Deakin Law Clinic

Submission to the Victorian Law Reform Commission on Artificial Intelligence in Victoria's Courts and Tribunals.

Prepared By

Deakin Law Clinic T3 2024 Policy Advocacy Students under the supervision of Rebecca Tisdale, Clinical Solicitor

Date of Submission

22 January 2025

deakinlawclinic.org.au

Deakin University CRICOS Provider Code: 00113B

Email: law-clinic@deakin.edu.au

ABN: 56 721 584 203

Waterfront Campus
Ground Floor, Building AD
Cunningham St
Geelong VIC 3220

Phone: 03 5227 8563

Deakin Downtown
Tower 2, Level 12
727 Collins St
Melbourne VIC 3008

Phone: 03 9244 3000

Deakin Law Clinic

Deakin Law Clinic serves as a practical Work Integrated Learning element of the LLB program. Under the supervision of qualified legal practitioners, Deakin Law Clinic offers opportunities for students to build the legal skills that underpin their academic training. This submission has been prepared by students undertaking the Policy Advocacy unit within the Deakin Law Clinic, which aims to provide students the opportunity to critically evaluate the adequacy of existing law and policy.

Deakin Law Clinic Policy Advocacy practice - Trimester 3 2024 students contributing to this submission:

Aastha Agarwal

Georgia Stadelmann

Ishrat Rayhan

Katie Robl

Marieta Ryan

Matthew George

Natalia Dembek Sanchez

Tim Brewer

Ysabeau Ellery-Slade

The students worked under the supervisor of Rebecca Tisdale, Clinical Solicitor.

Table of Contents

Introduction

Deakin Law Clinic welcomes the opportunity to respond to the Victorian Law Reform Commission's (**VLRC**) consultation paper on Artificial Intelligence (**AI**) in Victorian Courts and Tribunals. Our written submission is intended to assist VLRC with its review and recommendations.

As we are making this submission in an academic capacity, as future legal practitioners, we are in a unique position to provide the VLRC with a student perspective. With substantive academic training in research and an interest in AI, we are well-placed to analyse the intersection of technology and law.

In accordance with the Terms of Reference, this submission will examine the role AI has to play in Courts and Tribunals broadly and the potential regulation of AI in these roles, including lessons drawn from other jurisdictions, both domestic and international. We will address key issues such as benefits and risks and the need for transparency and accountability to ensure the adoption of AI aligns with the principles of judicial independence and procedural fairness.

To avoid confusion during our submission, when we mention AI, we refer to the definition provided by OECD which describes AI as 'machine-based system that, for explicit or implicit objectives, infers from the input it receives how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.'[1]

[1] OECD, Recommendation of the Council on Artificial Intelligence (Report No OECD/LEGAL/0449, 2024) 7.

Executive Summary

The integration of Artificial Intelligence (**AI**) into the Victorian legal system presents substantial opportunities to improve efficiency, reduce costs, and enhance access to justice. AI technologies are already being utilised in various aspects of court operations, including administrative tasks such as case management and document review, as well as providing basic legal advice to the public. However, the introduction of AI also raises significant concerns about fairness, transparency, security, and the potential for bias, which can undermine public trust in the legal system. This submission provides an in-depth analysis of AI's role in the legal system, the associated risks, and recommends a balanced regulatory approach to ensure AI's responsible integration.

I Opportunities and Risks

AI offers significant benefits for Victorian courts, tribunals, and the broader legal profession. By automating routine administrative tasks such as case management, e-filing, and document processing, AI can enhance efficiency, reduce delays, and address backlogs. Legal professionals stand to benefit from AI's ability to streamline legal research, document review, and other tasks, leading to faster and more accurate legal proceedings. AI-powered tools, such as chatbots and virtual legal advisors, can also enhance access to justice, particularly for self-represented litigants or those with limited financial resources, by providing quick, low-cost legal information. These advancements in AI technology have the potential to democratise legal services, making it easier for individuals to navigate the legal system, access legal advice, and understand their rights, thus bridging the justice gap for underserved populations.

However, the integration of AI into the legal system presents several risks that must be carefully managed. A major concern is the potential for embedded biases within AI systems, which may perpetuate or exacerbate existing biases in data, resulting in unfair outcomes. Predictive algorithms, such as those used in bail applications or sentencing, could unintentionally discriminate against certain groups, undermining fairness in judicial processes. Additionally, the lack of transparency and accountability in AI systems, particularly those operating as "black-boxes," raises concerns about public confidence and trust in the legal system. Security and confidentiality are also critical issues, as the use of AI involves processing sensitive personal data, raising the risk of breaches or manipulation. Perhaps most concerning is the potential for AI to replace human judgement in judicial decision-making. While AI can assist in decision support, it cannot replicate the nuanced, context-driven decisions made by judges, which require

empathy, moral reasoning, and discretion. There is a risk that over-reliance on AI could compromise judicial independence and fairness in the legal system.

II International Approaches to AI Regulation

Looking to international examples, there are valuable lessons that can inform the regulation of AI in Victoria's legal system. Brazil has made significant strides in AI implementation in its courts, emphasising bias reduction, ethical guidelines, and centralised AI platforms. Canada, similarly, emphasises judicial independence, transparency, and security, highlighting the importance of educational programs and regular impact assessments. China's approach, rooted in information control, has led to a multi-faceted regulatory framework that addresses specific issues like algorithmic recommendations, deepfakes, and generative AI, with an emphasis on political and social stability. The United States (**US**) adopts a sector-specific, risk-based approach, emphasising innovation alongside accountability, transparency, and human oversight, which could guide Victoria's AI governance in judicial settings. The European Union (**EU**) has implemented the AI Act, a comprehensive regulation with a focus on high-risk areas such as justice, yet has faced delays in AI implementation within courts, suggesting that regulatory frameworks alone may not suffice without broader stakeholder engagement. The UK follows a flexible, principles-based framework with a focus on industry collaboration, regulatory sandboxes, and assurance techniques to ensure AI safety and fairness. Lessons from these regions highlight the importance of adaptable, transparent frameworks, proactive stakeholder engagement, and robust governance, including clear educational and security protocols, to successfully integrate AI in judicial systems while maintaining public trust and accountability.

III Recommendations for a Balanced Regulatory Approach

This submission recommends a balanced regulatory approach, while principles alone are insufficient for meaningful reform, targeted guidelines and tools are necessary to ensure AI technologies are implemented responsibly. Drawing on international examples, particularly China's approach to AI regulation, the submission suggests that Victoria should adopt a similar incremental strategy, beginning with specific reforms/policy that address key issues before developing broader legislative frameworks. Key recommendations include:

- retaining judicial autonomy over AI systems;
- ensuring transparency and accountability;
- mitigating bias through disclosures; and

- providing ongoing training for legal professionals and judicial officers.

This submission also advocates for a tiered approach to transparency, differentiating between high-risk AI applications used by judicial officers and lower-risk tools employed by administrative staff, to maintain public trust, judicial independence, and fairness in legal processes. In addressing disclosure, the submission recommends a layered disclosure framework, public education, and the establishment of a centralised regulatory body to ensure safe and ethical AI implementation in the legal sector. As although disclosure enhances oversight and confidence, it also poses risks such as over-disclosure, operational challenges, and misinterpretation by non-technical stakeholders. Additionally, it highlights the need for a unified set of guidelines for AI use in both criminal and civil law, focusing on fairness, accuracy, and efficiency, while allowing for future adjustments as AI technology evolves.

The assessment of AI frameworks for Victorian courts and tribunals reveals that while the NSW AI Assurance Framework provides a useful model for ensuring safe, secure, and accountable use of AI, its broad application to government agencies and reliance on self-assessment limit its effectiveness for the legal system. The framework's lack of clear risk category definitions and its requirement for human oversight could lead to a false sense of security in AI tools. As a result, it may not be fully suitable for Victorian courts and tribunals, though the creation of an AI Review Committee in NSW is a valuable aspect that Victoria should consider adopting. A better model could be the CEPEJ Risk Assessment framework used in the EU, as it is more tailored to judicial systems and offers a simpler, more focused approach. It is recommended that Victoria establish an AI Assessment Review Committee and consider adopting a sector-specific risk assessment framework to ensure accountability, minimise bias, and safeguard judicial independence.

By implementing these recommendations, Victoria can responsibly harness the potential of AI while mitigating its risks, ensuring the Courts and Tribunals remain fair, transparent, and accountable. For a summarised version of the recommendations provided by the submission please see pages 74-76.

Chapter 3: Benefits and Risks of AI

Question 3: What are the most significant benefits and risks for the use of AI by Victorian courts and tribunals, legal professionals and prosecutorial bodies, and the public (including court users, self-represented litigants and witnesses)?

I Introduction

The legal profession is plagued with administrative issues which in turn has reduced the quality of judiciary services.[2] These issues cause backlogs, lengthy wait times and costly proceedings.[3] However, the introduction of AI has shown to be more than useful in legal contexts.[4] Not only has AI proven its ability to take over tasks originally performed by humans, but it has also demonstrated an ability to complete tasks humans cannot do.[5] Arguments have been put forward by scholars and legal professionals to introduce AI into the legal profession to assist courts, legal practitioners, and the broader public to resolve disputes in a faster and cheaper manner.[6] This part will engage in a discussion on the benefits of AI, focusing on its ability to promote justice, enhance the efficiency and quality of judicial proceedings.[7] However, these benefits will be balanced against the risks that AI presents. Namely, the risk of security and confidentiality breaches, and risks surrounding fears of an over-reliance being placed on AI systems by legal professionals and the broader public. Ultimately, AI can produce numerous benefits for courts, legal practitioners, and the broader public. However, AI systems must not operate without human oversight and that AI should replace the work of legal professionals.[8]

---

[2] Kalliopi Terzidou, 'The Use of Artificial Intelligence in the Judiciary and its Compliance with the Right to a Fair Trial' (2022) 31 *Journal of Judicial Administration* 154, 154-168 ('The use of Artificial Intelligence in the Judiciary').
[3] Ibid
[4] Ray Worthy Campbell, 'Artificial intelligence in the courtroom: The delivery of justice in the age of machine learning' (2023) 15 *Revista Forumul Judecatorilor* 8, 1-27 ('Artificial Intelligence in the Courtroom')
[5] Ibid.
[6] Ibid; The use of Artificial Intelligence in the Judiciary (n 2).
[7] The use of Artificial Intelligence in the Judiciary (n 2) 155.
[8] Artificial Intelligence in the Courtroom (n 4) 10.

II Benefits

A *Efficiency*

The implementation of AI into Victoria's courts and tribunals could assist in increasing efficiency by automating daily administrative tasks and workflows.[9] These tasks involve drafting documents, managing correspondence and allocating and planning for cases.[10] The use of AI in these settings could also extend to the distribution of legal information through chatbots, organising and triaging registries for legal aid or *pro bono* support, or the building of intranets and extranets for communication.[11] The transferring of these tasks to AI systems would reduce the workload of court and tribunal administrative staff. In turn, this could look to reduce backlogs, lengthy wait times and costly proceedings,[12] and inevitably increasing the general public's right to be heard.[13]

The introduction of AI has also proven to assist in increasing legal professionals' efficiency. AI systems have the power to streamline legal research and document review tasks.[14] AI systems have shown an ability to analyse statues and legal decisions in a manner which allows lawyers to move more quickly and accurately to find relevant law.[15] Further, AI systems and software have proven strong in their ability to assist with document review.[16] By assisting in discovery, AI not only allows legal professionals to increase their efficiency, but it also reduces the costs to the client.[17]

AI also has the power to help the public to be more efficient when compiling their case. The use of chatbots can allow individuals to access more information at a faster rate than administrative staff.[18] Further, it can help individuals streamline court processes by providing them with direct and accurate information on what is required to start proceedings in a court or tribunal.[19]

---

[9] The use of Artificial Intelligence in the Judiciary (n 2) 155.
[10] Ibid.
[11] *Study on the Use of Innovative Technologies in the Justice Field* [2020] EU 142/111.
[12] The use of Artificial Intelligence in the Judiciary (n 2) 155.
[13] Nikolaos Aletras et al, 'Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective' (2016) 2(93) *PeerJ Computer Science* 1, 7 ('Predicting Judicial Decisions').
[14] Manjari Singh, 'Review on Role of Artificial Intelligence in The Life of Legal Profession' (2024) 6(3) *International Journal of Legal Science and Innovation* 1087, 1087-1096 ('Role of Artificial Intelligence').
[15] Artificial Intelligence in the Courtroom (n 4) 8.
[16] Ibid.
[17] Sanna Luoma, 'AI Improving the Delivery of Justice and How Courts Operate' (2018) *How Will AI Shape the Future of Law* 72, 63-100 ('AI Improving the Delivery of Justice')
[18] Australian Human Rights Commission, 'Mandatory Guardrails for AI in High-risk Settings' (Submission to Department of Industry, Science and Resources, Australian Human Rights Commission, 04 October 2024) 1-10 ('Mandatory Guardrails for AI in High-risk Settings').
[19] Ibid.

However, discretion and care must be exercised when looking to utilise AI systems to automate certain tasks.[20] Attempts to implement AI to automate tasks to improve efficiency must not be engaged in a strictly business sense.[21] While it is acknowledged that there may be strong cost saving reasons to implement AI to increase efficiency, these reasons must be balanced with the potential sacrifice of judicial values.[22] Further, while there have been discussions of implementing AI in the context of Automated Decision Making (**ADM**) to increase efficiency,[23] to reap any gains these systems must function with a high degree of accuracy.[24] However, there are currently few avenues in the way of checking the validity of automated workflows to identify errors that may have been missed.[25] To use an AI system in ADM which is anything but highly accurate can cause harm and damage to the parties and legal professionals involved.[26] A court in the UK used a pre-approved divorce software that generated an error which was unable to be recognised due to the absence of software able to identify such inaccuracies in the workflow.[27] This error required 2,235 cases to be reopened and inevitably resulted in a massive backlog of cases.[28]

It is acknowledged that the implementation of AI in legal processes can lead to an increase in efficiency for courts and tribunals, legal practitioners and the public. However, these benefits must also be weighed against the potential risks of implementing such automated systems present.

B *Access to Justice*

The implementation of AI into courts and tribunals has strong prospects in increasing access to justice. The utilisation of chatbots by courts and tribunals could allow the distribution of more information on proceedings at a faster and more efficient rate.[29] Courts and tribunals could also improve access to justice where these chatbots can answer questions about how to lodge an application or what documents are required for different proceedings.[30] This would be

---

[20] *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonized on Artificial Intelligence (Artificial Intelligence Act)* [2021] OJ C 206 ('Proposal for a Regulation of the European Parliament').
[21] Ibid.
[22] Falicity Bell, Lyria Bennett Moses, Michael Legg, Jacob Silove and Monika Zalnieriute, 'AI Decision-Making and the Courts: A Guide for Judges, Tribunal Members and Court Administrators' (2023) *Australasian Institute of Judicial Administration* 1, 57 ('AI Decision-Making and the Courts').
[23] Mandatory Guardrails for AI in High-risk Settings (n 18) 23 [97].
[24] AI Decision-Making and the Courts (n 22) 58.
[25] The use of Artificial Intelligence in the Judiciary (n 2) 58.
[26] Proposal for a Regulation of the European Parliament (n 20); The use of Artificial Intelligence in the Judiciary (n 2) 58.
[27] Amnon Reichman, Yair Sagy and Shlomi Balaban, 'From a Panacea to a Panopticon: The Use and Misuse of Technology in the Regulation of Judges' (2020) 71(3) *Hasting Law Journal* 589, 597.
[28] Ibid
[29] The use of Artificial Intelligence in the Judiciary (n 2) 159.
[30] AI Decision-Making and the Courts (n 22) 57.

accomplished by training the chatbots with data pulled from court files and previous case outcomes.[31] By allowing the public, but more specifically self-represented litigants, access to specific information regarding their matter in a fast and efficient manner, this in turn would assist more individuals to access legal recourse.[32] Further, providing a virtual space for individuals to access information and tools to assist in their dispute can also improve access to justice.[33] These online spaces allow individuals who do not have the resources or income to engage a lawyer the opportunity to obtain legal recourse.[34]

As previously discussed, AI has shown to have the ability to increase legal professional efficiency and productivity while decreasing costs to cliental.[35] By implementing AI systems to assist with research and discovery, legal professionals could offer their services at a decrease rate to new and existing clients. Adopting this process would increase access to justice as it would allow individuals who once could not afford legal representation to engage a lawyer for their dispute.[36]

However, there are concerns regarding the 'rebound effect' – that is, if the court process becomes faster and cheaper, people could abuse the process and submit claims with little to no merit.[37] However, it is unlikely that this would happen due to processes in place such as genuine steps certificates which inhibit individuals to submit claims where they have not made genuine steps to resolve it. Further, it is all Australian's human right to be heard in their legal matter.[38] Therefore it is the courts and tribunals duty to ensure that all individuals have the opportunity to exercise that right.[39]

We also acknowledge that caution must be taken when considering using AI systems and chatbots to assist in legal proceedings. Defective or bias training of the AI system could lead to false or inaccurate information being given to those who engage with the chatbots.[40] Further, AI systems and chatbots could also impede access to justice if the systems were to provide longer wait times or inaccurately predict the outcome of a case.[41] These errors could result in individuals

---

[31] The use of Artificial Intelligence in the Judiciary (n 2) 159.
[32] Amy J Schmitz and John Zeleznikow, 'Intelligent Legal Tech to Empower Self-Represented Litigants (2021) 23 *Columbia Science & Technology Law Review* 142, 142.
[33] AI Improving the Delivery of Justice (n 17) 73.
[34] Ibid.
[35] Role of Artificial Intelligence (n 14) 1087; AI Decision-Making and the Courts (n 22) 57; Artificial Intelligence in the Courtroom (n 4) 9.
[36] AI Decision-Making and the Courts (n 22) 57.
[37] Josh Becker, 'Data Analytics & E-Discovery' (Speech, Litigation Cravath Panel Discussion at 'The Future of Law – the Case for Analytics', 29 March 2018).
[38] Predicting Judicial Decisions (n 13) 7.
[39] Ibid
[40] The use of Artificial Intelligence in the Judiciary (n 2) 159.
[41] AI Decision-Making and the Courts (n 22) 56; The use of Artificial Intelligence in the Judiciary (n 2) 159.

being disincentivised from pursuing their legal cases.[42] This was the case in the European Court of Human Rights where a machine learning system built by Altras et al to predict decisions prioritised the matters most likely to proceed.[43]

III Risks

A *Security and Confidentiality Risks*

It is essential that legal professionals, both judges and practitioners, maintain their confidentiality obligations and ensure there are no security risks to personal information obtained in the course of their work.[44] As technology and AI continues to reshape the legal profession, it is imperative that judicial officers and practitioners balance the use of AI algorithms with their ethical obligations and professional responsibilities.[45] The use of AI by legal professionals raises privacy concerns, especially where the client is unaware that their personal data is being fed into an AI algorithm.[46] This use of personal client information impedes on confidentially obligations legal professionals are under.[47] Further, as AI relies on large sets of data to pull its information from, AI algorithms are currently unsanctioned on what information may be used during this processes.[48] It is imperative that judges, tribunal members and legal practitioners are transparent about how they use personal data and inform individuals about how their data is collected and processed.[49]

B *AI's Inability to Exercise Discretion*

Arguments put forward suggest that a 'robojudge' (AI acting as a judicial decision maker) would be beneficial as it would create outcomes and decisions that are more impartial and efficient, while not being slowed down by human needs such as fatigue or hunger.[50] However, discussions of AI in this setting and the judicial context as a whole appear to oversimplify the role of judges.[51] The assertion that AI could prevent bias in decisions does not consider that the AI's training could

---

[42] The use of Artificial Intelligence in the Judiciary (n 2) 156.
[43] Predicting Judicial Decisions (n 13) 7; AI Decision-Making and the Courts (n 22) 56.
[44] L Brown, 'Navigating ethical concerns in the era of digital lawyering' (2018) 19(1) *Journal of Legal Ethics* 27, 30; Sadikov Ruslan, 'Challenges and Opportunities for Legal Practice and the Legal Profession in the Cyber Age' (2023) 1(4) *International Journal of Law and Policy* 1, 5 ('Challenges and Opportunities for Legal Practice').
[45] R Adams, 'Ethical obligations and the digital age' (2020) 45(2) *Legal Tech Journal* 132, 140; Challenges and Opportunities for Legal Practice (n 44) 5.
[46] Daniel Necz, 'Rules over words: Regulation of chatbots in the legal market and ethical considerations' (2024) *Hungarian 64*(3) *Journal of Legal Studies* 472, 478 ('Rules over words').
[47] Ibid.
[48] Ibid.
[49] Ibid.
[50] AI Improving the Delivery of Justice (n 17) 64.
[51] Artificial Intelligence in the Courtroom (n 4) 12.

be tainted by its programmers own biases.[52] Further, human experience and discretion is a core judicial value in the decision making process.[53] When arriving at a conclusion on a matter, judges engage their discretion and problem solving skills to evaluate the full range of factors involved in the case.[54] This cannot, at this point in time, be meaningfully exercised by any (known) AI algorithms.[55] Further, it is not morally desirable to allow a machine system to make judgements regarding peoples freedoms or even their lives.[56] Allowing a machine to make judgements which have such a significant impact on a person's life without engaging or consulting a human could result in a failure to treat the individuals with dignity and lead to the dehumanisation of the individuals of the case.[57]

While engaging AI systems to assist legal professionals with research and discovery, as discussed, this presents several risks.[58] Engaging AI to take over tedious tasks such as research and discovery may be appealing, however AI can be unreliable and return different results to the same question.[59] Therefore, it is essential that legal professionals proceed with caution and engage their discretion when using AI in the course of their legal research.[60]

Similarly to the risks involved with using AI in the course of legal work, caution must also be taken when engaging AI to provide the public with information on legal processes and the law.[61] As previously discussed, AI systems' inability to implement discretion into its processed could result in it returning inaccurate or false information to individuals who are doing their own research.[62] Where this information is unfavourable to the individuals case, it could dissuade them from proceeding with their application.[63]

The "Robodebt" case, which occurred between 2016 to 2019, demonstrated the dangers of allowing an AI algorithm to run without human oversight.[64] Therefore, while it is acknowledged that AI presents several favourable benefits for legal professionals and the broader public, it is

---

[52] The use of Artificial Intelligence in the Judiciary (n 2) 161.
[53] AI Decision-Making and the Courts (n 22) 56.
[54] Artificial Intelligence in the Courtroom (n 4) 12.
[55] Meena Hanna, 'Robo-Judge: Common Law Theory and the Artificial Intelligent Judiciary' (2019) 29 Journal of Judicial Administration 22, 39.
[56] AI Improving the Delivery of Justice (n 17) 65.
[57] AI Decision-Making and the Courts (n 22) 56.
[58] Artificial Intelligence in the Courtroom (n 4) 8.
[59] Ibid.
[60] Proposal for a Regulation of the European Parliament (n 20).
[61] AI Decision-Making and the Courts (n 22) 56; The use of Artificial Intelligence in the Judiciary (n 2) 159.
[62] The use of Artificial Intelligence in the Judiciary (n 2) 159.
[63] Ibid.
[64] Caroline Gans-Combe, 'Automated Justice: Issues, Benefits and Risks in the Use of Artificial Intelligence and Its Algorithms in Access to Justice and Law Enforcement' (2022) *Ethics, Integrity and Policymaking: The Value of the Case Study* 175, 176.

essential that AI is only engaged with human oversight.[65] AI cannot substitute the discretion required to be exercised by judicial officers and legal professional.[66] AI should not replace lawyers or judges, but only serve as a tool which can be utilised in the course of practice.[67]

Recommendations

1. Establish a "humans-in-the-loop" approach in all facets of AI use within courts and tribunals, ensuring that AI serves as a supplementary tool to enhance, not replace, human judgment. Emphasise the importance of human qualities such as empathy, moral reasoning, and contextual understanding in judicial decisions.
2. Additionally, courts must disclose AI usage in judicial processes and provide clear warnings and disclosures about the potential inaccuracies of AI tools or chatbots to users before engagement.

[65] Role of Artificial Intelligence (n 14) 1087; AI Decision-Making and the *Courts* (n 22) 57; Artificial Intelligence in the *Courtroom* (n 4) 9.
[66] Mandatory Guardrails for AI in High-risk Settings (n 18) 23 [97].
[67] Artificial Intelligence in the Courtroom (n 4) 10.

Chapter 4: AI in Courts and Tribunals

Question 5: Current use of AI in Victorian Legal System

I Courts and Tribunals

Victorian courts and tribunals have made use of automated processes for several years, such as e-filing and case management. However, the utilisation of AI in Victorian courts and tribunals has not begun to take shape until recent years. Guides have been created to aid judges, tribunal members and court administrators in their usage of AI tools.[68] Discussion surrounding how AI will interact with judicial values is a key concern in Victoria and Australia as a whole. Highlighting that there are certain problems as well as opportunities that arise with the inclusion of these tools into the legal system. Judicial administrators have been provided guidance in the treatment of AI. However, it is not currently being used in the decision-making process.

II Legal Professionals and the Victorian Public

Legal professionals in Victoria have begun to use AI in their work, including in tasks such as drafting initial letters, summarising documents and answering a set of questions about documents produced. This can help boost the efficiency of legal work but also calls into question the accuracy of the tools meeting requirements. The Victorian public has also had an opportunity to use AI, this is of relevance in self-represented litigation. The Supreme Court of Victoria has provided guidance to litigants regarding the usage of AI in the litigation process of the court. In relation to AI, principles and application of these principles have been established by the Supreme Court in these guidelines.[69]

Principles include the recognition that AI should be treated with caution due to its natural limitations. The disclosure of the use of AI by practitioners is required where the tool has given assistance in the required legal tasks. Additionally, the guidelines stress the need for litigants to take heed of the privacy and confidentiality of the information that is given to an AI tool. Other principles include not misleading the court through the means of AI, and the use of these tools is made under the ethical obligations placed on legal professionals. Application of these principles

[68] AI Decision-Making and the Courts (n 22).
[69] Supreme Court of Victoria, *Guidelines for Litigants: Responsible Use of Artificial Intelligence in Litigation* (June 2022).

encourages the use of AI as a relevant technology for practitioners to utilise, with a focus on the adoption of legal AI tools for more tailored application.

Recommendations

3. Should Victorian courts and tribunals adopt AI tools in decision making processes in the future, strong regulations must be developed. Such regulations must aim to maintain high rates of accuracy in outcomes. In addition, regulations should be designed to deter decisionmakers from placing excessive reliance on AI tools during the decision-making process.

Question 6: Are there uses of AI that should be considered high-risk, including in courts and tribunal administration and pre-hearing processes, civil claims, or criminal matters?

Currently, high-risk use of AI is determined through a principles-based approach as opposed to an exhaustive list.[70] Because of the emerging and everchanging sophistication of AI technologies, it is more appropriate to define 'high risk' in terms of the magnitude of its consequences. This has been considered largely through a human rights framework.[71] In the legal sector, the relevant high-risk use of AI is when an individual's legal rights and entitlements are significantly impacted. The protection of such rights is an overarching principle in minimising harm when AI is deployed. More broadly, the input of sensitive data and the inherent privacy risks associated with AI use in courts and tribunals is particularly relevant in this context as well.

AI tools can be utilized in court or tribunal pre-hearing processes. For example, 'Automated Decision-Support' systems are being designed for use in pre-hearing processes such as bail applications. Specific to NSW, a "Bail Assistant program [is] being developed by the Judicial Commission of NSW which seeks to guide decision-makers through the complexities of the *Bail Act 2013* (NSW)."[72] Data sets from previous bail decisions will be used to train the AI system which can then provide the decision maker with recommendations for bail decisions.[73] The use of such programs pose a direct risk for the personal liberties of bail applicants. A significant risk stems from the presence of unfair prejudice or algorithmic bias built into the systems from the training

[70] Mandatory Guardrails for AI in High-risk Settings (n 18) 9 [24].
[71] Ibid 1 [2].
[72] AI Decision-Making and the Courts (n 22) 56.
[73] Ibid.

data from which the predictions are derived.[74] The outcomes can be discriminatory in nature and have been found to target vulnerable groups such as low socioeconomic or racial groups in other jurisdictions. The intrinsic bias coupled with the opaque nature of AI decision making systems could result in recommendations that may undermine judicial integrity.[75] Should Victorian courts and tribunals allow ADM in the future, guidelines should be developed to prevent decision makers from placing disproportionate weight on AI recommendations during the decision-making process.[76] Care should be taken by the decision makers to ensure that their rationale and will is not influenced by AI tools. This is critical in maintaining judicial impartiality and hence, judicial independence.

In regard to administrative processes, AI tools can be deployed with seemingly low risk. For example, AI analytics has greatly improved the efficiency of case management through aiding triaging and processing the allocation of cases.[77] It can also aid in identifying any irregularities in case management as well.[78] However, there are circumstances where its use may be considered high risk. The protocols and procedures of the courts and tribunals are frequently subject to change. If AI software algorithm is not up to date with the most recent changes, this will inevitably cause disruption to case management.[7980] This delay may have a cascading effect, causing subsequent delays in access to justice or prolonging time in custody for individuals awaiting trial. This risk can be mitigated by employing human auditors to check that AI software is updated concurrently with any changes in regulation within a jurisdiction.

In the criminal law context, AI systems have supported decision makers with recommendations that have been, at least partially, relied upon when handing down sentence in other jurisdictions. One example is the use of AI recidivism risk tools which can estimate an offender's likelihood of reoffending; perhaps the most notorious being COMPAS (Correctional Offender Management Profiling for Alternative Sanctions.) The *State of Wisconsin v Loomis*[81] case highlights the challenges in understanding how the AI algorithm reaches its conclusions. The proprietary rights within the software prevented any meaningful insight as to what factors were considered in the

---

[74] *Loomis v State of Wisconsin* (2017) 371 Wis 2d 235.
[75] Alexis Morin-Martel, 'Machine learning in bail decisions and judges' trustworthiness' (2023) 39 *AI & Soc* 2033, 2044.
[76] Law Council Australia, *Safe and Responsible AI in Australia* (2023) 27, 1-39 [117] ('Safe and Responsible AI in Australia').
[77] AI Decision-Making and the Courts (n 22).
[78] Victoria Jennett, Fighting Judicial Corruption: Topic Guide (Report, 31 October 2014) 1, 6.
[79] *Hemmett v Market Direct Group Pty* Ltd [2018] WASC 214
[80] AI Decision-Making and the Courts (n 22).
[81] *State of Wisconsin v Loomis* (2016) 881 N.W.2d 749 (*'Loomis'*).

AI tool's recommendations.[82] This has significant implications for the overall sentencing of offenders including the potential imprisonment of offenders or the prolonging of imprisonment.[83] Courts can reduce these risks through placing guidelines on the use of such tools by a decision maker. For example, requiring that the decision makers are educated on the inherent limitations regarding AI tools and therefore "adjust their expectations of the tool to ensure appropriate use."[84] Decision makers would then be compelled to interrogate the basis and integrity for the recommendation, minimising the risk of judicial partiality. In high-stake instances where AI assisted or fully automated decisions suffer from a significant lack of explainability and increased opacity, such tools should be avoided by the judiciary entirely.

Recommendations

4. Require AI systems used in courts and tribunals to provide sufficient explanations of how they arrive at specific conclusions or recommendations.

Question 7: Should some AI uses be prohibited at this stage?

Uses of AI have been categorised into risk categories measured by their ability to impact a matter and whether the consequences of the use of AI will impact fundamental human rights or have irreversible consequences. Certain uses of AI such as predictive analytics, judicial analytics, virtual legal advice and judicial determination are examples of high-risk uses of generative AI. This is because these uses of AI within the legal system may lead to sever impacts on fundamental human and legal rights and may also have irreversible consequences.

I Predictive Analytics

Predictive analytics is a tool used in law enforcement, with major progression and usage being undertaken by government and policing bodies for predictive policing and risk assessment. Its use in the criminal justice system as a risk assessment, in principle, allows law enforcement agencies 'to better provide for public safety with the least restrictive means necessary', predict

[82] Monika Zalnieriute, Submission No 3 to Australian Law Reform Commission, *Technology and the Courts: Artificial Intelligence and Judicial Impartiality* (04 June 2021) 1, 5 ('Technology and the Courts submission')*.*
[83] Paul Daly, 'Artificial Administration: Administrative Law, Administrative Justice and Accountability in the Age of Machines' (2023) 30 *Australian Journal of Administrative Law* 95, 106.
[84] Monika Zalnieriute, Lyria Bennett Moses and George Williams, 'The Rule of Law and Automation of Government Decision-Making' [2019] 14 *The Modern Law Review* 1, 9 ('The Rule of Law and Automation')*.*

patterns, and allow for more accurate and efficient use of police force and intelligence resources.[85]

However, there are concerns regarding the accuracy, fairness and transparency of the AI tools,[86] which has led to violations of human rights and have undermined the rule of law. Berk asserts that the best example of the controversy surrounding the use of AI in predictive analytics is Correctional Offender Management Profiling for Alternative Sanctions (**COMPAS**).[87] COMPAS is an automate decision making process utilised by judges in some jurisdictions in the United States which uses historic data to infer which convicted defendants post the highest risk of reoffending, particularly violent re-offending.[88] COMPAS is used by judges to assist in determining the provision of bail and length of sentence. However, an article published by ProPublica in 2016 highlighted how the instrument was racially and gender bias.[89] Within this article, ProPublica provided various examples of cases across the jurisdictions of the United States such as Florida, where African American defendants, charged with lesser crimes and a lower criminal history to Caucasian defendants being rated at a higher risk of re-offending by COMPAS and therefore, either sentenced to longer imprisonment sentences or denied bail.[90] The formula used by COMPAS, which Northpoint, the company behind the software denied to publicly disclose, falsely flagged African American defendants 77% more likely to commit future violent crimes and 45% more likely to commit future crimes of any kind.[91] Then US Attorney, Eric Holder, warned that COMPAS may be 'injecting bias into courts' and 'inadvertently undermining efforts to ensure individualised and equal justice'.[92]

Further, in the case of *Loomis* gender discrimination was raised as an issue by the defendant who argued that the algorithm used in COMPAS was trained and formulated to use historic data to of statistics of violent re-offending in men and women, would not maximise predictive accuracy.[93] This ultimate impact on individual liberty, a fundamental human right, with serious and possibly

[85] Richard Berk 'Artificial Intelligence, Predictive Policing, and Risk Assessment for Law Enforcement' (2021) 4(1) *Annual Review of Criminology* 209, 209.
[86] Ibid.
[87] Ibid 210.
[88] The Rule of Law and Automation (n 84) 437.
[89] Julia Angwin et al, 'Machine Bias' *ProPublica* (online, 23 May 2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
[90] Ibid.
[91] Ibid.
[92] Ibid.
[93] The Rule of Law and Automation (n 84) 437.

irreversible consequences requires instruments such as COMPAS when used in sentencing to be of a higher degree of transparency.[94]

In London a system called Gangs Matric was a database and predictive tool developed by the Metropolitan police to assist in predicting gang-related violence.[95] In 2022, The Metropolitan Police faced legal action due to Gangs Matrix, where the plaintiff's successfully argued that the use of the tools was unlawful based on grounds that it was racially discriminatory and contravened individual rights as provided under Article 8 of the Human Rights Act 1998 in the UK.[96] The action led to the Metropolitan Police settling the claim and admitting that the operation of the database was unlawful due to it being discriminatory.[97]

Recommendations

5. Prohibit the use of predictive analytic tools to operate as a standalone instrument for judicial determination.
6. Where predictive analytics are being utilised by judges and tribunals, ensure they are used in combination with independent and competent human oversight and merely as a tool to assist decision-making rather than a determinative instrument for judicial determination.

II Judicial Analytics

Predictive judicial analytics have been used by courts in international jurisdictions to predict outcomes using patterns drawn from prior judicial determinations. This has enabled lawyers and court users to predict how cases will be decided. It is particularly helpful in the context of judges and other decisionmakers, as it enables them to determine whether or not to pursue a claim and how best to argue it.[98] It aims to enhance public confidence by holding Court systems and judges accountable and enhancing the efficiency of the judicial system by eliminating causes of action with a pattern demonstrating low chances of success.

[94] Ibid.
[95] Harriet Green, 'Consciousness over Code: How Judicial Review can Address Algorithmic Decision-making in Policing' (2024) 5(1) *York Law Review* 8, 18.
[96] Ibid.
[97] Rachel Pain 'Escaping the Matrix: Met Admits Gangs Matric Unlawful' *Mountford* (Online,18 November 2022) <https://www.mountfordchambers.com/escaping-the-matrix-met-admits-gangs-matrix-unlawful/>
[98] Pamela Stewart and Anita Stuhmcke 'Judicial Analytics and Australian Courts: A Call for National Ethical Guidelines' (2020) 45(2) *Alternative Law Journal* 82, 84-85.

However, the use of judicial analytics has been criticised to undermine the rule of law by using data and patterns that ignore legal precedent and the specifics of individual cases. Pamela Stewart and Anita Stuhmcke argued 'This use of data, or legal information as data, has the potential to determine litigation and courtroom tactics, the legal arguments advanced by counsel and may influence lawyers approaches to legal doctrine. The prediction of judicial outcomes will influence final litigation results, as litigants withdraw or settle claims on the basis of predictions. Ultimately over time, these subtle influences can shape legal principle.'[99]

For this reason, the French Government banned the use of judicial analytics on the basis that 'permitting judicial profiling could lead to undesirable pressures on judicial decision making and strategic behaviour by litigants'.[100] This is because AI is using machine learning rather than legal reasoning to assumed facts rather than facts found at trial inevitably asserting truths based on 'science' and possibly denying access to justice to those relying on judicial predictions for their cause of action.[101]

Furthermore, these AI tools use data based on past decisions, meaning that settled cases are excluded, and leading to datasets composed primarily of outliers. Legg and Bell argue that in a small jurisdiction such as Australia, which has far fewer judgements in general on which predictions are based on, could lead to further misleading and skewed data.[102]

Whilst it may assist in providing insights as to the likelihood of success for a possible claim and assist in holding judges accountable, similar legal advice about the likelihood of success may be obtained from experienced and knowledgeable legal professionals, and courts are able to be held accountable through the system of appeals.

Recommendations

7. Prohibit the use of judicial analytic tools.

[99] Ibid.
[100] Jena McGill and Amy Salyzyn 'Judging by the Numbers: Judicial Analytics, the Justice System and its Stakeholders' (2021) 44(1) *Dalhousie Law Journal* 249, 250.
[101] Michael Legg and Felicity Bell, 'Artificial Intelligence and the Legal Profession: Becoming the AI-Enhanced Lawyer' (2019) 38(2) *University of Tasmania Law Review* 34, 37 ('Artificial Intelligence and the Legal Profession'); Pamela Stewart and Anita Stuhmcke, 'Judicial Analytics and Australian Courts: A Call for National Ethical Guidelines' (2020) 45(2) *Alternative Law Journal* 82, 84-85.
[102] Artificial Intelligence and the Legal Profession (n 101) 50.

Chapter 5: Regulating AI: The Big Picture

Question 8: International Perspectives

I Brazil

A *Overview of Brazil's Approach*

Brazil has been proactive in developing and regulating AI in various sectors including the judiciary. In support of AI's development, the AI intelligence plan[103] included $23 billion dollars in funds for research and development of AI over the next couple of years. [104] A key reason Brazil is embracing the use of AI in the judiciary is to combat high case overloads in its courtrooms and reduce costs.[105] Forty-seven Brazilian Courts have been implementing AI since 2019.[106]

Though there are various resolutions and ordinances mentioning AI in the judiciary,[107] the most relevant to the VLRC is resolution 332/2020[108] that the national Council of Justice (**CNJ**) produced to provide national guidelines for the use of AI in the judiciary. In its creation the Brazilians considered European materials including "White Paper on Artificial Intelligence: a European approach to excellence and trust"[109], and the "European Ethical Charter on the use of AI in judicial systems and their environment".[110]

---

[103] Ministério da Ciência, Tecnologia e Inovação [Ministry of Science, Technology and Innovation], *Plano Brasileiro de Inteligência Artificial (PBIA) 2024-2028 [Brazilian Artificial Intelligence Plan]* (August 2024) <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2024/07/plano-brasileiro-de-ia-tera-supercomputador-e-investimento-de-r-23-bilhoes-em-quatro-anos/ia_para_o_bem_de_todos.pdf/view>.

[104] Ibid [17].

[105] Attorney General's Office, *'AGU Passa a Utilizar Ferramentas de Inteligência Artificial na Produção de Documentos Jurídicos'* [AGU Starts Using Artificial Intelligence Tools to Produce Legal Documents] (September 2024) <https://www.gov.br/agu/pt-br/comunicacao/noticias/agu-passa-a-utilizar-ferramentas-de-inteligencia-artificial-na-producao-de-documentos-juridicos>. ; 'Reuters, *'Brazil Hires OpenAI to Cut Costs of Court Battles'* (Web Page, 11 June 2024) <https://www.reuters.com/technology/artificial-intelligence/brazil-hires-openai-cut-costs-court-battles-2024-06-11/>.

[106] Luis Felipe Salomão,FGV Conhecimento *Tecnologia Aplicada à Gestão Dos Conflitos No Âmbito Do Poder Judiciário Brasileiro* (Report, 2nd ed, December 2020)(<https://ciapj.fgv.br/publicacoes> 10.('FGV Report')

[107] *Resolution no. 370* [Resolução Nº 370] (Brazil) January 2021; *Resolution no. 363* [Resolução Nº 363] (Brazil) January 2021; *Ordinance 271* [Portaria Nº 271] (Brazil), December 2020 (*'Ordinance* 271')

[108] *Resolution no. 332/2020* [Resolução Nº 332] (Brazil) August, 2020 *('Resolution no. 332/2020'*).

[109] European Commission, *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust* (Report, February 2020); Thiago Santos Rocha, "Brazil" in Dariusz Szostek and Mariusz Zalucki (ed) *Legal Tech Information technology tools in the administration of justice (European Law Institute* 2023) 487, 496. <https://doi.org/10.5771/9783748922834-487> ('P Use of Information Technology Tools'*)*.

[110] European Commission for the Efficiency of Justice (CEPEJ), *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment* (Report, Council of Europe, December 2018).
P Use of Information Technology Tools (n 109).

There are various articles in the resolution which Victoria should consider when developing its own policy approach. The relevant articles and the reasoning as to why they should be considered in Australia are listed in the table below:

Table 1: Key articles from resolution 332/2020 and why they should be implemented in Victoria

Article number	Topic	What it says?	Why should it be considered in Victoria?
7	Bias reduction measures	(1) AI must preserve; "equality, non-discrimination, plurality and solidarity, assisting in fair judgment, creating conditions that aim to eliminate or minimise oppression, marginalisation of human beings and errors of judgment resulting from prejudices." [111] (2) Approval is needed for an AI model before production in order to determine whether prejudices/ discriminatory tendencies exist [112] (3) If prejudices cannot be addressed then the AI model will be discontinued. [113]	Having AI models approved by a body to pre-screen for bias and prejudice before it is used is a way to mitigate the risks of AI. Additionally, by discontinuing the program if prejudice cannot be resolved is a way to permit for the development of AI and allow more people to access the judicial system while reducing AI's inherent risks. Additionally, it upholds the fairness of judgement and in this way protects those who use the justice system.[114] Victoria should consider having a similar structure in its policies as a way to tackle bias.
20	Bias reduction measures	Brazil looks to ensure diversity in gender, race, ethnicity, colour, sexual orientation and people with disabilities in its teams from research and developments to implementation of AI tools. In this way, Brazil tries to minimise the possibility of inherent biases that come with AI in all of its stages including algorithms. [115]	As bias is a risk in the use of AI,[116] it is important to address this issue early. By creating specialised teams to design and implement AI that are diverse you minimise the possibility of bias as the designers themselves will have various perspectives. This article will be most effective in conjunction with other measures to reduce bias such as the ones in article 7.

[111] *Resolution no. 332/2020* (n 108) art. 7 (1).
[112] Ibid art. 7 (2).
[113] Ibid art. 7 (3).
[114] *Charter of Human Rights and Responsibilities Act 2006* (Vic) s 24 ('The Charter').
[115] *Resolution no.* 332/2020 (n 108) art. 20.
[116] Md Abdul Malek, 'Criminal courts' artificial intelligence: the way it reinforces bias and discrimination' (2022) 2(1) *AI and Ethics* 233, 233-234 <https://doi.org/10.1007/s43681-022-00137-9>.

Article number	Topic	What it says?	Why should it be considered in Victoria?
22	Authorisation necessary for facial recognition development	The CNJ must authorise the use of facial recognition before it is implemented. This is not necessary for other AI models rather courts must only notify the CNJ of its development/ implementation. [117]	By keeping the national body in charge of the judiciary informed and aware of the use of AI in courts it is easier to track AI's progress and its impact. Additionally, by requiring authorisation before implementing facial recognition a separate body could oversee the ethics before allowing or preventing its use and in this way, there is a harmonised system in place to control risks.
23	AI in criminal matters	"The use of Artificial Intelligence models in criminal matters should not be encouraged, especially with regard to the suggestion of predictive decision models." With the exception of "calculation of sentences, prescription, verification of recidivism, mappings, classifications and sorting of records for collection management purposes" [118]	Decision making outputs by AI should never be unsupervised by a judge or be authority alone. Even so, criminal matters are sensitive as they deal with the freedom of a human. Therefore, they need extreme care which is what Brazil seems to be trying to achieve. However, it must be noted that the article states "not be encouraged" but does not specifically prevent the use of predictive decision models in criminal cases being used. If applied in Victoria predictive models should not be used.
19	Explanation of AI decision making	When using AI models for judicial decisions, it is a criterion that the AI "define the technique used", explain "the steps that led to the result, in addition to allowing the supervision of the competent Judge" [119]	Explainability features in AI models are vital for the integration of AI in a courtroom. [120] This is because judges over seeing outputs from AI must be able to follow the logic to their outcome to make sure that it is appropriate and that it can be explained in appeals and judicial reviews.

[117] *Resolution no. 332/2020* (n 108) art. 22.
[118] Ibid art. 23.
[119] Ibid art. 19.
[120] Ashley Deeks, 'The Judicial Demand for Explainable Artificial Intelligence" (Essay, *Columbia Law Review*, 2019).

B Lessons Learned

1 *Centralisation of AI programs used for the judiciary*

In addition to the 332/2020 resolution, Ordinance no 271[121] the CNJ established a centralised platform called 'Sinapses' for the development of AI tools specifically designated for the judiciary. [122] By having a centralised platform, it is easier for organisations and national law bodies to observe the type of technologies available and prevent the same technologies from being developed. Moreover, it helps set up a database of AI tools which could be helpful for tracking the impact of AI while also being convenient for those who use the tools to have AI programs in one platform. However, it must be noted that in Brazil there are other AI projects beyond Sinapses.[123] Australia should try to adhere to one platform only.

2 *Developing AI software for judiciary*

According to a survey by Felipe Salomão[124] beyond Sinapses, Brazil has 64 AI projects for the judiciary of which 47 are developed by the internal teams within the courts themselves. [125] While Brazil does have projects being created by private companies, it is a good sign that many courts are developing the AI projects themselves. This is because, in this way, there are fewer commercial interests and possibly less security risks in doing so. Victoria should aim to have its own teams within the courts developing AI programs specifically designed for the courts internal needs rather than outsourcing the project to private companies, if the budget permits this.

3 *Future Actions by Brazil*

Brazil may also be successful in implementing a risk-based regulation[126] to manage and protect individual rights of those who have been affected by AI. Such rights could include:

1) right to be informed if AI was used in the individual's case
2) right for the individual to challenge the decision made by AI and seek human review instead.

Victoria should follow *Projeto de Lei No. 2.338/2023* to see whether Victoria should implement similar rights.

---

[121] *Ordinance 271* [(n 108).
[122] Ibid art. 4.
[123] FGV Report (n 106).
[124] Ibid.
[125] Ibid 10
[126] *Projeto de Lei No. 2.338/2023 [Bill No. 2,338/2023] (Brazil).*

II Canada

A *Overview of Canada's Approach*

In September 2024 the Canadian Judicial Council (**CJC**) released 7 guidelines for the use of AI in the judiciary.[127] These guidelines were created to "raise awareness about the risks of AI"[128] and have been summarised in a table below along with possible recommendations for Victoria.

[127] Martin Felsky and Karen Eltis, *Guidelines for the Use of Artificial Intelligence in Canadian Courts* (Published Guidelines No 1, September 2024) ('Guidelines for the Use of Artificial Intelligence in Canadian Courts')
[128] Canadian Judicial Council, 'Canadian Judicial Council issues Guidelines for the Use of Artificial Intelligence in Canadian Courts' (Press Release, October 2024) 1 < https://cjc-ccm.ca/en/news/canadian-judicial-council-issues-guidelines-use-artificial-intelligence-canadian-courts>.

Table 2: Canadian guidelines on AI regulation in the courtroom and how they can implemented in Victoria

	Guideline	Further information	Identified risks or mitigation if any posed in document	Recommendation for Victoria
1	"Protect Judicial Independence"[129]	Must be careful that AI and its regulations do not interfere with judicial independence and agency	**Risk:** with the use of AI judicial independence can be lost like for example with dispute resolutions[130] **Mitigation**: therefore, judges must have oversight over dispute resolution that happens beyond the court.[131]	Victoria and Canada share the same independent branches of government, and it should be a priority for Victoria that Victorian courts remain independent from other branches of government.[132] Therefore, the judiciary needs involvement in developing guidelines and policies with the use of AI and be able to have oversight over their uses regardless of laws introduced by government.
2	"Use AI Consistently with Core Values and Ethical Rules" [133]	If a judge is to use AI, it must be in line with the core values of the court including "Independence, integrity and respect; diligence and competence; equality and	**Mitigation:** help with bias detection and maintaining trust and order for the public [135]	Using AI ethically and in line with the values of the court should be at the forefront for framing Victoria's policy on AI use in the judiciary. This is because by maintaining values (such as the ones described in the Canadian Guidelines) judges will have an anchor point from which he or she may be able to decide

[129] Guidelines for the Use of Artificial Intelligence in Canadian Courts (n 127) 6.
[130] Ibid.
[131] Ibid.
[132] Commonwealth of Australia *Constitution Act* (Cth) 1901 ss 1, 61, 71
[133] Guidelines for the Use of Artificial Intelligence in Canadian Courts (n 127) 7.
[135] Ibid.

	Guideline	Further information	Identified risks or mitigation if any posed in document	Recommendation for Victoria
		impartiality, fairness, transparency accessibility, timeliness and certainty."[134]		whether the application of AI for a particular use is appropriate. In holding to the court's values, the judge may more easily be able to detect biases if they exist (which are a known risk in the use of AI in courtrooms) and act accordingly.
3	"Have Regard to the Legal Aspects of AI Use" [136]	"Court administrators and Chief Justices must ensure that the integration of AI into any court process must consistently adhere to applicable laws, including those governing privacy, intellectual property, and criminal activities"[137]	**Risks** 1) Courts should pay attention to the material used to train AI programs. Sensitive data may be used. 2) Copyright allegations over AI without permission may occur 3) Data used to train generative AI may have been unlawfully obtained in some jurisdictions while not in others.	When drafting new policies Victoria should consider making policies that will mitigate the risks of AI breaking the law via copy right, or data infringements for example. One way of mitigating this risk is by the court having control over what material is being used when training AI programs. Sensitive data of sealed court documents should not be used to train the programs. There needs to be a balance between been able to train the AI for it to be effective without laws or ethics being breached.

[134] Ibid.
[136] Ibid.
[137] Ibid.

	Guideline	Further information	Identified risks or mitigation if any posed in document	Recommendation for Victoria
4	"AI tools must be subject to stringent information security standards (and output safeguards)" [138]	AI has inherent security and privacy risks that need careful protection and oversight	**Risks:** 1) Revealing sensitive data 2) Algorithm tampering 3) Non-secure third-party AI providers [139]	Victoria should consider these risks and vet third-party providers and have cybersecurity teams or other methods in place to mitigate these risks. Additionally, judges should be trained not to input sensitive information into AI software so as not to risk it being used inappropriately or exposed to unauthorised parties.
5	"Any AI tool used in court applications must be able to provide understandable explanations for their decision-making output"[140]	Explain-ability of AI allows for there to be accountability for its outputs. [141]	**Mitigation:** A feature to explain AI decision making prevents there from being uncertainty from the outputs and whether juris prudence for example is being correctly upheld. [142]	Explain -ability is yet another risk and deficit of some AI tools.[143] By allowing only AI tools that can explain decision outputs to courts and the public there is a greater sense of accountability and transparency in the judicial process. This is particularly important for appeals and judicial reviews. Additionally, by AI being able to explain its decision a judge or member of the court may be able to follow the AI's

[138] Ibid 8.
[139] Ibid 8.
[140] Ibid 8-9.
[141] Ibid.
[142] Ibid.
[143] Hans de Bruijn, Martijn Warnier and Marijn Janssen, 'The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making' (2022) 39(2) *Government Information Quarterly* 101666, 3 <https://doi.org/10.1016/j.giq.2021.101666>.

	Guideline	Further information	Identified risks or mitigation if any posed in document	Recommendation for Victoria
				"logic" and determine whether the reasoning applied to the decision is appropriate and therefore should be maintained or changed. Victoria should have a similar policy to maintain public trust and its values in the court room.
6	"Courts must regularly track the impact of AI deployments" [144]	Before AI is introduced it must be tested to see whether it maintains "judicial independence, workload, backlog reduction, privacy, security, access to justice, and the court's reputation"[145]	**Mitigation:** Run a pilot program and run impact assessments in intervals to assess how AI is doing. [146]	Victoria should implement some program that allows the impact of AI to be continuously monitored. In this way, the ethics and risks to AI can be assessed on a continual basis and if circumstances change where AI is no longer benefiting as much as it is posing a risk to the court or security etc. than it can be stopped before a situation develops.
7	"Develop a program of education and provide user support" [147]	"AI should not be employed without users undergoing a comprehensive educational	**Mitigation:**	Implementing some educational system to teach users such as judges how to use the AI programs is crucial.[150] They should

[144] Guidelines for the Use of Artificial Intelligence in Canadian Courts (n 127) 9.
[145] Ibid.
[146] Ibid.
[147] Ibid.
[150] Shukhrat Chulliev, 'Transforming Judicial Competencies: A Framework For Judge Training And Qualification In Ai-enhanced Court Systems' (2024) (3(2) *Elita* 1,2.

	Guideline	Further information	Identified risks or mitigation if any posed in document	Recommendation for Victoria
		process and understanding best practices for interacting with the technology" [148]	Guideline made to minimise the risks posed by the use of AI. [149]	be taught what the risks are with the AI tool and how to use it in a way that upholds the values and ethics of the court. Users should also be educated on errors and warnings to look out for when using AI like for example inherent bias in coding.

[148] Ibid.
[149] Ibid.

B Lessons Learned

Along with the guidelines, the document's preamble outlines Canada's stance on AI decision-making powers. Their stance being that the authority to make judicial decisions should never be handed over to AI systems but rather should be used as a tool to help judges reach decisions.[151] This stance seems to be an international one and should be replicated in Victoria in order to ensure that judicial independence is maintained.

III China

A *Overview of China's Approach*

China's approach to AI regulation has been through the creation of regulation on specific issues within the industry. This has then led to the creation of a multi-faceted platform which serves as the groundwork for the creation of further national legislation in the near future. When evaluating China's approach, three key reforms elucidate the impactful position that China is building. These reforms include the 2021 regulation on recommendation algorithms, the 2022 rules for deep synthesis (synthetically generated content), and the 2023 draft rules on generative AI.[152] A discussion on these three reforms and the requisite performance that is provided will follow.

China's first foray into AI regulation stemmed from concerns of the Chinese Communist Party (**CCP**) regarding algorithms disseminating information online that threatened the ability of the CCP to set the agenda of public discourse.[153] At the same time, these capabilities raise significant risks, including introducing the ability to mass produce plausible misinformation, amplify hate speech campaigns, propagate bias, and displace jobs.[154] The draft, released in 2021, of the *Provisions on the Management of Algorithmic Recommendations in Internet Information Services* (*"Algorithmic Recommendations"*),[155] provided regulation that required 'algorithmic recommendation service providers "uphold mainstream value orientations" and "actively transmit positive energy".[156] Embedded in the *Algorithmic Recommendations*[157] was a new

---

[151] Guidelines for the Use of Artificial Intelligence in Canadian Courts (n 127) 3; Commission on Judicial Conduct, *Improper Delegation of Judicial Authorities* (Annual Report 1 March 2019) 22-24.
[152] Matt Sheehan, 'China's AI Regulations and How They Get Made' (2023*) Carnegie Endowment for International Peace* 4 ('Sheen').
[153] Ibid 12.
[154] Gilad Abiri, Yue Huang, 'A Red Flag? China's Generative AI Dilemma' (2023) 37 *Harvard Journal of Law & Technology* 2 ('Abiri').
[155] 互联网信息服务算法推荐管理规定 [Provisions on the Management of Algorithmic Recommendations in Internet Information Services] (People's Republic of China) Ministry of Industry and Information Technology of the People's Republic of China, 31 December 2021 ('Algorithmic Recommendations').
[156] Sheehan (n 153) 12.
[157] Algorithmic Recommendations (n 156).

requirement, an algorithm registry function. The Algorithm Registry is an online database of algorithms, with developers required to 'submit information on how their algorithms are trained and deployed, including which datasets the algorithm is trained on and perform a security self-assessment report'.[158]

The CCP's second regulation, at a similar time in 2022, was to contend with deepfakes, or 'deep-synthesis' as the party named it. Subsequently, the *Provisions on the Administration of Deep Synthesis Internet Information Services* ('Deep Synthesis')[159] was drafted. The new regulation included requirements that 'content conform to information controls, that it is labelled as synthetically generated, and that providers take steps to mitigate misuse'.[160] Further, the *Deep Synthesis*[161] regulation included various ambiguous provisions surrounding content, including the requirement to adhere to the correct political direction and not disturb economic and social order.[162] Finally, content creators must also register their products with the Algorithm Registry.

The third seminal reform was the *Measures for the Management of Generative Artificial Intelligence Services* (***AI Management***)[163] in 2023, the first in the world.[164] This regulation requires registering content with the Algorithm Registry and to 'embody core socialist values'.[165] However, additional requirements which may prove difficult to adhere to were also included. The *AI Management*[166] regulation requires creators ensure the 'truth, accuracy, objectivity, and diversity of their training data',[167] setting a potentially impossible standard for AI that is trained on data and images from across the internet. Further, generated content must not infringe on IP rights.[168]

Through the three regulatory reforms listed above, China has implemented tools that can be used for future regulatory creation. This includes the Algorithm Register, a standardised tool that can be rolled out to assist in the implementation of further regulatory reform. Further, the approach of each reform has been vertical in nature, meaning the proposed reforms specifically target

---

[158] Sheehan (n 153) 13.

[159] 互联网信息服务深度合成管理规定 [Provisions on the Administration of Deep Synthesis Internet Information Services] (People's Republic of China) Ministry of Industry and Information Technology of the People's Republic of China, 25 November 2022 ('Deep Synthesis').

[160] Sheehan (n 153) 13.

[161] Deep Synthesis (n 160).

[162] Sheehan (n 153) 13.

[163] 生成式人工智能服务管理办法 [Measures for the Management of Generative Artificial Intelligence Services] (People's Republic of China) Cyberspace Administration of China, 11 April 2023 ('Generative AI').

[164] Barbara Li and Amaya Zhou, 'Navigating the Complexities of AI Regulation in China,' Reed Smith In-depth (Blog Post 7 August 2024).

[165] Sheehan (n 153) 14.

[166] Generative AI (n 164).

[167] Sheehan (n 153) 14.

[168] Ibid.

applications of a technology.[169] Any concerns of the CCP are identified and regulations imposed specific to those concerns.[170] As each reform is in itself iterative in nature, the CCP can tailor a new regulatory reform if a flaw is identified, thus building a series of vertical pillars to serve as a foundation. This process has led to the formulation of an Artificial Intelligence Law, which is still in the drafting stages, to sit horizontally across the vertical pillars, much as in the European Union.

B *Lessons Learned*

China's inherent motivation is information control. The technology needs to serve the CCP agenda which results in political and social stability.[171] For the CCP, a 'technology to be productive must first be tamed'.[172] Victoria need not be so imposing in the implementation and creation of regulatory reform. To follow the model of China, vertical and iterative reform that addresses specific technological issues, would be beneficial as this would develop a guideline of principles. These principles would then form the pillars for a legislative piece to be built upon once the time arose.

IV United States of America

A *Overview of the United States' Approach*

The United States adopts a policy-driven, sector-specific approach to AI regulation. Rather than enacting a comprehensive federal law for AI, governance relies on existing regulatory frameworks, sector-specific guidelines, and executive actions. Regulation occurs through sectoral agencies such as the Federal Trade Commission (FTC) and the Food and Drug Administration (FDA), as well as executive orders. The 2023 AI Executive Order, for instance, mandates federal agencies to implement principles of accountability, fairness, and transparency in AI use.[173] This approach emphasises innovation while addressing risks associated with AI deployment through oversight mechanisms tailored to the context of use.[174]

The US framework highlights several core principles:

---

[169] Ibid 16.
[170] Ibid.
[171] Ibid 17.
[172] Ibid.
[173] Vicki Birchfield, 'From Roadmap to Regulation: Will There Be a Transatlantic Approach to Governing Artificial Intelligence?' (2024) 46(7) *Journal of European Integration* 1053, 1056-8 ('From Roadmap to Regulation').
[174] Filippo Pesapane et al, 'Artificial Intelligence as a Medical Device in Radiology: Ethical and Regulatory Issues in Europe and the United States' (2018) 9 *Insights into Imaging* 745, 749-751 ('Artificial Intelligence as a Medical Device in Radiology').

- **Risk-Based Oversight:** Applications are regulated based on their risk, with higher-risk uses (e.g., in healthcare or criminal justice) subject to stricter oversight.
- **Transparency and Explainability:** Ensuring AI systems are understandable and decisions contestable remains central to US policy, as emphasised by the 2023 AI Executive Order.[175]
- **Human Accountability:** Judicial systems retain ultimate accountability for decisions involving AI, emphasising AI as a tool to assist, not replace, human judgment.[176]

B Lessons Learned

The US employs a tiered regulatory framework that matches oversight to the level of risk posed by an AI application. High-risk uses, such as AI systems involved in criminal justice or healthcare decisions, undergo more rigorous scrutiny, including pre-implementation assessments, impact evaluations, and ongoing audits. Lower-risk applications are subject to lighter regulatory touchpoints to promote efficiency and innovation.[177] For Victorian courts, this model could provide a structure where judicial applications of AI—such as predictive tools or sentencing aids—are held to higher standards of testing and accountability, while administrative uses like scheduling or document review are regulated more flexibly.

Transparency is a cornerstone of the US approach to AI governance. The 2023 AI Executive Order highlights the importance of making AI systems understandable and accessible to the public and stakeholders.[178] This is particularly relevant for judicial settings, where trust and accountability are paramount. Transparency in AI systems is foundational; black-box models inherently conflict with judicial values.[179] Victorian courts should adopt mandatory disclosure requirements, ensuring that AI-assisted decisions are accompanied by clear, plain-language explanations. This would enable litigants and court users to understand how AI has been used in their cases, safeguarding procedural fairness and contestability.

The US model emphasises human oversight to prevent over-reliance on AI. AI tools in judicial contexts are intended to assist decision-making, not replace human judgment. For example, the FDA's oversight of medical AI requires that human operators remain accountable for decisions made with AI support.[180] Victorian courts can adopt a similar approach by ensuring

---

[175] From Roadmap to Regulation (n 175) 1056-8.
[176] Giusella Finocchiaro, 'The Regulation of Artificial Intelligence' (2022) 39(4) *AI & SOCIETY* 1961, 1964.
[177] Ibid; Artificial Intelligence as a Medical Device in Radiology (n 176) 749.
[178] From Roadmap to Regulation (n 175) 1058.
[179] John Zeleznikow, 'Algorithmic Justice Symposium' (Conference Paper, University of Newcastle, 14 July 2023), 7 ('Algorithmic Justice Symposium').
[180] Artificial Intelligence as a Medical Device in Radiology (n 176) 751.

that judicial officers retain full responsibility for AI-assisted decisions. Limiting AI to tasks such as case filtering or document organisation, rather than substantive decision-making, will preserve judicial independence and integrity.

Recognising the evolving nature of AI, the US approach involves collaboration between government agencies, private-sector innovators, and international bodies. This co-regulatory framework allows for the development of dynamic standards that adapt to emerging technologies while maintaining consistency across sectors.[181] For Victoria, engaging industry experts, legal practitioners, and academia in regulatory development could ensure that the framework remains current and relevant, fostering innovation while upholding ethical standards.

The United States' approach to AI regulation provides valuable lessons for Victorian courts. A risk-based framework, coupled with transparency, human oversight, and adaptive co-regulation, can ensure that AI enhances judicial processes without undermining trust or fairness. By drawing on these principles, Victoria can create a regulatory system that balances innovation with the values at the core of its justice system.

V European Union

A *Overview of the European Union's Approach*

The European Union (EU) has been a global leader seeking to address the risks of AI and being the first major jurisdiction to have a legally binding Artificial Intelligence Act ("EU AI Act")[182] that came into force on 1 August 2024. The consultation paper has effectively incorporated consideration of the regulation of AI in the EU. In summary, the EU AI Act sets out the requirements for AI systems based on the risks posed. Risk categorisation is based on the potential of use to threaten health, safety and fundamental rights, with the EU AI Act primarily applying to 'prohibited' and 'high-risk' use, which specifically includes the administration of justice, but not ancillary administrative tasks. This comprehensive legislation includes oversight structures and provision of enforcement resources. Similar to efforts in Australia, design of AI regulation in the EU has been designed for consistency with the US, prioritising the need to facilitate business and trade between the regions. Criticisms include practical concerns for the standards set for 'trustworthiness' of AI outputs, the costs of compliance may be stifling for small organisations,

[181] Maria Cantero Gamito and Christopher Marsden, 'Artificial Intelligence Co-Regulation? The Role of Standards in the EU AI Act' (2024) 32(1) *International Journal of Law and Information Technology* 14.
[182] *Regulation (EU) 2024/1689 (Artificial Intelligence Act)* [2024] OJ L 2024/1689, Ch I Art 3(63).

and that the burden of compliance will limit innovation. Despite the benefits of certainty that come with comprehensive legislation, as we enter 2025, no significant AI technologies are known to have been implemented in EU courts.

Specifically, applicable when considering the use of AI in the court system, the European Commission for the Efficiency of Justice (**CEPEJ**) drafted the *European Ethical Charter on the Use of AI in Judicial Systems and their Environment* (**Charter**) in December 2018.[183] The Charter sets out five principles that require compliance monitoring, record keeping, human oversight, transparency obligations, and a reporting system for incidents. Additionally, GenAI must publish a summary of copyright material used as training data and demonstrate efforts to prevent illegal content. It is unclear if there are any bodies in Australia that intend to develop a similar national charter, or even at a state and territory level.

Since 2019, to thwart misuse of AI, France banned the use of some analytics when processing judicial decisions, such that "the identity data of judges, prosecutors and court staff cannot be reused to evaluate, analyse, compare or predict their real or supposed professional practice". At this stage, this is the only EU member state known to do so by specific legislative clause.

In 2023, the CEPEJ developed an AI risk assessment tool for use in the judicial setting.[184] This is applicable for the assessment of the potential impact and risk of an AI system to comply with the EU ethics Charter, with utility across the member states despite the variation in their respective court systems. Being focused only on the uses and peculiarities of how the judicial system operates, this evaluation tool is less complex. The consultation paper suggests a generic AI risk assessment tool borrowed from the NSW government and developed for use across the NSW public service. It would be valuable to have the respective tools reviewed by those intending to provide governance of AI risk assessment and users of the Victorian judiciary system to understand the feasibility and balance of effort to develop a sector specific AI assessment tool in Australia.

The CEPEJ also provides extensive sector reporting on the Council of Europe member states' judicial systems on various themes such as court efficiency, justice budgets or access to justice, as well as statistical data sheets for each country. The 2024 report provided the main trends

---

[183] European Commission for the Efficiency of Justice (CEPEJ), *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment* (Report, Council of Europe, December 2018) 7–12.
[184] European Commission for the Efficiency of Justice (CEPEJ), *Assessment Tool for the Operationalisation of the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment* (Report No CEPEJ (2023)16 final, Council of Europe, 4 December 2023).

including artificial intelligence in the field of justice, as well as the reduction of the backlog of court cases, but the most recent evaluation report covered 2022 data so findings were significantly impacted by COVID. Court Services Victoria (**CSV**) appears to be the current local body with responsibility for providing the most similar statistics. However, the most recent report indicates that much more limited data is published, and unknown if efforts have commenced to collect information that aligns with needs for setting foundations for decisions to implement AI, as well as being able to measure the benefits obtained from the investment.

The CEPEJ also has a number of working groups to facilitate indepth investigation of issues. Both the Quality of Justice and the Cyberjustice and AI working groups have been working jointly on; the implementation of pilot evaluations of selected AI tools, the development of an AI and Cyberjustice Resource Centre, and an AI Advisory Bureau (**AIAB**). Whilst the purpose is to enhance the adoption of AI, typically the communications from the CEPEJ are overwhelmingly promote an air of caution, to an already conservative audience. It is unknown what the Victorian judiciary will require, and from who, to forge ahead and confidently commit to AI deployment in their courts.

B *Lessons Learned*

As stated above, despite a comprehensive EU AI Act, none of the European judiciaries have completed implementing any AI projects in their courts, concluding it is more than clear legislation that is also required in Australia. In October 2024 Carnat published her research on the EU AI Act in the context of the judiciary and flags the importance of the judiciary being actively engaged with the implementation of AI-based information systems from the outset, and particularly through the design phase.[185] The CEPEJ has been encouraging the importance of digital literacy and facilitating discussion including with users. It is imperative that members of the judiciary to take active responsibility to learn and engage with AI systems, to ensure they understand the potential impact, but also clear on their purpose and objectives for implementing an AI system.[186] The advice of Carnat and the CEPEJ is consistent with the foreword comments by the Hon Justice Jenny Blokland to the survey report that has been frequently referenced in the VLRC consultation paper.[187]

---

[185] Irina Carnat, 'Addressing the risks of generative AI for the judiciary: The accountability framework(s) under the EU AI Act. (2024) 56 *Computer Law & Security Review: The International Journal of Technology Law and Practice* 13
[186] Ibid 14.
[187] AI Decision-Making and the Courts (n 22) 3.

In a recent retrospective of the implementation of information and communication technology (**ICT**), including AI, in the European judiciaries over the past 25 years, the experienced researcher provided interesting insights. Empirical data reported in 2003 identified that projects to implement new technology that were associated with complex projects "that seek to achieve fundamental changes in working practices, … lengthy and costly learning curves and training for judges, the practicing bar, and court staff, are frequently doomed to failure".[188] Fabri makes the point that the European Commission has made significant investment in ICT for the judiciary, which is expected to continue with focused reform on digital transformation.[189] With reference to the EU e-justice strategy 2024-2028, "the major challenges, as usual, are not merely the technology; we need consistent legislation, effective working practices, and a supporting organisation."[190] Fabri emphasises the importance of the role of a single institutional governance structure to the development of ICT in the judicial system, where many Nordic countries have experienced fewer obstacles and delays in implementing successful ICT projects.[191] Finally, the author laments the lack of information sharing between jurisdictions. This is particularly a lost opportunity, as it is more difficult for the judiciary system to attract and retain experienced technical staff, who are enticed to the higher salaries of the private sector, and where law firms lead the implementation of AI tools. The added consequence is the reliance of the private ICT providers to be the source of evaluative information and expected outcomes in the procurement process, increasing the threat to judicial independence. It is unknown if a similar review has been conducted to investigate the successes and failures of past implementation of ICT in Victorian courts, or elsewhere in Australia. Quantitative and qualitative data to what extent implementation of new technologies over the past decades delivered on meeting the objectives for the Victorian judiciary and users, and the reasons for deviations, would be valuable to identify projects most likely to succeed and to mitigate chances of failure.

The VLRC consultation paper and our submission is in the context of the recent proposal for mandatory guardrails for AI in high-risk settings. The similar voluntary guardrails came into effect on 1 September 2024, though it is unknown to what extent these have been implemented. Of note is that the first proposed safeguard is to "establish, implement and publish an accountability

---

[188] Marco Fabri, 'From Court Automation to e-Justice and beyond in Europe' (2024) 15(3) *International Journal for Court Administration 5*.
[189] Ibid *6*.
[190] Ibid 8.
[191] Ibid 15.

process including governance, internal capability and strategy for regulatory compliance".[192] Reflection of the above EU experiences in respect of this priority is worthy of consideration in the Victorian setting and that the consultation paper (Part D) clarifies that the potential governance of AI in the courts is far from settled. The joint research by the Australasian Institute of Judicial Administration (**AIJA**) and UNSW Faculty of Law and Justice[193] has been a valuable resource for the consultation paper. In particular, the view that Australian courts will not be subject to legislative regulation under enforcement by the executive, if a new AI Act is enacted, differentiating the circumstance of the EU. Therefore, the consultation paper proposes four alternatives for governance of AI selecting a body from within the existing system; Court Services Victoria (CSV), its governing body Courts Council, the Judicial College of Victoria, and the Judicial Commission of Victoria. CSV is the employer of all court administrative staff and manages 75 buildings in Victoria, across 66 locations. In 2021 the Victorian Attorney-General's Office (**VAGO**) conducted an audit of CSV. Scathingly, VAGO found that "CSV's governing body, Courts Council, didn't adequately direct CSV's strategy, governance and risk management."[194] Recognising that the audit did prompt immediate improvements, it is unknown if either body has the willingness or competency to meet either or both the governance and provision of comprehensive performance management data needed to provide foundations on which to identify and implement AI safely into Victorian courts and tribunals. Even with the benefits of extensive data provided by the CEPEJ, and established governance structures, the EU has still not yet been able to complete an AI project. In this context, the viewpoints of judiciary systems across Australia would be valuable, for example, if AI is intended to deliver increased efficiency and quality, who is responsible for setting the goals, holding the judiciary systems to account, and what are the relevant measures for success.

VI United Kingdom

A *Overview of United Kingdom's (UK) Approach*

The UK's regulatory framework for AI has been designed to accommodate the complexities and risks of AI across various sectors, opting for a pro-innovation, sector-specific model.[195] It allows

---

[192] Department of Industry, Science and Resources (Cth), *Safe and responsible AI in Australia: Proposals Paper for Introducing Mandatory Guardrails for AI in High-Risk Settings* (Proposal Paper No 1, September 2024) ('Safe and responsible AI in Australia: Proposal Paper').
[193] AI and Decision-Making in the Courts (n 22) 9.
[194] Administration of Victorian Courts. Performance Audit: Court Services Victoria. (Tabled 13 October 2021) *Victorian Auditor General's Office* (Web Page) <https://www.audit.vic.gov.au/sites/default/files/2021-10/AOC_transcript.pdf>
[195] HM Government, *National AI Strategy* (Cm 525, 2021) ('National AI Strategy'); Department for Science, Innovation and Technology, *A pro-innovation approach to AI regulation* (Cm 815, 2023) ('Pro-innovation approach').

for ongoing adjustments that reflect the unique characteristics of different industries, rather than enforcing a rigid, one-size-fits-all framework. It emphasises on principles-based regulation.[196] This approach enables evidence-based regulation, ensuring that adjustments are made based on actual data rather than perceived harm, where the probability of actual occurrence is limited or unknown.[197] The flexibility built into this framework allows for continuous assessment and proportional responses to the risks AI presents. However, the approach has faced criticism for being too fragmented and reactive.[198] Industry has argued that the lack of a comprehensive statutory framework leads to inconsistencies in interpretation and governance, which could hinder the ability to address emerging risks effectively.[199] Despite these criticisms, the approach has been considered pragmatic, particularly because ethical risks vary significantly depending on the AI application.[200] Nevertheless, to mitigate concerns, the UK has introduced a UK AI Bill into the House of Lords.[201] This legislation aims to formally enforce the principles and governance structures laid out in the White Paper. For Victoria, this indicates that while a flexible, principles-based approach can be effective in the short term, there may be a need for an overarching legislative change in the future to ensure consistency in AI regulation throughout the state.

To further address the regulatory gaps and uncertainty, the UK has supported the use of assurance techniques and regulatory sandboxes to help support regulatory compliance and risk management associated with using AI.[202] The UK intends to create an AI assurance ecosystem comprising of tools and services designed to provide meaningful information about AI systems to both users and regulators.[203] An assurance framework can provide a structured approach to monitoring and evaluating AI systems, ensuring they meet defined standards of safety, fairness, and transparency.[204] By setting benchmarks for performance and ethical compliance, these frameworks can help maintain public confidence in AI's role within the judiciary. Assurance techniques like implementing impact assessments before fully integrating AI into the judicial

---

[196] Pro-innovation approach (n 197) 21, 26.
[197] Ibid 19; Robert Glicksman and Sidney Shapiro, 'Improving Regulation Through Incremental Adjustment' (2004) 52 *University of Kansas Law Review* 1179, 1179.
[198] Pro-innovation approach (n 197) 15.
[199] Ibid.
[200] Hwu Roberts et al, Artificial intelligence regulation in the United Kingdom: a path to good governance and global leadership? (2023) 12(2) *Internet Policy Review* 1, 9; Asress Adimi Gikay, 'Risks, innovation, and adaptability in the UK's incrementalism versus the European Union's comprehensive artificial intelligence regulation' (2024) 32 *International Journal of Law and Information Technology* 1, 25.
[201] Artificial Intelligence (Regulation) HL Bill (2023) 11.
[202] Pro-innovation approach (n 197) 56.
[203] Centre for Data Ethics and Innovation, 'The roadmap to an effective AI assurance ecosystem' (Independent Report, 2021).
[204] Pro-innovation approach (n 197) 64-5.

system.[205] They can help identify potential risks and unintended consequences early in the development life cycle, allowing for timely adjustments and improving appropriate safeguards and governance mechanisms.[206] The implementation of assurance measures is essential for building trust in AI and supporting its broader adoption in legal settings.[207] The UK's clear direction for a market-based approach to assurance can facilitate regulatory oversight while encouraging the development of innovative compliance measures by industry leaders.[208] Similarly, the UK's pro-innovation focus has adopted the use of regulatory sandboxes.[209] Sandboxes can offer a safe space for innovators whilst enabling the UK government to understand how regulation interacts with new technologies and refine this interaction where necessary.[210] In these environments, AI tools can be tested in real-world scenarios under the watchful eye of regulators. This allows for the identification of risks and benefits before widespread adoption, enabling more informed decision-making in a relatively resource-efficient way.

However, notably, the UK's lack of a formal centralised AI regulatory framework exposes it to extra-territorial influence from jurisdictions like the EU and US.[211] In the absence of clear domestic regulations, UK companies may need to comply with foreign standards to access international markets, which can undermine the UK's regulatory autonomy and impose additional compliance burdens.[212] This influence risks shifting control over ethical and safety standards away from UK priorities, potentially stifling domestic innovation.[213] Nevertheless, overall, the UK's national approach to AI regulation offers a valuable example of how tailored, adaptive frameworks can be flexibility and adapt to evolving technologies. Building on this foundation, the judiciary has taken proactive steps to establish clear guidelines for the use of AI within the legal system.

B *Overview of the UK's Judiciary Response*

In 2023, the UK's Courts and Tribunals Judiciary issued guidelines which emphasises the need for clear, accessible guidance on the use of AI tools within the legal system.[214] Similarly to the

---

[205] Ibid 64.
[206] Ibid.
[207] Laura Freeman et al, 'The path to a consensus on artificial intelligence assurance' (2022) 55(3) *Computer (IEEE Computer)* 82, 82-3.
[208] Jack Clark and Gillian Hadfield, 'Regulatory markets for AI safety' (2019) *ArXiv* 1, 9-10.
[209] Pro-innovation approach (n 197) 60.
[210] Ibid 60-1; HM Treasury, 'Pro-innovation Regulation of Technologies Review: Digital Technologies' (2023).
[211] Roberts et al (n 202) 19-20.
[212] Ibid 20; Asress Adimi Gikay (n 202) 4.
[213] Roberts et al (n 202) 20.
[214] Courts and Tribunals Judiciary (UK), 'Artificial Intelligence (AI) Guidance for Judicial Office Holders' (2023).

national AI strategy, the UK guidelines are not static, they are subject to ongoing review and adaptation as technology evolves.[215] This flexibility is important to ensure that the guidelines remain relevant and effective. The development of these guidelines involved consultation with all judicial officer holders.[216] Involving stakeholders early in the process can also help identify potential challenges and areas of concern, such as ethical dilemmas, data privacy, and transparency in AI decision-making. For Victoria, engaging with key stakeholders such as judges, lawyers, AI developers, and advocacy groups will be essential in creating a regulatory environment that reflects the complexities of the Victorian legal system. Although these guidelines can be seen as simple or obvious, their publication establishes a clear framework for the judiciary, providing both transparency and accountability in the use of AI. By formalising expectations and standards, the guidelines foster public trust and reduce ambiguity about the role of AI in judicial processes.

C *Lessons Learned*

The UK's approach to AI regulation demonstrates a pragmatic balance between innovation and accountability, leveraging flexibility, stakeholder engagement and sector-specific frameworks to address the diverse challenges AI poses. While criticisms of fragmentation and external influence highlight the need for a more cohesive statutory framework, the use of assurance techniques, regulatory sandboxes and adaptive guidelines offers valuable lessons. For Victoria, UK's experience exhibits the importance of fostering a principles-based, adaptive regulatory environment to prioritise the competing interests of innovation and public trust and ethical compliance. By incorporating similar measures, Victoria can ensure that it's Courts and Tribunals remain responsive to technological advancements whilst maintaining transparency, fairness and accountability.

Question 9: What would the best regulatory response to AI use in Victorian courts and tribunals look like?

I Introduction

The implementation of artificial intelligence in Victorian courts and tribunals has the potential to enhance efficiency, reduce backlogs, and improve access to justice. However, it also presents

---

[215] '*Artificial Intelligence (AI) – Judicial Guidance*', Courts and Tribunals Judiciary (UK) (Web Page) <https://www.judiciary.uk/guidance-and-resources/artificial-intelligence-ai-judicial-guidance/>.
[216] Ibid.

significant risks, including bias, lack of transparency, and threats to judicial independence. A regulatory framework must balance these benefits and risks, ensuring public trust, fairness, and accountability while fostering innovation. Drawing on international practices, the most effective regulatory response for Victorian courts must incorporate risk-based frameworks, technological neutrality combined with specific regulations, and robust oversight mechanisms.

II Core Principles of Effective AI Regulation

The regulatory framework for AI in Victorian courts should be formed around the following principles:

1. **Transparency and Explainability:** Judicial AI systems must provide explainable outputs and auditable decision-making trails to ensure accountability.[217]
2. **Fairness and Bias Reduction:** Measures to prevent discrimination must include pre-approval processes to evaluate potential biases and diverse teams in AI development.[218]
3. **Human Oversight and Accountability:** AI should augment judicial decision-making, with judges retaining ultimate responsibility for AI-assisted decisions.[219]
4. **Adaptability and Collaboration:** The framework should allow for adaptability in response to technological advancements, engaging stakeholders from judiciary, academia, and civil society.[220]

[217] Canadian Judicial Council, *Guidelines for the Use of Artificial Intelligence in Canadian Courts*, 2024; Vicki Birchfield, 'From Roadmap to Regulation: Will There Be a Transatlantic Approach to Governing Artificial Intelligence?' (2024) 46(7) *Journal of European Integration* 1053.
[218] *Resolution no. 332/2020* (n 108); European Commission for the Efficiency of Justice, *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems*, 2018.
[219] Guidelines for the Use of Artificial Intelligence in Canadian Courts (n 127) 3; Neerav Srivastava, 'Liability for Chatbots: A Psychbot Negligence Case Study and the Need for Reasonable Human Oversight' (2023) 28 *Torts Law Journal* 155 ('Liability for Chatbots').
[220] Artificial Intelligence Co-Regulation (n183) 14.

III Lessons from International Approaches

A *Brazil*

Brazil's *Resolution 332/2020* emphasises bias reduction through pre-approval requirements and diverse development teams. A centralised platform, *Sinapses*, coordinates judiciary-specific AI tools, streamlining oversight and reducing redundancy.[221]

B *Canada*

Canada's *Guidelines for the Use of Artificial Intelligence in Canadian Courts* stress judicial independence, transparency, and ethical use. They recommend pilot programs, continuous evaluation, and explainability standards to maintain public trust and fairness.[222]

C *United States of America*

The US takes a sector-specific, risk-based approach, emphasising transparency, fairness, and human accountability. The 2023 AI Executive Order highlights the need for explainable AI systems and limits over-reliance on "black-box" models[223]. This model ensures higher scrutiny for applications affecting rights, such as judicial AI, while promoting innovation in low-risk contexts.[224]

D *European Union*

The EU's *Artificial Intelligence Act* employs a risk-tiered model, requiring strict scrutiny for high-risk applications, such as AI in judicial decision-making. It mandates pre-implementation assessments, transparency obligations, and monitoring mechanisms.[225] Co-regulation ensures flexibility while involving multiple stakeholders, though concerns about transparency in standard-setting remain.[226]

E *United Kingdom*

The UK employs a principles-based, sector-specific model, balancing flexibility and accountability. Assurance frameworks and regulatory sandboxes enable real-world testing of AI tools under controlled conditions, ensuring both compliance and innovation.[227]

---

[221] *Resolution no. 332/2020* (n 108)
[222] Guidelines for the Use of Artificial Intelligence in Canadian Courts (n 127) 3-6.
[223] From Roadmap to Regulation (n 175) 1058.
[224] Artificial Intelligence as a Medical Device in Radiology (n 176) 749–751.
[225] *Regulation (EU) 2024/1689 (Artificial Intelligence Act)* [2024] OJ L 2024/1689, Ch I Art 3(63).
[226] Artificial Intelligence Co-Regulation (n 183) 14–15.
[227] Pro-innovation approach (n 197) 21, 64.

IV Regulatory Tools for the Victorian Context

1. **Risk Management Frameworks:**
    - Adopt a risk-tiered model, like the EU AI Act, categorising AI applications based on their potential impact on rights and public trust.[228]
    - Require pre-implementation assessments, impact evaluations, and independent audits for high-risk applications, such as predictive tools in judicial contexts.[229]
2. **Tailored Guidelines and Standards:**
    - Develop explainability and disclosure requirements for judicial AI systems, ensuring transparency and procedural fairness.[230]
    - Administrative tools can have less stringent standards, focusing on operational efficiency.[231]
3. **Centralised Oversight Platform:**
    - Establish a centralised platform for judiciary-specific AI tools, modeled on Brazil's *Sinapses*, to standardise development, ensure harmonised oversight, and track AI applications.[232]
4. **Stakeholder Engagement and Education:**
    - Involve judiciary, academia, and civil society in co-regulation processes to ensure balanced oversight.[233]
    - Provide comprehensive training for judges and court staff on AI risks, limitations, and appropriate use.[234]

V Technological Neutrality vs. Specific Regulation

Victoria's regulatory response should employ a **hybrid approach**:

1. **Technological Neutrality**: General principles such as fairness, transparency, and accountability should apply universally to all AI systems, allowing flexibility for future advancements.

---

[228] *Regulation (EU) 2024/1689 (Artificial Intelligence Act)* [2024] OJ L 2024/1689, Ch I Art 3(63).
[229] From Roadmap to Regulation (n 175) 1058.
[230] Guidelines for the Use of Artificial Intelligence in Canadian Courts (n 127) 4.
[231] The Use of Artificial Intelligence in the Judiciary (n 2) 157.
[232] *Resolution no. 332/2020* (n 108)
[233] Artificial Intelligence Co-Regulation (n 183) 14.
[234] Guidelines for the Use of Artificial Intelligence in Canadian Courts (n 127) 6.

2. **Specific Regulation:** High-risk applications, such as predictive decision-making tools, require detailed and enforceable standards to address critical risks like bias and explainability.[235]

VII Best Responses: Closing Thoughts

The best regulatory response to AI in Victorian courts combines global best practices with local needs, fostering innovation while safeguarding public trust and judicial independence. A risk-tiered, co-regulatory framework with specific oversight for high-risk applications ensures AI enhances judicial processes without compromising fairness or accountability. By implementing adaptable, transparent, and inclusive regulations, Victoria can set a global benchmark for ethical AI governance in the judiciary.

Recommendations

8. Adopt a Principles-Based Regulatory Framework with a focus on overarching principles such as fairness, transparency, accountability, and human oversight to ensure adaptability as AI technology evolves. This approach allows for flexibility and adjustments based on actual data, rather than perceived harms, enabling regulations to remain effective in a dynamic AI landscape.
9. Consider the potential need for overarching legislation in the future. Such legislation can provide consistency across sectors and address emerging risks effectively, ensuring a comprehensive and unified regulatory framework for AI technologies.
10. Apply stricter regulations in the case of criminal matters, considering that they are more sensitive. Ensure that a judge presides over the case and that predicative AI is not used to for decision making unless it is being used for calculating sentencing.
11. Review ICT projects implemented in Australian courts and tribunals over the past 20 years as these may be value for identifying key factors to improve likelihood of successful AI projects.
12. Establish robust security protocols to mitigate risks, such as prohibiting the use of sensitive court data for AI training.
13. Establish working groups, such as Quality of Justice or Cyberjustice teams, to facilitate education, collect and report performance data, and share experiences across jurisdictions to build confidence and promote informed AI adoption.

[235] *Resolution no. 332/2020* (n 108) ; Artificial Intelligence as a Medical Device in Radiology (n 176) 749-751.

14. Develop an AI assurance framework that includes mandatory impact assessments, audits, and reporting, supported by regulatory sandboxes and pilot programs to test AI tools in controlled environments before full implementation. This framework will also allow ongoing monitoring and review of processes to reassess risks as AI technology evolves.
15. Evaluate AI systems and algorithms to ensure courts operate in compliance with copyright and other legal obligations, avoiding complicity in potential infringements by AI tools.
16. Implement Robust Bias Mitigation and Compliance Measures by establishing pre-approval processes for AI tools to evaluate biases, discontinuing tools that fail to address discrimination.
17. Stay informed about AI regulatory developments in international jurisdictions like the UK, EU, and US to anticipate concerns, adopt best practices, and ensure alignment with global standards.

Chapter 6: Principles for Responsible and Fair Use of AI in Courts and Tribunals

Question 12: Are principles sufficient, or are guidelines or other regulatory responses also required?

In evaluating international approaches, there is a clear method to implementation of reform. Principles themselves do not appear to be the final step in ensuring reform. The international jurisdictions that have, or intend to, legislate AI have used principles as a building block with the intention of growth from that position. Referencing China specifically, the CCP developed and implemented a range of reforms that targeted specific technological issues regarding AI. Once these were implemented, further principles were regulated to both address any potential failures and plug holes that had grown due to manifestation of the technology. From this position, which took close to seven years, China is now formulating and drafting a nation-wide legislation that will serve as an umbrella-model over AI technology in that jurisdiction. Using this as an example, in would be beneficial for Victoria to develop reform targeting specific issues which in turn would lead to a broader legislative requirement in due course.

Question 13: What regulatory tools, including guidelines, could be used to implement these high-level principles in Victoria's courts and tribunals?

To reach this stage, Victoria will need to develop a series of tools that are multi-jurisdictional and contain the relevant framework to operate in conjunction with the AI field. Using China as a specific example again, the development of an Algorithm Registry has served multiple purposes. On the surface, the tool has ensured compliance with regulation from relevant stakeholders, with some companies forced to complete over five separate filings for the same app.[236] As usage grows, the pool of information within the registry grows with it, simultaneously serving to develop the knowledge base of relevant bureaucrats. Secondly, the tool functions as a standardised disclosure tool which can be refined as required for future regulatory implementation, easing the construction of regulation.[237]

[236] Sheehan (n 153)15.
[237] Ibid.

Question 14: How can the use of AI by courts and tribunals be regulated without interfering with courts' independence, and what risks should be considered?

The separation of powers is a foundational principle of democratic governance, and any encroachment on the judiciary's autonomy by the executive or legislative branches poses a significant threat to the constitutional balance.[238] If the executive branch controls AI systems used in courts, it risks undermining the judiciary's ability to function independently and impartially. Regulatory frameworks must ensure that the judiciary retains full autonomy over the development, implementation, and oversight of AI technologies, thereby preserving the constitutional separation of powers and judicial independence.[239] Therefore, the use of open-source software is a practical solution to ensure transparency and public accountability, as it allows external scrutiny and fosters public trust in the judiciary.[240] These measures ensure that the judiciary retains control over how the technology is developed and applied, thus preserving its independence. Another way is through is an expert committee and creating a team within the Courts and Tribunals dedicated to creating and testing AI. For example, in Brazil, the National Council of Justice (CNJ) has played a leading role in developing and implementing AI solutions within the judiciary. They have established internal teams to oversee the development and deployment of these technologies, ensuring judicial control and expertise are at the forefront. The involvement of private entities or the executive branch in developing or managing AI systems, potentially exposes the judiciary to undue influence, leading to politically motivated outcomes that could compromise the impartiality of judicial decisions.[241]

AI should serve as a decision-support tool rather than a decision-maker, enabling judges to critically evaluate the system's recommendations and apply their own legal expertise to each case.[242] Ex-ante regulation—regulatory interventions designed to prevent harm before it occurs—may be particularly relevant for the judiciary.[243] For instance, conducting risk assessments at

---

[238] Tania Sourdin, 'Replacing, Supporting or Enhancing Judges? Judge AI Considerations for the Future' (2024) 98 *Australian Law Journal* 696, 704-5.
[239] Ibid 705; Paweł Marcin Nowotko, 'AI in judicial application of law and the right to a court' (2021) 192 *Procedia Computer Science* 2220, 2224; Viktor Alan Brekke, 'Artificial Intelligence in the judicial system: Maintaining the independency of the judiciary power in the development, implementation, and use of artificial intelligence' (Master thesis, University of Oslo, 2022) 31 ; Monika Zalnieriute and Felicity Bell, 'Technology and Judicial Role' in Gabrielle Appleby and Andrew Lynch (eds), *The Judiciary and the Court: Individual, Collegial and Institutional Judicial Dynamics in Australia* (Cambridge University Press, 2021) 1, 13 (*'The Judiciary and the Court'*).
[240] Nowotko (n 240) 2224; Technology and the Courts submission (n 82); *The Judiciary and the Court* (n 240) 24.
[241] Fundacja Moje Państwo, 'Algorithm of the System of Random Allocation of Cases finally disclosed! (Web Page, 22 September 2021) <https://mojepanstwo.pl/aktualnosci/773>.
[242] *The Judiciary and the Court* (n 240) 35.
[243] Safe and responsible AI in Australia: Proposal Paper (n 193) 13-4.

appropriate intervals and implementing robust oversight mechanisms to monitor the performance and impact of AI systems in legal contexts.[244]

In particular, training programs designed to enhance judges' understanding of AI's capabilities and skills in prompt engineering will help to identify errors and prevent overreliance on the technology. For judicial officers generally training should be developed per their specific roles and responsibilities, focusing on how AI tools might impact their tasks, such as case management, document review, and legal research. This could include modules on data privacy, bias, and the appropriate use of AI-generated summaries. For barristers and lawyers, training programs should be implemented to ensure they understand the ethical implications of using AI in litigation, including disclosure requirements, data security, and potential biases in AI tools. These programs should also cover best practices for using AI in legal research, document drafting, and case preparation, while emphasising the importance of human oversight and critical thinking. As previously mentioned, general guidelines implementing ethical standards when using AI should be published to provide clear guidance for all legal professionals, including judges, judicial officers, barristers, and lawyers. These guidelines should address issues such as transparency, accountability, fairness, and data privacy, and should be regularly updated to reflect advancements in AI technology and its application in the legal field.

Further, as discussed below, disclosure should become a requirement not only for judges but also other judicial officers and legal professionals. Judicial officers like court clerks, registrars, and judicial assistants assist judges by managing administrative tasks, conducting legal research, preparing case summaries, and drafting orders. Judges are also assisted by submissions by the representatives of both sides. Although they might conduct some of their own research, they are most reliant on these submissions, as they present the legal arguments and evidence necessary for the judge to make an informed decision within the adversarial system. Submissions generated with AI can then infect the rest of the judicial process if not disclosed because they may contain biased or inaccurate information, which could influence the judge's understanding of the case and lead to an unfair outcome. By applying these preventative measures early, courts and tribunals can mitigate risks associated with high-stakes decisions, ensuring that AI technologies enhance rather than undermine justice.

[244] Ibid.

Transparency is a fundamental principle in maintaining judicial independence and public trust. Litigants must be informed when AI is involved, and the technology must provide sufficient explanations of how it arrived at specific conclusions or recommendations.[245] Without transparency, AI's potential biases and limitations remain hidden, reducing the accountability of judicial decisions.[246] Intellectual property laws that protect proprietary AI algorithms often pose a challenge to transparency, as they prevent public access to the underlying mechanisms of the technology.[247] If the judge does not understand why AI is making a particular recommendation, they cannot meaningfully assess its validity in the context of the specific case. Consequently, courts must implement regulations that prioritise transparency over proprietary interests to maintain the integrity of the judicial process.[248] The use of AI systems with opaque algorithms can create public suspicion and undermine trust in judicial outcomes.[249] Regulatory frameworks should consider the broader implications of AI on public confidence in the judiciary, as any perceived bias or lack of transparency can erode trust in the legal system.

The risks associated with AI in the judiciary are not limited to questions of independence and impartiality. Automation bias is a significant risk and occurs when judges place excessive trust in AI-generated recommendations, potentially leading to a reduction in independent scrutiny and critical judgment.[250] Automation bias is particularly problematic in the judiciary, where each case must be evaluated based on its specific facts and context, rather than being treated as a mere data point within a larger pattern.[251] As research indicates, decision-makers are prone to defer to automated systems,[252] reducing independent scrutiny,[253] and struggle to effectively judge the quality of algorithmic outputs or determine when to override them.[254] It has also shown that when

[245] Safe and responsible AI in Australia: Proposal Paper (n 193) 17.
[246] Ibid 17, 35.
[247] *The Judiciary and the Court* (n 240) 19-20.
[248] Safe and responsible AI in Australia: Proposal Paper (n 193) 24, 35.
[249] Brian Barry, 'AI for Assisting Judicial Decision-making: Implications for the Future of Open Justice' (2024) 98 *Australian Law Journal* 656, 661-2.
[250] Giovana Lopes, 'AI and Judicial decision-making; Evaluating the role of AI in debiasing' (2024) *Journal for Technology Assessment in Theory and Practice* 28, 31; *The Judiciary and the Cour*t (n 240) 7.
[251] See Justice Melissa Perry and Sonya Campbell, 'AI and Automated Decision-Making: Are you just another number?' (Speech, Gilbert + Tobin Centre of Public Law, 21 October 2021); Brekke (n 240) 32.
[252] Kelly Hannah-Moffat, 'The Uncertainties of Risk Assessment: Partiality, Transparency and Just Decisions' (2015) 27(4) *Federal Sentencing Reporter* 244; Mary Dzindolet et al, 'The perceived utility of human and automated aids in a visual detection task' (2002) 44 *Human Factors: The Journal of the Human Factors and Ergonomics Society* 79, 88-91.
[253] Raja Parasuraman and Dietrich Manzey, 'Complacency and bias in human use of automation' (2010) 52(3) *Human Factors: The Journal of the Human Factors and Ergonomics Society* 381, 398-9.
[254] Ben Green, 'The flaws of policies requiring human oversight of government algorithms' (2022) 45 *Computer Law & Security Review* 1, 7-8.

presented with algorithmic risk assessments, decision-makers place undue emphasis on their results, overshadowing other relevant factors.[255]

In *Loomis*,[256] Eric Loomis challenged the use of COMPAS, a risk assessment tool, during his sentencing, arguing that its proprietary nature prevented him from understanding how it arrived at its risk assessment, thus violating his due process rights. Despite acknowledging the limitations of COMPAS, including its lack of transparency and potential biases,[257] the Wisconsin Supreme Court upheld the lower court's sentencing decision. This case highlights the intertwined problems of automation bias and hindsight bias in judicial decision-making. Automation bias is evident in the court's reliance on COMPAS despite its acknowledged limitations and even with 'warning labels,' judges may still be subconsciously influenced by the AI's recommendations.[258] This is compounded by hindsight bias, where it becomes easy to rationalise a decision after the fact, claiming the outcome would have been the same regardless of the AI's input. In *Loomis*, the court accepted the lower court's assertion that the sentence would have been the same,[259] demonstrating this retrospect justification. This interplay of automation bias (over-reliance on the tool) and hindsight bias (rationalising the decision after the fact) makes it difficult to ascertain the true impact of AI on judicial decisions and effectively address the potential for undue influence.

To address these concerns, courts must implement clear disclosure requirements that inform the public about the use of AI in judicial processes. The integration of AI into the judiciary also raises ethical considerations related to the role of human judgment in the legal process. While AI can enhance efficiency and consistency, it cannot replicate the human capacity for empathy, moral reasoning, and contextual understanding.[260] These qualities are essential for delivering justice in a manner that respects the individuality of each litigant and their case. To mitigate this risk, regulatory frameworks must emphasise that AI systems are intended to supplement human judgment, not replace it. Judges must remain the final arbiters in all cases, using AI as a tool to enhance their decision-making rather than the 'longarm of the algorithm.'[261]

---

[255] Ben Green and Yiling Chen, 'Algorithmic risk assessments can alter human decision-making processes in high-stakes government contexts' (2021) 5 *Proceedings of the ACM on Human-Computer Interaction* 1, 13-8.
[256] *Loomis* (n 81).
[257] Ibid 770, 774–5.
[258] Han-Wei Liu, Ching-Fu Lin and Yu-Jie Chen, 'Beyond State v Loomis: artificial intelligence, government algorithmization and accountability' (2019) 27(2) *International Journal of Law and Information Technology* 122, 130.
[259] *Loomis* (n 81) 769.
[260] Sangeet Sharma et al, '*A Study of AI-Based Systems in the Judicial Interpretation of the Law'* (Conference Paper, India, 2023) 53; Sourdin, (n 239) 700.
[261] Kyriakos Kotsoglou and Marion Oswald, 'The Long Arm of the Algorithm? Automated Facial Recognition as Evidence and Trigger for Police Intervention' (2020) 2 *Forensic Science International: Synergy* 86.

Recommendations

18. The judiciary must retain full autonomy over AI technologies by establishing judicial oversight committees, utilising in-house expertise (if feasible) such as a designated AI team and a preference for open-source software to ensure transparency, accountability and independence from external influence.
19. Create clear, practical guidelines for judges and tribunal members on AI usage, covering topics like data privacy, algorithmic bias, and human oversight. Mandate AI training for all court and tribunal judicial employees to ensure understanding of AI's role, risks, and limitations.
20. Mandate regular audits of all AI systems used in courts and tribunals to ensure alignment with current legal practices, identify and mitigate algorithmic biases, and address potential errors or prejudicial outcomes. Require independent third-party audits for high-risk applications to ensure compliance with ethical and technical standards.

Question 15: Is it appropriate to have varying levels of transparency and disclosure depending on the use of AI by courts and tribunals? (For example, use by administrative staff compared with judicial officers.)

The implementation of artificial intelligence (AI) in Victorian courts and tribunals necessitates a nuanced approach to transparency and disclosure. Distinguishing between AI applications used by administrative staff and those employed by judicial officers is essential to balance operational efficiency, public trust, and accountability.

I Administrative Staff

AI applications used by administrative staff—such as tools for scheduling, document processing, or resource allocation—carry relatively low risks of influencing case outcomes. As such, minimal disclosure requirements may be appropriate, provided these tools are used solely for operational purposes. AI operational use by administrative staff differs from judicial officers, demanding tailored disclosure based on risk. This approach aligns with the principles of proportionality, ensuring that resources are not unnecessarily diverted toward explaining low-risk systems.

However, baseline transparency remains essential, even for administrative tools, to ensure ethical use and avoid potential misuse. Stakeholders, including court users and legal

practitioners, should be informed about the existence of these AI systems and their functions. For example, AI systems that automate scheduling should disclose their criteria for prioritising cases to ensure fairness in resource distribution.[262] Courts should balance transparency with operational efficiency, ensuring that stakeholders are informed about the general functions of AI systems without revealing unnecessary technical details. For instance, layered disclosure mechanisms could be employed, where stakeholders receive high-level information about AI systems, while technical specifics remain available for internal or regulatory audits.[263]

Additionally, there should be internal mechanisms for monitoring these tools to identify errors or biases that could indirectly affect judicial processes. While these tools do not make substantive decisions, any failure in their operation could impact court efficiency or lead to unequal access to services, underscoring the need for basic oversight.[264] Monitoring mechanisms must also account for the risk of 'cascading errors', where operational mistakes could indirectly influence case outcomes or procedural fairness.[265] Additionally, Chaudhary notes that disclosure mechanisms must ensure ethical safeguards without overwhelming users or disrupting operations.[266]

By maintaining baseline transparency, stakeholders can remain informed about the existence of AI systems and their general functions without requiring technical details that may overwhelm users.[267] As such, courts can promote public trust and accountability without burdening administrative processes with disproportionate disclosure requirements.

II Judicial Officers

In contrast to administrative tools, AI systems assisting judicial officers—such as tools for predicting case outcomes, assessing risk, or providing sentencing recommendations—require robust transparency and accountability measures. These high-risk applications directly influence legal decision-making and have significant implications for the principles of fairness, equality, and public trust.

---

[262] The Use of Artificial Intelligence in the Judiciary (n 2) 157.
[263] Fabiana Di Porto, 'Algorithmic Disclosure Rules' (2021) 31(1) *European Journal of Risk Regulation* 13, 14 ('Algorithmic Disclosure Rules').
[264] Ashley Deeks, 'The Judicial Demand for Explainable Artificial Intelligence' (2019) 119(7) *Columbia Law Review* 1829 ('The Judicial Demand for Explainable Artificial Intelligence').
[265] The Use of Artificial Intelligence in the Judiciary (n 2) 157.
[266] Gyandeep Chaudhary, 'Unveiling the Black-box: Bringing Algorithmic Transparency to AI' (2024) 18(1) *Masaryk University Journal of Law and Technology* 93, 96 ('Unveiling the Black-box').
[267] The Use of Artificial Intelligence in the Judiciary (n 2) 157.

A *The Necessity of Robust Transparency*

AI systems employed by judicial officers must operate with full transparency, particularly due to their potential to influence rights and liberties. Unlike operational AI, these systems actively shape legal outcomes, meaning their use must be thoroughly understood and monitored. Differentiation in transparency levels ensures AI's utility without compromising judicial accountability.[268] This differentiation acknowledges the heightened responsibilities of judicial officers in applying AI compared to administrative staff.

Transparency serves multiple purposes:

1. **Safeguarding Public Confidence**: Public trust in judicial systems depends on their perceived fairness and impartiality. A lack of transparency in AI systems risks eroding confidence, particularly if decisions are perceived as being made by 'black-box' algorithms.[269]
2. **Promoting Contestability**: Courts must ensure that parties affected by AI-assisted decisions can understand the basis of those decisions and, if necessary, contest them. This aligns with the principles of procedural fairness, which underpin the rule of law.[270]
3. **Enabling Accountability**: The inability to scrutinise AI decisions undermines judicial accountability. As Chaudhary explains, the inability to scrutinise AI decisions erodes the foundations of judicial accountability and fairness.[271]

B *Specific Requirements for Full Disclosure*

Full disclosure should address several critical aspects of AI systems used by judicial officers:

1. Underlying Algorithms and Data Sets:
    - The algorithms employed must be explainable, providing insights into how decisions are made, transparency in algorithmic logic is crucial for ensuring accountability and trust.[272]
    - Data sets used to train these algorithms must be scrutinised to prevent biases that could result in systemic discrimination or inequity.[273]
2. Explanation of Decision-Making Processes:

---

[268] Neerav Srivastava, 'Liability for Chatbots: A Psychbot Negligence Case Study and the Need for Reasonable Human Oversight' (2023) 28 *Torts Law Journal* 155.
[269] The Use of Artificial Intelligence in the Judiciary (n 2) 157.
[270] Unveiling the Black-box (n 267) 95; The Judicial Demand for Explainable Artificial Intelligence (n 266) 1829.
[271] Unveiling the Black-box: Bringing Algorithmic Transparency to AI (n 267) 95.
[272] Ibid.
[273]The Judicial Demand for Explainable Artificial Intelligence (n 265) 1829.

- Judicial officers and parties to proceedings must be able to understand how AI tools reach their conclusions. This requires the implementation of tools that make algorithmic logic interpretable to non-technical stakeholders. As Deeks notes, explainability is essential for preserving the integrity of judicial processes.[274]

3. Auditable Decision Trails:
    - Courts must employ auditable decision trails to clarify how AI systems influence judicial decisions. This ensures traceability and allows for meaningful oversight, creating transparency in processes that involve AI.[275]
4. Mechanisms for Contestability:
    - AI-assisted decisions must remain subject to judicial oversight and human intervention. This includes clear pathways for challenging decisions where AI has been used. As Chaudhary argues, the ability to contest algorithmic outcomes is foundational to safeguarding human rights in judicial processes.[276]

C Risks of a Lack of Transparency

Failure to implement robust transparency measures risks undermining judicial integrity and public confidence. The reliance on opaque, 'black-box' algorithms poses several threats:

- **Bias and Discrimination:** Without proper scrutiny, AI tools may reinforce existing biases in legal systems, disproportionately impacting vulnerable groups.[277]
- **De-Skilling of Judicial Officers:** Over-reliance on AI could lead to the erosion of critical judicial skills, reducing the capacity for independent human decision-making over time.[278]
- **Public Distrust:** Non-transparent systems may undermine the perception of fairness, especially when parties cannot fully understand or contest decisions influenced by AI.[279]

IV Tailored Transparency Framework

A tiered approach to transparency is recommended, where disclosure obligations align with the risk and impact of the AI application:

- **High-Risk Systems:** Full transparency, including algorithmic explainability, traceability, and accountability measures.

---

[274] Ibid.
[275] Algorithmic Disclosure Rules (n 264) 16.
[276] Unveiling the Black-box (n 267) 96.
[277] The Use of Artificial Intelligence in the Judiciary (n 193) 157.
[278] Algorithmic Justice Symposium (n 180).
[279] Unveiling the Black-box: (n 267) 95.

- **Low-Risk Systems:** Basic disclosures to inform stakeholders of AI use, focusing on functionality and ethical safeguards.

This risk-based framework aligns with calls for 'algorithmic accountability' and ensures that courts maintain their commitment to fairness while leveraging AI for efficiency.[280]

V Closing Thoughts: Varying Levels of Disclosure

Adopting a differentiated approach to transparency and disclosure will enable Victorian courts and tribunals to harness the benefits of AI without undermining public trust or judicial integrity. By aligning disclosure requirements with the risk profile of each application, courts can achieve a balanced framework that supports both operational efficiency and accountability

[280] The Use of Artificial Intelligence in the Judiciary (n 193) 157.

Chapter 8: Developing Guidelines for the use of AI in Victoria's Courts and Tribunals

Guidelines for Courts and Tribunals

Question 29: What are the benefits and risks of disclosure? If mandatory, what form should disclosure take?

The use of AI in Victorian courts and tribunals introduces both opportunities and challenges in transparency and disclosure. Mandatory disclosure of AI systems fosters trust, accountability, and oversight. However, it also raises concerns about proprietary protection, operational complexity, and potential misuse of disclosed information. A balanced approach to disclosure is essential to ensure AI is implemented ethically and effectively without undermining judicial processes or public confidence. Here we explore the benefits and risks of disclosure and recommend strategies to address these concerns.

A *Benefits of Disclosure*

1 *Transparency Fosters Trust and Accountability*

Transparency is vital for maintaining public confidence and ensuring the integrity of AI systems in judicial settings. By providing clarity on how AI systems operate, courts can align their use with ethical and legal principles. Chaudhary highlights that transparency helps establish public trust and confidence in AI-driven systems, ensuring their alignment with ethical standards and legal principles.[281] This is particularly crucial in high-stakes applications, such as risk assessment or sentencing decisions, where fairness must be evident and indisputable.

2 *Facilitates Oversight*

Mandatory disclosure empowers regulators to monitor and assess AI systems effectively, minimising risks of systemic bias, discrimination, or error. Terzidou emphasises that regular reviews and audits of disclosed information can mitigate risks like data misuse or algorithmic discrimination.[282] Additionally, Mandatory disclosure of AI use in judicial contexts ensures systemic fairness and public confidence.[283]

---

[281] Unveiling the Black-box, (n 267) 95.
[282] The Use of Artificial Intelligence in the Judiciary (n 193) 157.
[283] Liability for Chatbots (n 220).

3 *Enhances Public Confidence*

Clear communication about AI's role in judicial processes builds public trust, particularly when the systems are used in high-risk applications. As noted at the Algorithmic Justice Symposium, transparency reassures stakeholders of the ethical and responsible use of AI systems.[284] Deeks adds that algorithmic transparency enhances accountability and fairness, enabling stakeholders to scrutinise the logic and data behind AI decisions.[285]

B *Risks of Disclosure*

1 *Over-Disclosure Risks*

Excessive disclosure may expose proprietary systems or sensitive algorithmic designs, creating vulnerabilities for misuse or adversarial attacks. As Chaudhary notes, the challenge lies in ensuring algorithmic transparency without compromising the proprietary nature of AI models or creating undue barriers for stakeholders.[286] Di Porto similarly warns of the risks posed by over-disclosure, stating that adversarial attacks on disclosed algorithms could compromise their functionality or security.[287]

2 *Operational Challenges*

Balancing the need for transparency with operational confidentiality is complex. Courts face logistical and financial challenges in maintaining detailed disclosures for frequently updated AI systems.[288] Furthermore, disclosing too much may hinder innovation or violate contractual obligations with AI vendors.[289]

3 *Potential for Misinterpretation*

Non-technical stakeholders may misinterpret the disclosed information, leading to scepticism or unwarranted distrust of AI systems. Chaudhary cautions that disclosing technical details of AI systems may overwhelm stakeholders with unnecessary complexity, creating barriers to comprehension and accessibility.[290]

---

[284] Algorithmic Justice Symposium (n 180).
[285] The Judicial Demand for Explainable Artificial Intelligence' (n 265).
[286] Unveiling the Black-box (n 267) 94.
[287] Algorithmic Disclosure Rules (n 264) 14.
[288] The Judicial Demand for Explainable Artificial Intelligence (n 265).
[289] Algorithmic Disclosure Rules (n 264) 13.
[290] Unveiling the Black-box (n 267) 97.

Recommendations

21. Implement a tailored, layered approach to AI disclosures. For low-risk applications, provide public summaries that are transparent and accessible without overwhelming stakeholders. For high-risk systems, require detailed technical disclosures—including algorithmic design, training data, and oversight mechanisms—accessible to regulators and judicial authorities. This ensures accountability and transparency while maintaining operational confidentiality for proprietary systems.
22. Require the disclosure of the specific roles AI make in decision-making, provide high-level algorithm descriptions and their limitations, and outline mechanisms for challenging AI-influenced decisions to uphold procedural fairness. Introduce measures like watermarks or labels to identify AI interactions, balancing transparency with appropriate detail for public understanding and the specific risks associated with AI use.

Mandatory disclosure of AI use in Victorian courts and tribunals is essential to foster public trust, ensure accountability, and mitigate systemic risks. However, over-disclosure poses significant challenges, including operational burdens, proprietary risks, and stakeholder misinterpretation. A nuanced approach, combining a layered framework with clear guidelines, will ensure that disclosure remains a tool for transparency without undermining the effective and ethical use of AI in judicial contexts.

Question 30: Should courts and tribunals undertake consultation with the public or affected groups before using AI and/or disclose to court users when and how they use AI? What other mechanisms could courts and tribunals use to promote the accountable and transparent use of AI?

I Introduction

To allow responsible AI to flourish, it is essential that public sentiment is considered, particularly as the vast majority of the Australian public believe that widespread AI deployment within government institutions must be strongly regulated.[291] A recent survey in 2020 reported that only

[291] KPMG Australia and The University of Queensland, Trust in Artificial Intelligence: Australian Insights (Report, October 2020) 2.

one in three respondents felt the use of AI by public services could be considered trustworthy.[292] Following the 'Robodebt' saga, it is reasonable to assume that the public will remain apprehensive of another public service utilizing AI in its processes. This is of course problematic as the public will likely be the most affected stakeholder while AI use continues to grow.[293] In a post 'Robodebt' scheme era,[294] it is essential that the public regains trust in the use of automated software, particularly where outcomes will directly affect the legal rights of individuals. This must be achieved through purposefully prioritizing transparency and accountability when deploying AI by courts and tribunals.[295]

To counter this evident lack of trust in AI implementation, Law Council Australia has suggested that a public education strategy regarding use of AI should be included when preparing AI regulations in the legal sector.[296] Specifically, public forums between key stakeholders and teaching workshops to engage with the public may be beneficial.[297] In the same vein, it may be advisable that judicial bodies engage with the public in an educational context to quell the current wariness over AI implementation by government institutions. The apparent apprehension and distrust that the public feel towards AI must be overcome as it has detrimental consequences on public confidence in judicial institutions.

Drawing parallels with the Robodebt inquiry, a significant challenge pertaining to AI use in the legal sector stems from the lack of transparency regarding how the data is used and decisions are made. During the Robodebt scheme, the public could not access the reasoning behind the decisions made by the software nor could they contest those decision.[298] This issue will be particularly troubling if carried over to the deployment of AI by courts and tribunals. This past foray into partially or complete ADM use has detrimentally affected the public's trust in the use of AI technology.[299] However, it is important to strike a reasonable balance between transparency and avoiding overwhelming the public with an unnecessary level of technical detail.[300]The degree of disclosure regarding the explainability of AI use in various court processes must be done in a

---

[292] Ibid.
[293] Gabriel Lima, Nina Grgić-Hlača and Meeyoung Cha, 'Human Perceptions on Moral Responsibility of AI: A Case Study in AI-Assisted Bail Decision-Making', (2021) *CHI Conference on Human Factors in Computing Systems* 1,1.
[294] Royal Commission into the Robodebt Scheme (Final Report, 7 July 2023).
[295] Safe and Responsible AI in Australia (n 76) [151].
[296] Law Council Australia, Submission No 2 to Department of Indusrty, Science and Resources, *Introducing Mandatory Guardrails for AI in High-risk Settings: Proposals Paper* (9 October 2024).
[297] Safe and Responsible AI in Australia (n 76) 37 [153].
[298] Victoria Legal Aid, Submission No 1 to the Australian Human Rights Commission, *Human Rights and Technology Discussion Paper* (10 March 2020) 2,1-5
[299]Safe and Responsible AI in Australia (2023) (n 76) 5 [6].
[300] Unveiling the Black-box (n 267) 95.

way that is palatable for the public. The extent of disclosure must also be proportionate to the level of risk of harm associated with the specific AI use in specific tasks. For example, if ADM is used, there must be a conscious effort to ensure transparency of outcomes. This can ensure that individuals are able to understand the rationale behind decisions made and therefore, maintain the contestability of decisions where appropriate.[301]

Apart from disclosure to users, there are other means that can be used to safeguard transparency when courts and tribunals utilise AI technology[302]:

- **User Awareness:** Ensuring users know when they are interacting with AI or viewing AI-generated content. This could be done visually e.g. watermarks.[303]
    - The identification of AI use cannot be removed from the AI-generated content.[304]
- **Reporting AI System Details:** Publicly sharing information about the system's limitations, capabilities, and insight into how to use AI tools appropriately.
- **Sharing Model Data:** Providing insight into the data used to train AI systems.
- **Cross-disciplinary dialogue between all key stakeholders:** Legal practitioners, AI developers, court staff and other key stakeholders must be able to discuss the areas across various disciplines.

Practises for upholding accountability to ensuring responsible and safe AI could include:

- **Designated Roles for AI Safety:**[305] Appointing specific individuals or teams, within specific courts or tribunals, with clear responsibility for ensuring AI systems are safe and comply with ethical standards. This ensures that there is human oversight for all AI assisted processes in the courts.
- **Establishing a centralised regulatory body:**[306] To provide continuous oversight of the development and use of AI technologies as well as regular evaluations of AI systems. Courts

---

[301] Law Council Australia, Submission to Select Committee on Adopting Artificial Intelligence, *Inquiry into the opportunities and impacts of the uptake of artificial intelligence technologies in Australia (*20 May 2024) 29, 1-58 [67] ('Submission to the Select Committee on Adopting Artificial Intelligence').
[302] Department of Industry, Science and Resources, Australian Government, *Safe and Responsible AI in Australia Consultation: Australian Government's Interim Response* (Interim Response, 17 January 2024).
[303] Department of Industry, Science and Resources, Australian Government, 'Australia's Artificial Intelligence Ethics Principles', *Australia's AI Ethics Principles* (Publication 7 November 2019) <https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles>.
[304] Ibid.
[305] Ombudsman, *Automated Decision-Making: Report* (Report, Commonwealth of Australia, 2025) 33, 1-36 <https://www.ombudsman.gov.au/__data/assets/pdf_file/0029/288236/OMB1188-Automated-Decision-Making-Report_Final-A1898885.pdf>.
[306] Safe and Responsible AI in Australia (n 76) 27 [118] citing Dayal [2024] FedCFamC2F 1166.

and tribunals are currently using practice notes as a means of regulating AI use. However, this 'soft law' and voluntary approach to AI use can lead to greater uncertainty and opens its use to more risk as the technological landscape evolves.[307] A discrete regulatory body for AI use in the legal sector may lead to a streamlined, unified approach to its usage, specifically in the judicial context. The regulations must reflect the ethical guardrails proposed by the government, particularly where the use of AI is deemed high-risk.

In relation to practitioners, it is essential that they are made aware of what constitutes as ethical and responsible use of AI when engaging with the judicial system. Such regulations addressing responsible AI use should be set out by a central regulatory body. In Victoria, a legal practitioner was referred to the VLSBC for using AI tools for submissions which contained 'hallucinations' (that is, fictional legal citations).[308] The practitioner reasoned that he was unaware of the expectations surrounding AI use as guidelines have yet to be issued by the relevant court. This unfortunate incident illustrates that the current absence of clear and unified regulations across the judiciary can have potentially significant consequences for the legal sector and therefore lead to a diminished public trust in the Australian legal system.

- **Training Requirements:**[309] Mandating training for all AI deployers in the justice sector to ensure they understand safety protocols and best practices. A highly controversial use of AI is in its application within the judiciary (though, this practice has yet to reach Victoria.) Decision makers must be wary of the extent at which AI tools are supporting or assisting in decision making. An overreliance on AI can impact the principles of judicial independence and ultimately erode public trust.[310]

Recommendations

23. Balance transparency with considering the appropriate degree of detail for ordinary members of the public and the risk associated with the specific AI use.
24. Launch public education strategies, such as forums and workshops, to engage the public and key stakeholders in discussions about AI in the legal system. Encourage interdisciplinary dialogue between legal professionals, AI developers, and other

[307] Submission to Select Committee on Adopting Artificial Intelligence (n 302) 27 [34].
[308] *Handa v Mallick* [2024] FedCFamC2F 957.
[309] Department of Finance , Australian Government, *Implementing Australia's AI Ethics Principles in Government* (Online, 21 June 2024) <https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government/implementing-australias-ai-ethics-principles-government#6-transparency-and-explainability>.
[310] Technology and the Courts submission (n 82) 8.

stakeholders to ensure effective implementation and accountability throughout AI application stages.

25. Create a centralised regulatory authority, such as a Victorian AI Assessment and Review Committee, to oversee the development, deployment, and use of AI in the judiciary. This body should have the authority to prevent projects with high risks, ensure AI development aligns with the public interest, and uphold judicial independence.

Question 31: Should there be different guidelines or additional considerations for the use of AI in relation to criminal and civil law matters?

The current guidelines for the use of artificial intelligence (AI) in Victoria do not differentiate between criminal and civil law. Both the guidelines issued by the Supreme Court of Victoria and the County Court fail to specify whether they apply to criminal or civil matters.

In contrast, Queensland's guidelines for the responsible use of generative artificial intelligence by non-lawyers clearly state that they apply to "criminal and civil proceedings in Queensland courts and tribunals, including the Supreme Court, District Court, Planning and Environmental Court, Magistrates Courts, Land Court, Children's Court, Industrial Court, Queensland Industrial Relations Commission, and Queensland Civil and Administrative Tribunal."[311]

The lack of distinction between criminal and civil matters in Victoria's guidelines may warrant further consideration, particularly due to the differing standards of proof required in these areas and the potential impact on human rights. The distinct nature of these proceedings could have significantly different implications for individuals involved.

Criminal law is underpinned by essential human and legal rights principles, such as the presumption of innocence and the right to a fair and unbiased trial. Moreover, criminal sanctions can severely affect an individual's right to liberty. In contrast, civil law primarily deals with financial matters, contractual disputes, and the resolution of conflicts through compensation or injunctions. The outcomes of civil proceedings are generally less invasive of personal freedoms, focusing instead on monetary or specific performance remedies.

With the growing use of AI in courtrooms, particularly in sentencing, judicial determinations, predictive analysis, and policing, the need for distinct guidelines becomes more pressing. The

[311] Queensland Courts, *Artificial Intelligence Guidelines for Non-Lawyers* (Guideline 13 May 2024) 1 <https://www.courts.qld.gov.au/__data/assets/pdf_file/0012/798375/artificial-intelligence-guidelines-for-non-lawyers.pdf>.

use of AI in these high-risk areas could significantly affect human rights and legal outcomes. As discussed in Question 7, evidence from criminal justice systems internationally shows that AI tools may perpetuate racial, gender, or other biases, leading to inaccurate and unjust outcomes that violate fundamental principles of justice. Our recommendation for question 7 suggests that, at this stage, the use of AI tools in the legal system should either be prohibited or strictly regulated to ensure they operate within tightly controlled parameters with extensive human oversight.

While this submission acknowledges the potential for differing guidelines in the future as AI continues to evolve, the current use of AI in both criminal and civil law presents similar concerns. The present risks associated with AI do not vary significantly between these two areas, and therefore a unified approach may be more effective.

This submission recommends the establishment of a unified set of guidelines for the use of AI, applicable equally to both criminal and civil matters. These guidelines should focus primarily on procedural law to mitigate current risks and guide future implementation. The overarching goals should be fairness, accuracy, accountability, and efficiency, ensuring the ethical application of AI across all legal domains. Adopting unified guidelines would promote consistency between criminal and civil law, simplifying the regulatory framework and avoiding unnecessary complexity in their application.

Recommendation

26. Develop ethical guidelines to regulate AI usage, ensuring it aligns with principles of accountability, data privacy, and the public interest whilst addressing the differences in criminal and civil law matters.

Assessment Framework for Courts and Tribunals

Question 33: Does the NSW AI Assurance Framework provide a useful model for Victorian courts and tribunals? Why or why not? What other models or guidelines should be considered?

The NSW AI Assurance Framework is a recently developed framework that assists government agencies to design, build and use AI-enabled products and solutions. It is a framework that assists in building and using AI technology appropriately.[312] It aims to support the NSW Government in ensuring that AI technology is used safely, securely and with clear accountability for the design and use of AI systems.[313]

The framework aims to achieve this through a three-step process where firstly, the risk factors of the AI instrument is assessed, where risk factors of the instrument are assessed in accordance with the prescribed risk matrices within the framework. Secondly, the instrument is to analyse in accordance with the questions that consider whether the instrument should operate as is, with additional treatments, or be ceased in development or use. Finally, the instrument is subjected to either self-assessment based on the way the questions were answered in the second step or is submitted to the NSW AI review body who will determine whether the instrument is able to continue with or without amendments.[314] It is the only risk assessment and assurance framework created and mandated in Australia.

An interesting feature of the framework is that NSW has created an AI Review Committee to guide and oversee the use of AI in government.[315] Per Victor Dominello, the Minister for Customer Service and Minister for Digital, the Committee is pivotal in building community trust and has been instrumental in enabling assurance in AI projects.[316] Within the legal system, the NSW AI Assurance Framework aims to distinguish the use of AI in legal decisions into five levels of risk. Please see question 34 for the breakdown of the risk assessment categories.

The framework aims to assess and categorise all generative AI solutions within those particular risk categories. Dependent on the level of risk involved in the utilisation of an instrument will

---

[312] NSW Government, *Artificial Intelligence Assurance Framework* (Report 2022) 5 <https://www.digital.nsw.gov.au/sites/default/files/2022-09/nsw-government-assurance-framework.pdf>.
[313] Ibid 6.
[314] Ibid 8.
[315] Ibid 5.
[316] Ibid.

determine whether or not the use of AI within that decision needs to be submitted to the AI review Committee. The Committee then determines whether the instrument or decision is allowed to proceed with or without changes. This aims to ensure that high risk use of AI that may have either irreversible or significant consequences on key stakeholders within the government and legal systems are only actioned with thorough human oversight.

However, whilst the NSW AI Assurance Framework serves as a model for ensuring consistency with the use of AI in government through implementation of ethical principles, there is no mention is this framework is mandated for everyone besides government agencies.[317]

Additionally, a critical downfall of the NSW Assurance Framework is that it requires self-assessment of the risk factors of the AI instrument by developers and users of the instrument prior to either utilising it or submitting it to the AI Review Committee for their determination on the safety of the instrument. According to Green, there are several faults with human oversight where he opines that people are unable to competently perform the desired oversight on their own uses and developments of AI instruments due to fundamental embedded biases.[318] This fault provides a false sense of security in adopting algorithms and enable vendors and agencies to shrink accountability for algorithmic harms.[319] As a critical step in the framework requires self-assessment and human oversight of the AI instrument utilised, this framework may therefore, prove to provide a false sense of security within the tools that are being used. This may lead to the use of many tools falsely categorised as safe, prior to their use and examination being determined as safe by the AI Review Committee. This point is further exacerbated by the lack of a clear definition of the risk levels within the framework.[320]

Finally, it is yet to be seen how this framework transforms into an effective and mandatory safeguard for the general development and deployment of AI within the legal system. The NSW AI Assurance framework is a mandated for all government agencies, however, its specific applications to courts and tribunals is yet to be examined as it is currently not specific to courts and tribunals.

---

[317] Ahmed Imran, Zena Assad and Thaye Choden, 'A critical assessment of AI governance and policy gaps in Australia' (2024*) Australasian Conference on Information Systems* 1, 6 ('A critical assessment of AI governance and policy gaps in Australia').
[318] Ben Green, 'The flaws of policies requiring human oversight of government algorithms' (2022) 45 *Computer Law & Security Review: The International Journal of Technology Law and Practice* 1, 7.
[319] Ibid.
[320] A critical assessment of AI governance and policy gaps in Australia (n 322) 6.

Therefore, whilst the NSW AI assurance framework is the only AI assurance framework that has been mandated in Australia, it is very broad application to government agencies, lack of definition of risk categories, and requirement for human oversight, hinder its ability to provide a useful model in aid Victorian Courts and Tribunals.

However, the creation of the AI Review Committee in NSW is a useful consideration that should adopted in Victoria as it would provide for an institutional agent to oversee and assess the high risk uses of AI prior to them being deployed in Victoria, limiting the impact of errors, increasing accountability, decreasing the potential for bias and discrimination, and upholding judicial independence and public trust.

The CEPEJ Risk Assessment used in the European Union might provide a better model for Victorian Courts and Tribunals to adopt as it is sector specific to the courts and tribunals and has a less complex framework for assessment.[321] It focuses on the uses and peculiarities of how the judicial system operates, rather than the generic assessment tool, providing sturdier basis for risk assessment than the undefined terms within the NSW AI Framework. Whilst the CEPEJ Risk Assessment framework is tailored and aligned with the EU Framework, a similar adaptation could prove more beneficial to Victorian courts and tribunals.

Recommendations

27. Consider adopting a model similar to the CEPEJ Risk Assessment utilised in the EU.

Question 34: How can risk categories (low, medium and high) be distinguished appropriately? What should be considered high risk?

I Distinguishing Risks

Given the rapid evolution of AI technology, a principle-based approach to risk classification is essential. Static, exhaustive lists of high-risk applications may fail to account for emerging technologies and their associated risks. Therefore, several factors should be considered when assigning a risk category to an AI application. A factor to consider is the type of decision-making process in which the AI system is involved. The degree of human intervention in the AI system's operation is another critical factor. AI applications that function with routine human oversight are

[321] European Commission for the Efficiency of Justice (CEPEJ), *Assessment Tool for the Operationalisation of the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment* (Report No CEPEJ (2023)16 final, Council of Europe, 4 December 2023).

less likely to be categorised as high-risk because humans can identify and correct errors before they cause harm. However, systems that make autonomous decisions without human input—particularly in high-stakes scenarios—pose a greater risk and require stricter oversight to prevent adverse outcomes. AI technology is developing at an unprecedented pace, with new applications and risks emerging regularly. Errors in AI systems can have varying consequences, from minor inconveniences to severe violations of individual rights. A principle-based framework ensures that risk assessments remain flexible and adaptive to technological advancements. For example, AI applications that appear low risk today may evolve into high-risk systems as their role in decision-making expands. Some risks associated with AI systems may not be immediately apparent. A principle-based approach enables courts and tribunals to identify and mitigate hidden or systemic risks that could compromise fairness, transparency, or accountability.

As discussed, the NSW Government's Artificial Intelligence Assurance Framework is limited in the case for Courts and Tribunals. However, it can provide a foundation for distinguishing risk categories in the first instance for Courts and Tribunals to then refine and develop. Courts and Tribunals can utilise this framework to systematically assess AI applications, ensuring appropriate safeguards are implemented at each level (e.g., see Figure 1). Each risk category should have corresponding risk management and assessment protocols to continuously evaluate and address emerging risks, ensuring that AI tools align with legal, ethical, and procedural standards.

II Defining and Capturing High Risk Applications

Within a judicial context, high-risk should be defined on where AI is used within decision-based tasks such as legal research, evidence evaluation or sentencing. Errors in these contexts can have irreversible consequences. For example, they could lead to wrongful convictions, which undermine the principles of justice and equality. Transparency is a cornerstone of procedural fairness, enabling parties to understand and challenge decisions. However, many AI systems operate as "black-boxes," making it difficult to explain how decisions are made.[322] With known risks of biased data, this lack of transparency undermines the right to a fair trial and the ability of parties to seek redress for errors or biases. In addition, cognitive biases such as automation bias where the integration of AI into decision-making processes could limit judicial discretion, creating a perception of dependency on technology. This dependency may compromise the

[322] David Freeman Engstrom et al, 'Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies' (Public Research Paper No 20-54, NYU School of Law, 2020) 75.

autonomy of judges and diminish public confidence in the justice system's ability to deliver impartial outcomes. The judiciary must carefully consider the ethical implications of delegating certain aspects of judicial decision-making to AI and ensure that human judges remain central to the process.

Recommendations

28. Adopt a principle-based framework to identify risk categories rather than relying on static lists of applications. This allows for flexibility and adaptation as AI technology evolves. The core principles should prioritise protection of fundamental rights, procedural fairness and judicial independence.
29. Establish risk management and assessment protocols for each risk category. These protocols should address emerging risks and ensure alignment with legal, ethical, and procedural standards.

Figure 1: Example Risk Categorisation Protocol

Risk Level	Description	Examples of AI Applications	Implementation Guidance
Very Low	AI analyses data to produce insights or patterns for humans to interpret and use in contexts that do not directly affect operations or real-time decisions	Analytics tools reporting on historical case trends.	Can be implemented with minimal oversight, subject to regular monitoring and reviews.
Low	AI provides actionable information (e.g., alerts, predictions) to humans, who then decide how to act on it, with the risk of harm or error from these insights as very low.	Administrative tasks such as case scheduling or document management.	Suitable for implementation with standard oversight and operational guidelines.
Medium	AI independently makes decisions or provides recommendations that are acted upon with little or no human intervention and there is a limited potential for harm.	AI-powered tools for mediation or settlement negotiations; public-facing chatbots.	Implement with clear accountability mechanisms, bias detection protocols, and contingency plans for rectifying errors.
High	AI provides decisions or recommendations that directly influence operations, such as legal rulings, sentencing, or case allocation and human involvement is limited.	Legal research tools influencing case outcomes; evidence evaluation systems.	Requires rigorous risk assessments, continuous oversight, and mechanisms to challenge or override AI decisions where necessary.
Very High	AI operates independently, with no human intervention, in high-stakes contexts where decisions can lead to significant consequences, such as affecting a person's freedom, safety, or livelihood.	Autonomous systems for custodial sentence recommendations; or those operating as 'black-boxes'.	Should **not** be implemented due to the unacceptable risk of harm, bias, and erosion of judicial independence.

Summary of Recommendations

1. Establish a "humans-in-the-loop" approach in all facets of AI use within courts and tribunals, ensuring that AI serves as a supplementary tool to enhance, not replace, human judgment. Emphasise the importance of human qualities such as empathy, moral reasoning, and contextual understanding in judicial decisions.
2. Additionally, courts must disclose AI usage in judicial processes and provide clear warnings and disclosures about the potential inaccuracies of AI tools or chatbots to users before engagement.
3. Should Victorian courts and tribunals adopt AI tools in decision making processes in the future, strong regulations must be developed. Such regulations must aim to maintain high rates of accuracy in outcomes. In addition, regulations should be designed to deter decisionmakers from placing excessive reliance on AI tools during the decision-making process.
4. Require AI systems used in courts and tribunals to provide sufficient explanations of how they arrive at specific conclusions or recommendations.
5. Prohibit the use of predictive analytic tools to operate as a standalone instrument for judicial determination.
6. Where predictive analytics are being utilised by judges and tribunals, ensure they are used in combination with independent and competent human oversight and merely as a tool to assist decision-making rather than a determinative instrument for judicial determination.
7. Prohibit the use of judicial analytic tools.
8. Adopt a Principles-Based Regulatory Framework with a focus on overarching principles such as fairness, transparency, accountability, and human oversight to ensure adaptability as AI technology evolves. This approach allows for flexibility and adjustments based on actual data, rather than perceived harms, enabling regulations to remain effective in a dynamic AI landscape.
9. Consider the potential need for overarching legislation in the future. Such legislation can provide consistency across sectors and address emerging risks effectively, ensuring a comprehensive and unified regulatory framework for AI technologies.
10. Apply stricter regulations in the case of criminal matters, considering that they are more sensitive. Ensure that a judge presides over the case and that predictive AI is not used to for decision making unless it is being used for calculating sentencing.
11. Review ICT projects implemented in Australian courts and tribunals over the past 20 years as these may be value for identifying key factors to improve likelihood of successful AI projects.

12. Establish robust security protocols to mitigate risks, such as prohibiting the use of sensitive court data for AI training.
13. Establish working groups, such as Quality of Justice or Cyberjustice teams, to facilitate education, collect and report performance data, and share experiences across jurisdictions to build confidence and promote informed AI adoption.
14. Develop an AI assurance framework that includes mandatory impact assessments, audits, and reporting, supported by regulatory sandboxes and pilot programs to test AI tools in controlled environments before full implementation. This framework will also allow ongoing monitoring and review of processes to reassess risks as AI technology evolves.
15. Evaluate AI systems and algorithms to ensure courts operate in compliance with copyright and other legal obligations, avoiding complicity in potential infringements by AI tools.
16. Implement Robust Bias Mitigation and Compliance Measures by establishing pre-approval processes for AI tools to evaluate biases, discontinuing tools that fail to address discrimination.
17. Stay informed about AI regulatory developments in international jurisdictions like the UK, EU, and US to anticipate concerns, adopt best practices, and ensure alignment with global standards.
18. The judiciary must retain full autonomy over AI technologies by establishing judicial oversight committees, utilising in-house expertise (if feasible) such as a designated AI team and a preference for open-source software to ensure transparency, accountability and independence from external influence.
19. Create clear, practical guidelines for judges and tribunal members on AI usage, covering topics like data privacy, algorithmic bias, and human oversight. Mandate AI training for all court and tribunal judicial employees to ensure understanding of AI's role, risks, and limitations.
20. Mandate regular audits of all AI systems used in courts and tribunals to ensure alignment with current legal practices, identify and mitigate algorithmic biases, and address potential errors or prejudicial outcomes. Require independent third-party audits for high-risk applications to ensure compliance with ethical and technical standards.
21. Implement a tailored, layered approach to AI disclosures. For low-risk applications, provide public summaries that are transparent and accessible without overwhelming stakeholders. For high-risk systems, require detailed technical disclosures—including algorithmic design, training data, and oversight mechanisms—accessible to regulators and judicial authorities.

This ensures accountability and transparency while maintaining operational confidentiality for proprietary systems.

22. Require the disclosure of the specific roles AI make in decision-making, provide high-level algorithm descriptions and their limitations, and outline mechanisms for challenging AI-influenced decisions to uphold procedural fairness. Introduce measures like watermarks or labels to identify AI interactions, balancing transparency with appropriate detail for public understanding and the specific risks associated with AI use.
23. Balance transparency with considering the appropriate degree of detail for ordinary members of the public and the risk associated with the specific AI use.
24. Launch public education strategies, such as forums and workshops, to engage the public and key stakeholders in discussions about AI in the legal system. Encourage interdisciplinary dialogue between legal professionals, AI developers, and other stakeholders to ensure effective implementation and accountability throughout AI application stages.
25. Create a centralised regulatory authority, such as a Victorian AI Assessment and Review Committee, to oversee the development, deployment, and use of AI in the judiciary. This body should have the authority to prevent projects with high risks, ensure AI development aligns with the public interest, and uphold judicial independence.
26. Develop ethical guidelines to regulate AI usage, ensuring it aligns with principles of accountability, data privacy, and the public interest whilst addressing the differences in criminal and civil law matters.
27. Consider adopting a model similar to the CEPEJ Risk Assessment utilised in the EU.
28. Adopt a principle-based framework to identify risk categories rather than relying on static lists of applications. This allows for flexibility and adaptation as AI technology evolves. The core principles should prioritise protection of fundamental rights, procedural fairness and judicial independence.
29. Establish risk management and assessment protocols for each risk category. These protocols should address emerging risks and ensure alignment with legal, ethical, and procedural standards.

Conclusion

In conclusion, the integration of AI into the Victorian legal system presents a complex but ultimately manageable challenge. The potential benefits – increased efficiency, reduced costs, and improved access to justice – are significant and should not be dismissed. However, these benefits must be pursued cautiously, with a clear understanding of the inherent risks. The potential for a diverse range of biases, the lack of transparency in some AI systems, and the crucial need to preserve human judgment in judicial decision-making are all critical concerns that demand careful consideration and proactive mitigation.

As this submission has stated, a balanced regulatory approach is essential. This approach must move beyond general principles and embrace concrete guidelines and practical tools. By learning from international experiences and adopting an incremental strategy, Victoria can develop a robust and adaptable framework for AI governance in the legal sector. This framework must prioritise judicial autonomy, transparency, accountability, bias mitigation through disclosure, continuous training for legal professionals, and a tiered approach to transparency based on risk level. The establishment of a centralised regulatory body and an independent AI Assessment Review Committee, combined with sector-specific risk assessment frameworks adapted to the unique needs of the Victorian legal system, will be crucial in ensuring responsible and ethical AI implementation.

Ultimately, the goal is not to prevent the integration of AI into the legal system, but rather to ensure that this integration serves the interests of justice. This requires ongoing dialogue, collaboration among stakeholders, and a commitment to ongoing monitoring and evaluation as AI technology continues to evolve. From there, by carefully navigating the complex interplay of opportunities and risks, Victorian Courts and Tribunals can effectively regulate the integration of AI in the judicial system.

Bibliography

A *Articles/Books/Reports*

Adams, R, 'Ethical obligations and the digital age' (2020) 45(2) *Legal Tech Journal* 132

Abiri, Gilad, Huang, Yue, 'A Red Flag? China's Generative AI Dilemma' (2023) 37 *Harvard Journal of Law & Technology* 2

Aletras, Nikolaos et al, 'Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective' (2016) 2(93) *PeerJ Computer Science* 1

Barry, Brian, 'AI for Assisting Judicial Decision-making: Implications for the Future of Open Justice' (2024) 98 *Australian Law Journal* 656

Bell, Felicity et al, *AI Decision-Making and the Courts: A Guide for Judges, Tribunal Members and Court Administrators* (Report, Australasian Institute of Judicial Administration Incorporated, December 2023) 3

Birchfield, Vicki, 'From Roadmap to Regulation: Will There Be a Transatlantic Approach to Governing Artificial Intelligence?' (2024) 46(7) *Journal of European Integration* 1053

Berk, Richard, 'Artificial Intelligence, Predictive Policing, and Risk Assessment for Law Enforcement' (2021) 4(1) *Annual Review of Criminology* 209

Brekke, Viktor Alan, '*Artificial Intelligence in the judicial system: Maintaining the independency of the judiciary power in the development, implementation, and use of artificial intelligence*' (Master thesis, University of Oslo, 2022)

Brown, L, 'Navigating ethical concerns in the era of digital lawyering' (2018) 19(1) Journal of Legal Ethics 27

Campbell, Ray Worthy, 'Artificial intelligence in the courtroom: The delivery of justice in the age of machine learning' (2023) 15 *Revista Forumul Judecatorilor* 1

Cantero Gamito, Maria, and Christopher Marsden, 'Artificial Intelligence Co-Regulation? The Role of Standards in the EU AI Act' (2024) 32(1) *International Journal of Law and Information Technology* 14

Carnat, Irina, 'Addressing the risks of generative AI for the judiciary: The accountability framework(s) under the EU AI Act' (2024) 56 *Computer Law & Security Review: The International Journal of Technology Law and Practice*

Centre for Data Ethics and Innovation, '*The roadmap to an effective AI assurance ecosystem*' (Independent Report, 2021)

Chaudhary, Gyandeep, 'Unveiling the Black-box: Bringing Algorithmic Transparency to AI' (2024) 18(1) *Masaryk University Journal of Law and Technology* 93

Chulliev, Shukhrat, 'Transforming Judicial Competencies: A Framework For Judge Training And Qualification In Ai-enhanced Court Systems' (2024) (3(2) *Elita* 1

Clark, Jack and Gillian Hadfield, 'Regulatory markets for AI safety' (2019) *ArXiv* 1

Commission on Judicial Conduct, *Improper Delegation of Judicial Authorities* (Annual Report 1 March 2019)

Daly, Paul, 'Artificial Administration: Administrative Law, Administrative Justice and Accountability in the Age of Machines' (2023) 30 *Australian Journal of Administrative Law* 95

de Bruijn, Hans, Martijn Warnier and Marijn Janssen, 'The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making' (2022) 39(2) *Government Information Quarterly* 101666

Deeks, Ashley, 'The Judicial Demand for Explainable Artificial Intelligence' (2019) 119(7) *Columbia Law Review* 1829

Di Porto, Fabiana, 'Algorithmic Disclosure Rules' (2021) 31(1) *European Journal of Risk Regulation* 13

Downing, Mandy (ed), *The Routledge Handbook of Human Research Ethics and Integrity in Australia* (Routledge, 1st ed, 2024) Dzindolet, Mary, et al, 'The perceived utility of human and automated aids in a visual detection task' (2002) 44 *Human Factors: The Journal of the Human Factors and Ergonomics Society* 79

Engstrom, David Freeman, et al, 'Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies' (Public Research Paper No 20-54, NYU School of Law, 2020)

European Commission, *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust* (Report, February 2020)

European Commission for the Efficiency of Justice (CEPEJ), *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment* (Report, Council of Europe, December 2018)

European Commission for the Efficiency of Justice (CEPEJ), *Assessment Tool for the Operationalisation of the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment* (Report No CEPEJ (2023)16 final, Council of Europe, 4 December 2023)

Fabri, Marco 'From Court Automation to e-Justice and beyond in Europe' (2024) 15(3) *International Journal for Court Administration 7*

Finocchiaro, Giusella, 'The Regulation of Artificial Intelligence' (2022) 39(4) *AI & SOCIETY* 1961

Freeman, Laura, et al, 'The path to a consensus on artificial intelligence assurance' (2022) 55(3) *Computer (IEEE Computer)* 82

Gans-Combe, Caroline, 'Automated Justice: Issues, Benefits and Risks in the Use of Artificial Intelligence and Its Algorithms in Access to Justice and Law Enforcement' (2022) *Ethics, Integrity and Policymaking: The Value of the Case Study* 175

Gikay, Asress Adimi, 'Risks, innovation, and adaptability in the UK's incrementalism versus the European Union's comprehensive artificial intelligence regulation' (2024) 32 *International Journal of Law and Information Technology* 1

Glicksman, Robert and Sidney Shapiro, 'Improving Regulation Through Incremental Adjustment' (2004) 52 *University of Kansas Law Review* 1179

Green, Ben, 'The flaws of policies requiring human oversight of government algorithms' (2022) 45 *Computer Law & Security Review* 1

Green, Ben and Yiling Chen, 'Algorithmic risk assessments can alter human decision-making processes in high-stakes government contexts' (2021) 5 *Proceedings of the ACM on Human-Computer Interaction* 1

Green, Harriet, 'Consciousness over Code: How Judicial Review can Address Algorithmic Decision-making in Policing' (2024) 5(1*) York Law Review* 8

Hanna, Meena, 'Robo-Judge: Common Law Theory and the Artificial Intelligent Judiciart' (2019) 29 *Journal of Judicial Administration* 22

Hannah-Moffat, Kelly, 'The Uncertainties of Risk Assessment: Partiality, Transparency and Just Decisions' (2015) 27(4) *Federal Sentencing Reporter* 244

Imran, Ahmed, Zena Assad and Thaye Choden, 'A critical assessment of AI governance and policy gaps in Australia' (2024*) Australasian Conference on Information Systems* 1

Jennett, Victoria, *Fighting Judicial Corruption: Topic Guide* (Report, 31 October 2014) 1

KPMG Australia and The University of Queensland, Trust in Artificial Intelligence: Australian Insights (Report, October 2020)

Kotsoglou, Kyriakos and Marion Oswald, 'The Long Arm of the Algorithm? Automated Facial Recognition as Evidence and Trigger for Police Intervention' (2020) 2 *Forensic Science International: Synergy* 86

Lamchek, Jayson, and Shiri Krebs, 'Cybersecurity Research and Society: Considerations for Researchers and Human Research Ethics Committees', in Bruce M Smyth, Michael A Martin and

Legg, Michael and Felicity Bell, 'Artificial Intelligence and the Legal Profession: Becoming the AI-Enhanced Lawyer' (2019) 38(2) *University of Tasmania Law Review* 34

Liu, Han-Wei, Ching-Fu Lin and Yu-Jie Chen, 'Beyond State v Loomis: artificial intelligence, government algorithmization and accountability' (2019) 27(2) *International Journal of Law and Information Technology* 122

Lopes, Giovana, 'AI and Judicial decision-making; Evaluating the role of AI in debiasing' (2024) *Journal for Technology Assessment in Theory and Practice* 28

Luoma, Sanna, 'AI Improving the Delivery of Justice and How Courts Operate' (2018) *How Will AI Shape the Future of Law* 63

McGill, Jena and Amy Salyzyn, 'Judging by the Numbers: Judicial Analytics, the Justice System and its Stakeholders' (2021) 44(1) *Dalhousie Law Journal* 249

Malek, Abdul, 'Criminal courts' artificial intelligence: the way it reinforces bias and discrimination' (2022) 2(1) *AI and Ethics* 233

Morin-Martel, Alexis, 'Machine learning in bail decisions and judges' trustworthiness' (2023) 39 *AI & Soc* 2044

Necz, Daniel, 'Rules over words: Regulation of chatbots in the legal market and ethical considerations' (2024) 64(3) *Hungarian Journal of Legal Studies* 472

Nowotko, Paweł Marcin, 'AI in judicial application of law and the right to a court' (2021) 192 *Procedia Computer Science* 2220

Ombudsman, *Automated Decision-Making: Report* (Report, Commonwealth of Australia, 2025) 33

Parasuraman, Raja and Dietrich Manzey, 'Complacency and bias in human use of automation' (2010) 52(3) *Human Factors: The Journal of the Human Factors and Ergonomics Society* 381

Pesapane, Filippo et al, 'Artificial Intelligence as a Medical Device in Radiology: Ethical and Regulatory Issues in Europe and the United States' (2018) 9 *insights into Imaging* 745

NSW Government, *Artificial Intelligence Assurance Framework* (Report 2022) 5

Reichman, Amnon, Yair Sagy and Shlomi Balaban, 'From a Panacea to a Panopticon: The Use and Misuse of Technology in the Regulation of Judges' (2020) 71(3) *Hasting Law Journal* 589

Roberts, Huw, et al, Artificial intelligence regulation in the United Kingdom: a path to good governance and global leadership? (2023) 12(2) *Internet Policy Review* 1

Royal Commission into the Robodebt Scheme (Final Report, 7 July 2023)

Sadikov Ruslan, 'Challenges and Opportunities for Legal Practice and the Legal Profession in the Cyber Age' (2023) 1(4) *International Journal of Law and Policy* 1

Salomão, Luis Felipe, FGV Conhecimento *Tecnologia Aplicada à Gestão Dos Conflitos No Âmbito Do Poder Judiciário Brasileiro* (Report, 2nd ed, December 2020)

Santos Rocha, Thiago, "Brazil" in Dariusz Szostek and Mariusz Zalucki (ed) *Legal Tech Information technology tools in the administration of justice (European Law Institute* 2023) 487

Schmitz, Amy J and John Zeleznikow, 'Intelligent Legal Tech to Empower Self-Represented Litigants (2021) 23 *Columbia Science & Technology Law Review* 142

Sheehan, Matt, 'China's AI Regulations and How They Get Made' (2023*) Carnegie Endowment for International Peace*

Singh, Manjari, 'Review on Role of Artificial Intelligence in The Life of Legal Profession' (2024) 6(3) *International Journal of Legal Science and Innovation* 1087

Sourdin, Tania, 'Replacing, Supporting or Enhancing Judges? Judge AI Considerations for the Future' (2024) 98 Australian Law Journal 696

Spasojević, Dijana et al, 'Study on the Use of Innovative Technologies in the Justice Field (Final Report to the European Commission No 1, Directorate-General for Justice and Consumers, September 2020)

Srivastava, Neerav, 'Liability for Chatbots: A Psychbot Negligence Case Study and the Need for Reasonable Human Oversight' (2023) 28 *Torts Law Journal* 155

Stewart, Pamela and Anita Stuhmcke 'Judicial Analytics and Australian Courts: A Call for National Ethical Guidelines' (2020) 45(2) *Alternative Law Journal* 82

Terzidou, Kalliopi, 'The Use of Artificial Intelligence in the Judiciary and Its Compliance with the Right to a Fair Trial' (2022) 31 *Journal of Judicial Administration* 155

Zalnieriute, Monika and Felicity Bell, 'Technology and Judicial Role' in Gabrielle Appleby and Andrew Lynch (eds), *The Judiciary and the Court: Individual, Collegial and Institutional Judicial Dynamics in Australia* (Cambridge University Press, 2021)

Zalnieruite, Monika, Lyria Moses and George Williams 'The rule of Law and Automation of Government Decision Making' (2019) 82(3) *The Modern Law Review* 425

B Cases

Handa v Mallick [2024] FedCFamC2F 957

Hemmett v Market Direct Group Pty Ltd [2018] WASC 214

Loomis v State of Wisconsin (2017) 371 Wis 2d 235

State of Wisconsin v Loomis (2016) 881 N.W.2d 749

C *Legislation*

Artificial Intelligence (Regulation) HL Bill (2023)

生成式人工智能服务管理办法 [Measures for the Management of Generative Artificial Intelligence Services] (People's Republic of China) Cyberspace Administration of China, 11 April 2023

互联网信息服务深度合成管理规定 [Provisions on the Administration of Deep Synthesis Internet Information Services] (People's Republic of China) Ministry of Industry and Information Technology of the People's Republic of China, 25 November 2022

互联网信息服务算法推荐管理规定 [Provisions on the Management of Algorithmic Recommendations in Internet Information Services] (People's Republic of China) Ministry of Industry and Information Technology of the People's Republic of China, 31 December 2021

Charter of Human Rights and Responsibilities Act 2006 (Vic)

Commonwealth of Australia *Constitution Act* (Cth) 1901

Ordinance 271 [Portaria Nº 271] (Brazil) December 2020

Projeto de Lei No. 2.338/2023 [Bill No. 2,338/2023] (Brazil).

Regulation (EU) 2024/1689 (Artificial Intelligence Act) [2024] OJ L 2024/1689, Ch I Art 3(63).

Resolution no. 332/2020 [Resolução Nº 332] (Brazil) August 2020

Resolution no. 363 [Resolução Nº 363] (Brazil) January 2021

Resolution no. 370 [Resolução Nº 370] (Brazil) January 2021

D *Other*

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

Administration of Victorian Courts. Performance Audit: Court Services Victoria. (Tabled 13 October 2021) *Victorian Auditors General's Office* (Web Page) <https://www.audit.vic.gov.au/sites/default/files/2021-10/AOC_transcript.pdf>

'*Artificial Intelligence (AI) – Judicial Guidance'*, Courts and Tribunals Judiciary (UK) (Web Page) <https://www.judiciary.uk/guidance-and-resources/artificial-intelligence-ai-judicial-guidance/>

Australian Human Rights Commission, 'Mandatory Guardrails for AI in High-risk Settings' (Submission to Department of Industry, Science and Resources, Australian Human Rights Commission, 04 October 2024)

Attorney General's Office, *'AGU Passa a Utilizar Ferramentas de Inteligência Artificial na Produção de Documentos Jurídicos'* [AGU Starts Using Artificial Intelligence Tools to Produce Legal Documents] (September 2024) <https://www.gov.br/agu/pt-br/comunicacao/noticias/agu-passa-a-utilizar-ferramentas-de-inteligencia-artificial-na-producao-de-documentos-juridicos>.

Becker, Josh, 'Data Analytics & E-Discovery' (Speech, Litigation Cravath Panel Discussion at 'The Future of Law – the Case for Analytics', 29 March 2018).

Canadian Judicial Council, 'Canadian Judicial Council issues Guidelines for the Use of Artificial Intelligence in Canadian Courts' (Press Release, October, 2024) 1 < https://cjc-ccm.ca/en/news/canadian-judicial-council-issues-guidelines-use-artificial-intelligence-canadian-courts>

Canadian Judicial Council, *Guidelines for the Use of Artificial Intelligence in Canadian Courts*, 2024

Courts and Tribunals Judiciary (UK), 'Artificial Intelligence (AI) Guidance for Judicial Office Holders' (2023)

Department for Science, Innovation and Technology, *A pro-innovation approach to AI regulation* (Cm 815, 2023)

Department of Finance , Australian Government, (Cth), *Implementing Australia's AI Ethics Principles in Government* <https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government/implementing-australias-ai-ethics-principles-government#6-transparency-and-explainability>.

Department of Industry, Science and Resources (Cth), 'Safe and Responsible AI in Australia: Proposals Paper for Introducing Mandatory Guardrails for AI in High-Risk Settings' (September 2024) <https:///proposals_paper_for_introducing_mandatory_guardrails_for_ai_in_high_risk_settings.pdf>

European Commission for the Efficiency of Justice (CEPEJ), *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment* (Report, Council of Europe, December 2018)

European Commission for the Efficiency of Justice (CEPEJ), *Assessment Tool for the Operationalisation of the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment* (Report No CEPEJ (2023)16 Final, Council of Europe, 4 December 2023)

Felsky, Martin and Karen Eltis, *Guidelines for the Use of Artificial Intelligence in Canadian Courts* (Published Guidelines No 1, September 2024)

HM Government, *National AI Strategy* (Cm 525, 2021)

HM Treasury, 'Pro-innovation Regulation of Technologies Review: Digital Technologies' (2023)

Justice Perry, Melissa and Sonya Campbell, 'AI and Automated Decision-Making: Are you just another number?' (Speech, Gilbert + Tobin Centre of Public Law, 21 October 2021)

Law Council Australia, Submission No 2 to Department of Industry, Science and Resources, *Introducing Mandatory Guardrails for AI in High-risk Settings: Proposals Paper* (9 October 2024)

Law Council Australia, Submission No 1 to Department of Industry, Science and Resources, *Safe and Responsible AI in Australia* (17 August 2023)

Law Council Australia, Submission to Select Committee on Adopting Artificial Intelligence, *Inquiry into the opportunities and impacts of the uptake of artificial intelligence technologies in Australia* (20 May 2024)

Li, Barbara and Amaya Zhou, 'Navigating the Complexities of AI Regulation in China,' *Reed Smith In-depth* (Blog Post 7 August 2024) <https://www.reedsmith.com/en/perspectives/2024/08/navigating-the-complexities-of-ai-regulation-in-china>

Lima, Gabriel, Nina Grgić-Hlača and Meeyoung Cha, 'Human Perceptions on Moral Responsibility of AI: A Case Study in AI-Assisted Bail Decision-Making', *CHI Conference on Human Factors in Computing Systems* (2021)

Ministério da Ciência, Tecnologia e Inovação [Ministry of Science, Technology and Innovation], *Plano Brasileiro de Inteligência Artificial (PBIA) 2024-2028* [Brazilian Artificial Intelligence Plan] (August 2024) <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2024/07/plano-brasileiro-de-ia-tera-supercomputador-e-investimento-de-r-23-bilhoes-em-quatro-anos/ia_para_o_bem_de_todos.pdf/view>

Państwo, Fundacja Moje, 'Algorithm of the System of Random Allocation of Cases finally disclosed! (Web Page, 22 September 2021) <https://mojepanstwo.pl/aktualnosci/773>

Pain, Rachel, 'Escaping the Matrix: Met Admits Gangs Matric Unlawful' *Mountford* (Online,18 November 2022) <https://www.mountfordchambers.com/escaping-the-matrix-met-admits-gangs-matrix-unlawful/>

Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonized on Artificial Intelligence (Artificial Intelligence Act) [2021] COD 0106

Queensland Courts, *Artificial Intelligence Guidelines for Non-Lawyers* (Guideline 13 May 2024) 1 <https://www.courts.qld.gov.au/__data/assets/pdf_file/0012/798375/artificial-intelligence-guidelines-for-non-lawyers.pdf>

'Reuters**, '***Brazil Hires OpenAI to Cut Costs of Court Battles'* (Web Page, 11 June 2024) <https://www.reuters.com/technology/artificial-intelligence/brazil-hires-openai-cut-costs-court-battles-2024-06-11/>

Sharma, Sangeet et al, '*A Study of AI-Based Systems in the Judicial Interpretation of the Law'* (Conference Paper, India, 2023)

Supreme Court of Victoria, *Guidelines for Litigants: Responsible Use of Artificial Intelligence in Litigation* (June 2022)

Victoria Legal Aid, Submission to the Australian Human Rights Commission, Human Rights and Technology, Discussion Paper (2019)

Zalnieriute, Monika, Submission No 3 to Australian Law Reform Commission, *Review of Judicial Impartiality* (4 June 2021)

Zeleznikow, John, 'Algorithmic Justice Symposium' (Conference Paper, University of Newcastle, 14 July 2023)