing and Able?* (House of Lords Paper No 100, Session 2017-19) 36.
[56] World Economic Forum, *AI Governance: A Holistic Approach to Implement Ethics into AI* (White Paper, January 2019) 15.
[57] Hermstrüwer (n 26) 216 [56].
[58] Will Knight, 'If AI's So Smart, Why Can't it Grasp Cause and Effect?', *Wired* (online, 9 March 2020) <https://www.wired.com/story/ai-smart-cant-grasp-cause-effect/>
[59] *Pintarch v Deputy Commissioner of Taxation* [2018] FCAFC 79 [47].
[60] European Commission Independent High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI* (Guidelines, 8 April 2019) 7 <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.
[61] Australian Human Rights Commission (n 54) 87.

involved in the operation of the AI-informed system as decision-makers must establish, execute and maintain measures of a similar character to the following:

i. Setting and maintaining the Requirements;

ii. Compliance with all relevant laws;

iii. Compliance with minimum reasonable standards for transparency of the actions taken in course of setting the Requirements;

iv. Providing an effective means for the subjects of past decisions to challenge them on the grounds of violation of their Human Rights under relevant international covenants to which Australia is a signatory, and:

- if the challenge is not made out, provide a reasonable explanation of the reasons for the original decision and why the challenge is not made out; or

- if the challenge is made out, amend or cancel the decision; and

v. Proactively intervening before a decision is made if they, or their human associates, reasonably suspect, regardless of the source of the evidence informing that suspicion, that a violation of the Human Rights of a subject(s) of the decision may occur.

55. This process should be implemented in addition to best practice guidance from respected international or supranational organisations, such as (bodies within the) the European Commission (especially the recommended mechanisms for human oversight),[62] the World Economic Forum[63] and the OECD.[64]

56. The Committee recognises that these recommendations cover more than merely the point-in-time intervention of dedicated 'human decision-makers' in an AI-informed decision. The inclusion of developers in this model of the proposed requirement is significant to ensuring systems which are transparent by design, and to ensuring accessible enforcement mechanism for human and administrative rights across sectors.

57. By placing the legal obligation on the overseeing or commissioning body, and not just the designers, this measure seeks to ensure continued transparency beyond initial impact reports. The Committee, however, also recognises this process of intervention will often be interdisciplinary, being driven by the efforts of the developers, testers and maintainers of the AI-informed system as well as the dedicated decision-makers. This echoes the Recommendation of the OECD Council on AI, which champions multi-stakeholder collaboration.[65]

---

[62] European Commission Independent High-Level Expert Group on Artificial Intelligence (n 60).
[63] See eg World Economic Forum (n 55).
[64] OECD, *Recommendation of the Council on Artificial Intelligence* (OECD Legal Instrument No OECD/LEGAL/0449, 22 May 2019).
[65] Ibid 2.5(b)-(c).

58. Human intervention in AI-informed decision-making is not a panacea for rationality.[66] We would invite the Commission as a well-placed and essential stake-holder to do further research on this issue, given the importance of human agency and user autonomy in relation to AI systems.[67]

Question F

59. **Summary**: The CET Committee submits that:

    i. a regulatory sandbox may be a significant measure to respond to the rapid development of the sector;

    ii. technological neutrality, or a multi-sector specific system would be necessary to ensure universal Human Rights;

    iii. in addition to Human Rights, key areas will include privacy, data governance, and cyber security.

    iv. Regarding criteria of entry, the CET Committee submits that requirements of cyber security expertise and solvency will be necessary to preserve efficacy;

    v. a multi-disciplinary, multi-stakeholder approach is necessary;

    vi. the aims of the sandbox, on balance, should supersede other legal rights such as trade secrets; and

    vii. the sandbox should be evaluated by reference to the Commission's proposed 'Requirements' at minimum;

60. The Committee supports the use of a Human Rights regulatory sandbox as one aspect of the development of appropriate regulation. It agrees with the Commission about the rapid, unpredictable development of AI and the delay inherent to legislative reform to catch up.[68] This form of regulation is a key part of an effective regulatory landscape, especially considering the need to in-build protections in the developmental and design stages.[69]

(a) What should be the scope of operation of the regulatory sandbox, including criteria for eligibility to participate and the types of system that would be covered?

61. The CET Committee considers that the regulatory sandbox must be technology neutral. It must not exclude any particular type of AI-informed system or use case. If such a sandbox would be too broad or unwieldy, however, for relevant government agencies to control, the CET Committee recommends that the Commission research the potential for multiple sandboxes that are either catered to the type of AI involved

[66] Australian Human Rights Commission (n 54) 101-2.
[67] European Commission Independent High-Level Expert Group on Artificial Intelligence (n 58).
[68] Australian Human Rights Commission (n 54) 108.
[69] Australian Human Rights Commission (n 54) 108-109.

in the decision-making system (be it merely simple machine learning algorithms or exceedingly complex artificial neural networks) or the sector to which the system is targeted.

62. The CET Committee views that operation of the sandbox should, again, consist of a multidisciplinary approach including input and supervision from legal representative(s) advising the developer, legal counsel from the relevant government agency or agencies administering the sandbox, and experienced cyber security experts. The sandbox may also require the completion of specific supervised outcomes including ethics/Human Rights risk training (at a standard which is approved by the Commission), a cybersecurity and privacy protection plan in relation to its participation, and a statement of objectives.

63. The Commission may consider these criteria in addition to those governing the sandboxes mentioned in the Discussion Paper,[70] as well as financial services regulatory sandboxes from government agencies such as the United Kingdom's Financial Conduct Authority; especially the criterion of 'genuine innovation'.[71]

64. Given both the nature of the technology and the rights involved, an international or multinational approach may help maximise the learnings and benefits from a sandbox setting, however, may also diminish the administrative capacity of a participating government entity. The Commission could model such a sandbox after the financial services regulatory equivalent in the Global Financial Innovation Network.[72]

(b) What areas of regulation should it cover?

65. The CET Committee submits that the sandbox should cover Human Rights, privacy, data governance and cybersecurity as a minimum, given the strong relevance of these areas to AI (see our earlier submission on AI).[73] Any additional areas of regulation that the sandbox must cover would be those that are directly and primarily engaged by the use case(s) represented by the system or the needs of any particular sector.

(c) What controls or criteria should be in place prior to a product being admitted to the regulatory sandbox?

66. With respect to criteria of admission, the CET Committee considers the following issues to be significant to avoiding undue breaches of privacy and Human Rights:

    - The applicant entity must have passed a cybersecurity and data governance audit for compliance with relevant standards under the *Privacy Act 1988* (Cth), ISO/IEC 27001 and the Australian Government Information Security Manual, and

---

[70] See eg Australian Human Rights Commission (n 54) 118.
[71] See eg 'Applying to the Regulatory Sandbox', *Financial Conduct Authority* (Web Page, 17 January 2020) <https://www.fca.org.uk/firms/innovation/regulatory-sandbox-prepare-application>.
[72] 'Global Financial Innovation Network (GFIN)', *Financial Conduct Authority* (Web Page, 27 February 2020) <https://www.fca.org.uk/firms/innovation/global-financial-innovation-network >.
[73] Ellen Brown et al, Submission to Submission to Department of Industry, Innovation and Science, Commonwealth of Australia, *Artificial Intelligence: Australia's Ethics Framework* (11 June 2019).

- The applicant entity or their sponsor must have sufficient funds available to remain a going concern during their participation.

(e) What body or bodies should run the regulatory sandbox?

67. The CET Committee notes (in addition to the Commission) the following bodies as possible governing entities:

    - *Office of the Australian Information Commissioner;*
    - *Australian Cyber Security Centre;*
    - *Australian Competition and Consumer Commission;*

(f) How could the regulatory sandbox draw on the expertise of relevant regulatory and oversight bodies, civil society and industry?

68. Multidisciplinary support from regulatory bodies is significant in the framing of initial policy objectives, eligibility criteria, controls and processes for the sandboxes, as well as the criteria for evaluation of the sandboxes.

69. The nature of this consultation and liaison, and the particular stakeholders prioritised in the process, should vary with the nature of the policy objectives, community sector(s) or area(s) of law primarily engaged by the specific cohort of entities participating in the sandbox.

(g) How should it balance competing imperatives eg, transparency and protection of trade secrets?

70. The Committee provides that Human Rights, and existing legal rights of the individual are of paramount importance, and as previously discussed, hinge upon transparency and accountability.

71. While the protection of intellectual property is an important and necessary measure, it must not overshadow what is the primary objective of the regulatory sandbox — the protection and enhancement of the Human Rights of the subjects of decisions made by the AI-informed system. The potential freedom offered by the regulatory sandbox to test and refine the Human Rights-compliance of innovative technologies in a controlled environment must not be unduly limited by the imperative of protecting commercially sensitive information. A properly functioning sandbox can be conducive to the generation of better, more valuable, technology which is more capable of fulfilling the sandbox's policy objectives, versus a more limited one: the benefits of the sandbox is tied with the quality and depth of the testing.

(h) How should the regulatory sandbox be evaluated?

72. Standards of evaluation should consider the practices of the sandboxes mentioned in the Discussion Paper,[74] as well as financial services regulatory sandboxes. The European Commission Independent High-Level Expert Group on Artificial Intelligence pilot list for 'Trustworthy AI Assessment' would be a useful source of evaluation criteria as well. [75] This would be in addition to the previous **Requirements** as discussed above. Additional considerations may include:

    1. *Were the policy objectives of the sandbox met by the results of the participation of the applicant entity?*

    2. *Were the specific objectives set by the applicant entity aligned with the policy objectives of the sandbox?*

    3. *Were any cybersecurity risks in relation to the AI-informed system successfully mitigated, or incidents averted or reasonably managed?*

    4. *Were any privacy risks in relation to the AI-informed system successfully mitigated, or incidents averted or reasonably managed?*

    5. *What was the financial impact of the AI-informed system? Did it operate within the applicant entity's budget, developed in consultation with the experts from the sandbox that advised the entity?*

Proposal 10

73. **Summary**: The CET Committee is in favour of Proposal 10 and supports such a right as the logical and practical extension of accountability.

74. This is a logical extension of the requirement of human intervention and broader human involvement in the development and deployment of AI-informed systems for decision-making. The sheer potential for harm arising from the use of AI[76] necessitates accountability of the most serious kind. Liability provides a potentially significant deterrent against both negligent and deliberate harms to Human Rights. Actors would be required, by implication, to put in reasonable measures to prevent, and not merely to respond to such harms at all stages, from development to decision, and thereafter.

75. This is of particular significance to decisions which impact on the administrative rights, but as discussed, the administrability of an issue should not limit the regulation. All actors relying on AI have the potential to impact the rights of communities and individual, and should be responsible for those harms, just as they would without the use of AI systems.

---

[74] See eg Australian Human Rights Commission, (n 54) 118.
[75] European Commission Independent High-Level Expert Group on Artificial Intelligence, (n 60).
[76] See eg Australian Human Rights Commission, (n 54) 67; Alana Maurushat, 'BD Use by Law Enforcement and Intelligence in the National Security Space: Perceived Benefits, Risks And Challenges '(2016) 21 *Media and Arts Law Review* 229, 250-1.

THE LAW SOCIETY OF NEW SOUTH WALES
young LAWYERS

Concluding Comments

NSW Young Lawyers and the Committee thank you for the opportunity to make this submission. If you have any queries or require further submissions please contact the undersigned at your convenience.

Contact:

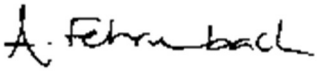
David Edney

President

NSW Young Lawyers

Email: president@younglawyers.com.au

Alternate Contact:



Ashleigh Fehrenbach

Chair

NSW Young Lawyers Communications, Entertainment and Technology Committee

Email: ashleigh.fehrenbach@younglawyers.com.au