Office of the Victorian Information Commissioner

Phone: 1300 00 6842
Email: enquiries@ovic.vic.gov.au
PO Box 24274
Melbourne Victoria 3001

12 December 2024

The Hon. Jennifer Coate AM
Acting Chair
Victorian Law Reform Commission

By email only: law.reform@lawreform.vic.gov.au

Dear Ms Coate,

Thank you for the opportunity to provide a submission in response to the Victorian Law Reform Commission's (**VLRC**) consultation paper, *Artificial Intelligence in Victoria's Courts and Tribunals* (**consultation paper**).

The Office of the Victorian Information Commissioner (**OVIC**) is the primary regulator of information privacy, information security and freedom of information for the Victorian public sector, administering both the *Privacy and Data Protection Act 2014* (Vic) (**PDP Act**) and *Freedom of Information Act 1982* (Vic) (**FOI Act**). OVIC's role includes upholding, and advocating for, the privacy rights of the Victorian community, uplifting information security practice across the Victorian public sector, and promoting public access to government-held information.

OVIC hopes this submission will help the VLRC to further understand the privacy, information security and information access implications and risks of using Artificial Intelligence (**AI**) systems, particularly in a judicial setting, and OVIC's minimum expectations for VPS organisations in managing these risks.

OVIC's submission covers the following:

- the application of the FOI Act and PDP Act to courts and tribunals
- the definition of AI
- the different types of Generative AI, and why we need to understand the difference
- the privacy and information security risks relevant to courts and tribunals' use of AI
- ways in which privacy and information security risks can be managed
- principles for safe and responsible use of AI, and guidance for courts and tribunals
- a summary of OVIC's recommendations to the VLRC.

I have no objection to this submission being published by VLRC, subject to my signature being removed. I also propose to publish a copy of this submission on OVIC's website.

If you have any questions about the comments in this submission, or would like to discuss these issues further, please contact Fathia Tayib, Senior Policy Officer, at fathia.tayib@ovic.vic.gov.au.

Yours sincerely

[redacted]

Sean Morrison
Information Commissioner

The application of the FOI Act and PDP Act to courts and tribunals

This section explains the circumstances in which the statutory obligations found in the FOI Act and PDP Act apply to courts and tribunals and recommends expanding the application of the PDP Act to a broader range of court and tribunal functions.

FOI Act

The FOI Act creates a legally enforceable right for any person to request access to documents in the possession of Victorian Government agencies and official documents of Victorian Government Ministers. The right to access information is recognised as essential to building effective, accountable and participatory democracy at all levels of government.[1]

The right, however, is not absolute and has limited application to courts and tribunals.

When the FOI Act will not apply

Courts and tribunals are not subject to the FOI Act when exercising judicial functions (such as hearing and deciding cases).[2] This means that there is no right to request access to documents relating to a court or tribunal's use of an AI system that supports the exercise of judicial functions (such as summarising court documents, preparing briefs or recommending outcomes).

Section 29B of the FOI Act also exempts from access, documents of Court Services Victoria (**CSV**) that relate to the exercise of a judicial or quasi-judicial function of a Victorian court or the Victorian Civil and Administrative Tribunal (**VCAT**).[3] This means that where the elements of the exemption in section 29B are met, it can be applied to withhold a document from release under the FOI Act.

When the FOI Act will apply

The FOI Act does apply to documents in the possession of a court or tribunal that relate to non-judicial functions, such as a court or tribunal's administrative functions.[4]

'Administrative function' means the apparatus supporting the exercise of judicial or quasi-judicial functions, such as the management and administration of registry and office resources, logistical support, infrastructure, physical necessities, travel and accommodation or the platform that enables judicial and quasi-judicial functions to occur.[5]

---

[1] For more information on the right and its purposes and benefits, see OVIC's submission (submission 55) to the Integrity and Oversight Committee's Inquiry into the operation of the *Freedom of Information Act 1982* (Vic); United Nations Sustainable Development Goal 16.10: 'Ensure public access to information and protect fundamental freedoms, in accordance with national legislation and international agreements'; United Nations Office of the High Commissioner for Human Rights, Freedom of opinion and expression: Report of the Office of the United Nations High Commissioner for Human Rights, 10 January 2022, A/HRC/49/38, res 44/12.

[2] FOI Act, sections 6. See the FOI Guidelines for further information.

[3] FOI Act, section 29B. See the FOI Guidelines for further information.

[4] FOI Act, sections 6 and 29B. See the FOI Guidelines for further information.

[5] *Kline v Official Secretary to the Governor General* [2013] HCA 52 [41], [47], [71]-[72], [74], [77].

This means that an individual has a right to request access to documents relating to a court or tribunal's use of an AI system that supports its administrative functions.

The FOI Act only applies to documents in the possession of an agency. 'Possession' means actual physical possession, as well as constructive possession (for example, a contractual or legal right to require someone to provide the document to the agency).[6]

Constructive possession is relevant where documents are held by a third party contracted service provider, or third-party owner or developer of a technology, system or tool, that is used by the agency in the performance of its functions. The ability for courts and tribunals to call for access to documents held by AI developers and deployers is relevant to issues of accountability and transparency, discussed later in this submission.

PDP Act

The PDP Act contains 10 Information Privacy Principles (**IPPs**), which govern how Victorian public sector organisations should handle personal information.[7] The privacy obligations of Victorian public sector organisations are outlined in Part 3 of the PDP Act.

Personal information is information that identifies an individual or could reasonably identify an individual (either on its own, or when combined with other information). This includes opinions, and it does not matter whether the information is true or not.[8] Certain attributes can serve as proxies for personal information, also known as 'indirect identifiers', and can include things such as post code, gender, occupation, and tertiary education.

Sensitive information is a subset of personal information, and is subject to stricter requirements around its collection, and in some cases its use and disclosure. Sensitive information includes racial or ethnic origin, political opinions or associations, and health or genetic information.[9]

The IPPs set out:

- when and how an organisation can collect personal information, including what individuals should be told about the collection of their personal information (IPP 1)
- when an organisation can use and disclose personal information for a secondary purpose (IPP 2)
- how personal information should be handled:
    - to ensure it is accurate, complete and up to date (IPP 3)

[6] See FOI Guidelines, section 5 [1.19].
[7] Personal information is defined in section 3 of the PDP Act. While OVIC regulates the collection and handling of personal information by Victorian public sector organisations, the Health Complaints Commissioner has oversight of health information privacy under the *Health Records Act 2001*.
[8] PDP Act, section 3.
[9] PDP Act, Schedule 1. See the Guidelines to the IPPs for further information.

  - to ensure it is protected from misuse, loss and unauthorised access, modification or disclosure (IPP 4.1)
  - to ensure it is destroyed when no longer needed for any purpose (IPP 4.2)

- general transparency requirements to prepare and publish privacy policies, and to be open with individuals about the personal information an organisation holds, and how it handles their personal information (IPP 5)
- when an individual can seek access to and/or correction of their personal information from an organisation (IPP 6)
- when an organisation can use unique identifiers (IPP 7)
- when an organisation must allow a person to be anonymous when transacting with the organisation (IPP 8)
- when personal information can be transferred outside of Victoria (IPP 9), and
- stricter requirements for when sensitive information can be collected (IPP 10).

Part 4 of the PDP Act sets out the legislative framework for monitoring and assuring the security of public sector information. Referred to in the PDP Act as 'public sector data', public sector information is any information (including personal information) obtained, received or held by an agency or body to which Part 4 applies.[10]

Under Part 4 of the PDP Act, the Information Commissioner has developed the Victorian Protective Data Security Framework and issued the Victorian Protective Data Security Standards (VPDSS). The VPDSS set out mandatory requirements to protect Victorian public sector information, including oversight of limited privacy and information security incidents.

When the PDP Act will and will not apply

Courts and tribunals carrying out a judicial or quasi-judicial function are exempt from complying with the IPPs and VPDSS under section 10 of the PDP Act. The exemption is operationally broad, applying to courts and tribunals, a holder of a judicial or quasi-judicial office, a registry or other office of a court or tribunal, or the staff of such a registry or other office, in relation to those matters which relate to the judicial, or quasi-judicial, functions of the court of tribunal. The effect of this broad exemption is that the information handling practices and processes of courts and tribunals in most instances are not covered by the PDP Act or subject to OVIC's oversight.

The PDP Act applies only to the administrative functions of courts and tribunals, such as maintenance of staff records and general administrative matters.[11]

[10] PDP Act, section 3.
[11] See Guidelines to the IPPs.

Expanding the application of the PDP Act to courts and tribunals

OVIC holds the view that courts and tribunals should not be broadly exempt under section 10 of the PDP Act from complying with the IPPs and VDPSS, due to the nature of the public sector information they handle, which includes (but is not limited to) personal, sensitive and health information.

Courts and tribunals are responsible for delivering critical functions and services to the Victorian community. A privacy or information security incident relating to the information that courts and tribunals hold could result in serious harm to individuals and to courts and tribunals themselves. The potential for incidents can be exacerbated by the use of AI.

The consultation paper notes that courts and tribunals often voluntarily comply with federal and state privacy laws,[12] which OVIC commends. However, this does not mean that courts and tribunals adhere to privacy principles in all circumstances, and they are not subject to regulatory oversight for their privacy or information security practices. OVIC's position is that legislative change is required to provide greater coverage of the PDP Act to courts and tribunals. OVIC recommends that:

- section 10 of the PDP Act is repealed
- a new section be inserted into section 15 of the PDP Act, setting out that courts and tribunals are exempt from the IPPs only where they are acting in their judicial capacity
- section 84 is amended to include courts and Court Services Victoria in the application of Part 4 for the handling of public sector information under the VPDSS.

The removal of section 10 and insertion of a new section into section 15 would have the effect of creating a 'functional' approach to the privacy exemption for courts and tribunals, providing an exemption only 'when' a court or tribunal is 'acting in a judicial capacity', instead of the current broad wording of section 10 which excludes information handling 'which relates to' judicial or quasi-judicial functions.

In OVIC's view, the current wording of section 10 is unnecessarily broad as it excludes most decisions and actions undertaken by a court or tribunal, even those that are often far removed from the court or tribunal exercising its judicial or quasi-judicial functions. This results in certain acts or practices which ordinarily would be considered an interference with privacy not being treated as such.

For example, functions such as communication with individuals relating to their future availability to attend proceedings, CCTV images and footage from security cameras within the court precinct, and information about lawyers (where not connected to a specific case) are likely to be considered 'related to' judicial or quasi-judicial functions.

The privacy laws of other Australian jurisdictions do not have such a broadly worded exemption. For example, section 6 of the *Privacy and Personal Information Protection Act 1998* (NSW) uses narrower wording, and there is no equivalent exemption in the *Privacy Act 1988* (Cth). Further, several nations

[12] Consultation paper, page 89, paragraph 7.16.

subject to the European Union's General Data Protection Regulation (GDPR)[13] that have an exemption for courts and tribunals limit its application to a 'functional' approach determined by 'when' a court or tribunal is 'acting in their judicial capacity.'

OVIC recommends adopting this functional approach, which differentiates between the functions undertaken by a court or tribunal for its operations where the processing of personal information in a judicial capacity is required (for example, verdicts or decisions and civil and criminal proceedings), and where a court or tribunal is undertaking functions to process personal information outside of its judicial capacity, such as registry officers dealing with personal information.

Adopting this functional approach will mean that OVIC can assist individuals with privacy complaints in more instances. Currently, because of the broad construction of section 10 of the PDP Act, OVIC is limited in its ability to conciliate the privacy complaints it receives arising from courts and tribunals.

With respect to OVIC's recommendation to amend section 84, OVIC does not believe that extending coverage of Part 4 to courts and tribunals would challenge the independence of the judiciary in making decisions in its judicial capacity. Instead, it would ensure the information handling practices and processes relied on by courts and tribunals are consistent with those which are expected of other Victorian public sector organisations.

In the absence of expanded coverage for courts and tribunals under the PDP Act, for reasons of judicial independence or otherwise, it will continue to be up to the courts and tribunals to self-regulate their information handling practices. As it relates to AI, courts and tribunals need to ensure staff are adequately trained and appropriate governance structures are put in place, including policies, guidelines and procedures, to safely manage the usage of AI systems. Judicial independence does not negate the need for transparency and accountability of AI usage within the courts.

What is AI?

This section considers the definition of AI, and the differences in Generative AI tools available.

OVIC's submission focuses on Generative AI in order to provide the VLRC with some important distinctions on how privacy and information security are affected by the different environments in which Generative AI can be used. OVIC acknowledges that there are other AI technologies that will also be relevant to courts and tribunals, as outlined in the consultation paper on pages 9-12.

Definition of AI

The consultation paper adopts the OECD definition of AI.[14] OVIC agrees with the use of this definition, as it provides an accurate description of the technology. It acknowledges that AI makes predictions, and different technologies vary in their autonomy.

[13] Austria, Croatia, Cyprus, the Czech Republic, Estonia, France, Germany, Greece, Ireland, Liechtenstein, Lithuania, Luxemburg, the Netherlands, Poland, Portugal, Slovenia, Spain, and Sweden take the functional approach to courts and tribunals' application.
[14] Consultation paper, page 9, paragraph 2.5.

OVIC cautions against the use of definitions that attribute human-like features to AI systems, and anthropomorphise what they are able to do. Such definitions may impact how humans interact with AI systems, the expectations they place on them, and their potential reliance on outputs to make or inform decisions. AI technologies do not have cognitive abilities, and cannot be relied upon as a source of truth.

When discussing AI, the kinds of tools that are in discussion are important to be aware of. Tools developed through machine learning that have been refined for specific purpose use (for example, software that has been trained to optimise energy use in a specific kind of building) are very different in scope and application than general-purpose tools that are applied in multiple contexts and information domain (for example, a Generative Pre-trained Large Language Model). The latter, by its nature, is prone to errors when it is prompted to generate information not previously encountered in its data set. Such errors are commonly referred to as 'hallucinations' by their developers. Such errors, outside of clearly defined circumstances, are impossible to eradicate. The probabilistic nature of Large Language Model outputs means they are inevitable.

Publicly available vs enterprise Generative AI tools

There are different settings in which Generative AI tools can be accessed and used by organisations, including courts and tribunals. The risks involved with each differ. For example, publicly available Generative AI tools (non-tenanted) tend to carry higher privacy and information security risks. Tools that are used at an enterprise level, that may be customised to suit an organisation's needs, trained on internal data, and securely managed by the organisation, are likely to be more privacy-enhancing. Enterprise tools that are securely managed by an organisation do not carry the same level of risk of unauthorised access, use and disclosure as publicly available tools.

Enterprise tools are approved and managed by organisations and operate in secured environments. Organisations have full control over what and how information is used and disclosed. Publicly available tools, such as a personal ChatGPT account, are used outside of a secured environment. Given the wide range of free and open web and app based Generative AI tools, organisations may not be aware of if and how their staff are accessing such tools, and what information is being submitted. This is a risk that needs to be managed appropriately by the organisation.

While enterprise tools based on internal data are in some senses better than publicly available tools, they are nevertheless usually based on proprietary models that form the basis of the 'pre-training' the model uses. Where these models are not transparent, the degree of risk associated with their use is impossible to determine, even by a skilled data scientist.

It is OVIC's position that public sector information that is not already published should not be entered into publicly available Generative AI tools.[15]

OVIC's expectations for Victorian public sector organisations in managing the risks of Generative AI tools is discussed later in this submission.

[15] See Public statement: Use of personal information with ChatGPT.

Privacy and information security risks for courts and tribunals

This section sets out the key privacy and information security risks of AI tools. This section is relevant to ensuring that courts and tribunals comply with their obligations when using AI for administrative functions, and can be used to inform the design and development of appropriate governance arrangements for its use in judicial and quasi-judicial functions. This section responds generally to Chapters 3 and 4 of the consultation paper.

Privacy and information security risks associated with AI

OVIC considers that there are significant privacy and information security risks associated with the use of AI, that are important to manage in the context of courts and tribunals. Many of these risks are not new, but are exacerbated by the use of these technologies due to the unprecedented speed and scale with which these technologies can analyse information and produce outputs and decisions that cause harm.

Some of these risks, and the relevant IPPs, are outlined below:

- Collection of personal information: AI can generate or infer new personal information about individuals. This constitutes a new 'collection' by the AI system of personal information, which may not be necessary, lawful, or fair, and may result in inaccurate information or opinions being generated and subsequently used or disclosed (see IPPs 1.1, 1.2, 3.1 and 10).

- Transparency: individuals may not understand how their personal information is collected, used and disclosed by AI systems. Transparency and explainability can be difficult to achieve with AI. This limits individuals' ability to make choices about how their personal information is collected, used and disclosed (see IPPs 1.3 and 5).

- Use of personal information: personal information can be used for secondary purposes, that the individual is not aware of (see IPP 2). For example, individuals may not reasonably expect the courts to use their personal information to train AI systems or for decision-making purposes.

- Disclosure of personal information: personal information can be easily shared with third parties that the individual is not aware of. For example, the company that supplies the AI technology and any other parties with whom it shares information (see IPP 2).

- Consent: it is difficult for individuals to meaningfully consent to the collection and use of their personal information in an AI system.[16] For consent to be valid, individuals must be informed about how their personal information will be collected, used and disclosed. It can be difficult for individuals to understand how AI technologies work, to enable them to make choices about the handling of their personal information.[17]

[16] See the Guidelines to the IPPs for discussion on the elements of consent.
[17] Consultation paper, page 12, paragraph 2.19.

- **Accuracy of personal information**: it can be difficult to ensure the accuracy of personal information, for both that which is entered into AI systems and generated by AI (see IPP 3). When it comes to Generative AI, these tools are not a reliable source of accurate information – they simply make predictions based on the prompt they receive and the training they've had. They may confidently confabulate when they encounter concepts not in their training data. They are not accurate at summarising complex information, such as that which would be relevant in the judicial system.[18]

- **Protection of personal information**: confidentiality of personal information or other public sector information may be compromised if entered into AI systems, if it is subject to unauthorised disclosure (see IPP 4.1).

- **Transborder data flows**: depending on where data used in AI systems is stored, personal information may travel outside Victoria and not be covered by similar protections provided by the IPPs (see IPP 9).

High-risk use of AI systems

The consultation paper considers whether there are some uses of AI that should be considered high-risk, or prohibited altogether.

OVIC holds the view that when the use of an AI system or technology is categorised as high-risk, the use should be prohibited.[19]

The EU Artificial Intelligence Act categorised several types of AI systems as prohibited, including:

- subliminal, manipulative or deceptive techniques to maliciously influence behaviour and decision-making

- exploiting vulnerabilities due to a person's age, disability, social or economic status

- biometric categorisation systems

- social scoring i.e., evaluating or classifying individuals or groups based on social behaviour or personal traits

- assessing the risk of an individual committing criminal offenses

- compiling facial recognition databases by untargeted scraping of facial images from the internet or CCTV footage

- inferring emotions in workplace or educational institutions

[18] See 'AI worse than humans in every way at summarising information, government trial finds', Crikey, September 2024; 'When ChatGPT summarises, it actually does nothing of the kind', R&A IT Strategy & Architecture, May 2024.
[19] See OVIC's submission to the Department of Industry, Science and Resources on mandatory guardrails for AI in high-risk settings, page 1.

- 'real-time' remote biometric identification in publicly accessible spaces for law enforcement.[20]

The use of ChatGPT by a Child Protection worker

In October 2024, OVIC published an investigation report into the use of ChatGPT in child protection settings. The investigation found that personal and delicate information[21] had been entered into ChatGPT, resulting in the breach of two IPPs.

The relevant Child Protection Worker used ChatGPT to assist in preparing a report for the Children's Court. The report contained inaccurate statements that downplayed the severity of the harms the child could have faced, which could have put the child in danger. The investigation found that the Department of Families, Fairness and Housing (DFFH) failed to take reasonable steps to ensure the accuracy of personal information under IPP 3.1, and to protect personal information from unauthorised disclosure under IPP 4.1.

In a compliance notice, OVIC prohibited DFFH from using Generative AI tools in child protection matters, noting that 'child protection, by its nature, requires the very highest standards of care.'[22] In OVIC's view, child protection is a setting that carries such a high risk to the individuals involved, that the risk cannot be adequately mitigated through measures such as organisational policies and training.

The use of AI in decision-making

Similarly, OVIC recommends that the use of AI tools should be prohibited in the courts and tribunals for decision-making, or for providing material to be used in arriving at a decision. As the consultation paper notes, 'some AI systems and models can simulate legal reasoning, but none can exercise reason.'[23] The outputs of AI models should not be confused with human intelligence, and the skill that judicial officers and legal professionals have. To do so could lead to significant harm to individuals and communities, particularly those who are disenfranchised and vulnerable.

The consultation paper also notes that AI may provide opportunities to automate repetitive administrative tasks that involve little discretion.[24] While this may be a low-risk use of AI in many cases, OVIC cautions against using AI for tasks that require discretion, legal reasoning, or tasks that will inform an outcome that affects an individual.

Given the significant accuracy issues with the use of Generative AI, and issues with being able to explain how an AI tool produces specific outputs or reaches conclusions, OVIC considers that its use to inform decision-making, such as writing a judgement, should be prohibited.

[20] See https://artificialintelligenceact.eu/high-level-summary/.
[21] The term 'delicate information' is used in place of what could, in common usage, be described as 'sensitive information'. This is because sensitive information has a specific definition under the PDP Act. What individuals may think of as information that is sensitive to them (for example, information they regard as embarrassing or secret), may not fall within that definition. The term 'delicate information' is used to refer to such information.
[22] See Investigation into the use of ChatGPT by a Child Protection worker, page 30.
[23] Consultation paper, page 15, paragraph 2.32.
[24] Consultation paper, page 19, paragraph 3.4.

In addition to the above examples, the VLRC may wish to consider whether there are other settings in which the use of Generative AI (or AI more broadly) would be deemed as a high-risk use case for courts and tribunals.

Managing privacy and information security risks

This section considers steps that organisations, including courts and tribunals, can take to manage the privacy and information security risks that AI poses. It looks at OVIC's expectations to minimise the risks involved with Generative AI tools, and other important measures such as education and training, AI assessment frameworks, human oversight, business continuity, accountability and transparency, and access to information.

Publicly available Generative AI tools

In February 2024, OVIC issued a public statement on the use of personal information with ChatGPT.[25] The statement addresses the privacy and information security risks of using ChatGPT, but can also apply to other publicly available Generative AI tools.

OVIC makes clear in this statement that:

- Personal information should not be entered into publicly available tools.
- Public sector information (that is not already publicly available) should not be entered into publicly available tools.
- If personal information or public sector information is entered into publicly available tools, organisations should treat it as an information security incident and notify OVIC immediately.
- Publicly available tools must not be used to formulate decisions, undertake assessments, or be used for other administrative actions that may have consequences for individuals. For example, evaluations, assessments, or reviews.
- Using personal information with publicly available tools means disclosing that information to the company that owns the tool. The information may then be subsequently used or accessed for unauthorised purposes by individuals outside of the organisation or outside of Victoria.
- Personal information entered into a publicly available tool may be irretrievable and indefinitely retained by the company that owns the tool.

Enterprise Generative AI tools

OVIC's public statement on the use of Microsoft 365 Copilot in the Victorian public sector sets out OVIC's minimum expectations for organisations before adopting an enterprise Generative AI tool, such

[25] See Public statement: Use of personal information with ChatGPT.

as Microsoft Copilot.[26] The guidance in this statement also applies to other enterprise Generative AI tools.

OVIC's expectations include that organisations should:

- Assess the maturity of their existing information security program prior to signing up for the integration of Copilot. This includes:
    - identifying existing information holdings and systems that may be impacted by the introduction of Copilot, including consideration of their security value
    - considering how any newly generated information by Copilot will be assessed, valued and securely managed, including applying appropriate protective markings
    - conducting an updated information security risk assessment considering the integration of Copilot
    - implementing updated treatment plans by rolling out any new or changed controls
    - implementing a formal process for the ongoing monitoring and review of risks and controls, especially critical given the dynamic development and release of enhancements and new features of Copilot.
- Explicitly disable data sharing with Microsoft in platforms such as Power Platform and Dynamics 365.
- Undertake a privacy impact assessment (**PIA**) to understand the ways in which Copilot will be utilised, the risks it presents to the privacy of individuals, and how to mitigate those risks.
- Develop and implement clear guidance and training on the use of Copilot, that:
    - ensures all users are aware of how the tools may generate, collect, use or disclose personal information and public sector information, and understand how to ensure the quality and security of the information
    - ensures personnel and systems do not have the option to send feedback or report content or bugs to Microsoft, as doing so may result in the unauthorised disclosure of public sector information
    - covers prompt engineering, risk assessment and management, human review of generated language, and active monitoring and reviews to identify potential misuse.
- Ensure all users undertake training with a Microsoft Accredited partner prior to accessing and using Microsoft Copilot features.

[26] See Public statement: Use of Microsoft 365 Copilot in the Victorian public sector.

- Develop an incident response plan for dealing with inadvertent disclosures or misuse of information through the use of Copilot.

Education and training

As part of managing the privacy and information security risks associated with AI, OVIC considers that user training is critical.

Organisations should understand the impact of the AI systems they use on people, processes and procedures, as well as unintended or unexpected consequences. Users should be trained to test, validate and question the outputs of AI models. If those using AI within a judicial setting to aid their work are unaware of the risks involved in relying on an output without verification, this can lead to a raft of issues. Issues of accuracy and reliability are then exacerbated when using datasets provided by external vendors.

The consultation paper notes that ongoing education is needed for court staff, administrators, lawyers, judicial decision makers, and the legal profession generally.[27] OVIC agrees with this sentiment. Guidance for self-represented litigants will also be important. The consultation paper highlights that the risk of relying on inaccurate legal advice generated by AI is worse for self-represented litigants, which may result in misguided legal action.[28]

Similarly, judicial officers should also receive tailored education on AI. In addition to receiving general education on the risks and limitations of AI, their training should also cover the matters that the judiciary should be considering, such as deciding whether to accept something into evidence. Factors like who designed the AI tool, the algorithm used, and what data it was trained on, will be relevant when deciding how much weight the judiciary should give to the evidence. Issues such as bias and inaccuracy can be present in the outputs, and should be taken into consideration.

Guidance on the use of Generative AI

OVIC recommends that any policies or guidance issued by courts and tribunals on the use of Generative AI include a distinction between publicly available (non-tenanted) and enterprise (tenanted) tools. The risks involved with each will differ, as outlined above, as will the types of information that will be appropriate to enter into these tools.

The advice of courts and tribunals may also vary for different users, such as lawyers, self-represented litigants, judicial officers and court staff. Lawyers, for example, require specific education and guidance around what they should not enter into a Generative AI tool, such as a client's personal information, legally privileged or confidential information. The consultation paper notes that concerns with the accuracy of outputs are relevant for legal professionals.[29] While OVIC agrees with this, there are also other important factors for the legal profession to consider when contemplating using Generative AI, and in particular, a publicly available tool.

[27] Consultation paper, page 115, paragraph 9.8.
[28] Consultation paper, page 40, paragraph 4.37.
[29] Consultation paper, page 95, paragraph 7.47.

Given the types of information handled by courts and tribunals, OVIC recommends that careful consideration is given to whether it is appropriate for publicly available Generative AI tools to be used in these settings at all. Where personal, sensitive and health information is relevant, OVIC recommends prohibiting their use.

OVIC recommends aligning with existing policies, procedures and guidelines for the use of Generative AI, including the:

- Victorian Government's guidance for the safe and responsible use of generative artificial intelligence in the Victorian public sector[30]
- Commonwealth's interim guidance on government use of public generative AI tools[31]
- Public Record Office Victoria's Artificial Intelligence (AI) Technologies and Recordkeeping Policy.[32]

The courts and tribunals should act consistently with the Victorian Government and Commonwealth guidance on the use of Generative AI, unless a deviation is demonstrated as necessary.

AI assessment framework

The consultation paper seeks stakeholder views on the development of an AI assessment framework, to guide the assessment of the suitability of AI technology in Victorian courts and tribunals.[33] OVIC agrees that an AI assessment framework for courts and tribunals should be developed. As noted above, in developing a framework, OVIC recommends reviewing existing guidance and resources on this topic.[34]

There are a number of elements that OVIC considers important to include in an AI assessment framework. These elements are couched in terms of privacy and information security, given OVIC's remit, however there will also be other relevant components to include.

- PIAs and security risk assessments (SRA): prior to procuring AI tools, courts and tribunals should conduct PIAs and SRAs.[35] These assessments will allow courts and tribunals to identify and mitigate privacy and information security issues before they implement an AI tool. If the risks are too great and it does not appear they can be appropriately managed, courts and tribunals should opt for a different tool (that offers better protections for individuals) or avoid using AI altogether for that particular activity. Legal professionals should also have a professional obligation to undertake PIAs and SRAs on new tools they are considering, prior to procuring and using them in their work.

[30] See Guidance for the safe and responsible use of generative artificial intelligence in the Victorian public sector.
[31] See Interim guidance on government use of public generative AI tools.
[32] See AI Technologies and Recordkeeping Policy.
[33] Consultation paper, page 112, paragraph 8.60.
[34] See National framework for the assurance of artificial intelligence in government; NSW Artificial Intelligence Assessment Framework.
[35] See OVIC's guidance on PIAs and a PIA template.

- Charter of Human Rights impact assessment: courts and tribunals should assess their proposed use of an AI tool against the Charter of Human Rights and Responsibilities. The Charter includes a right to privacy, that is broader than the right to privacy under the PDP Act.[36]

- Due diligence: courts and tribunals will need to undertake due diligence when procuring an AI tool, to ask questions about what foundation model has been used, what is in the training data, whether it was lawfully obtained, and what has been done to treat the training data (such as removing personal information). Courts and tribunals should check which laws the technology provider is subject to and ensure it is to the same standard that the PDP Act provides.

Business continuity

If an AI system is trained to operate parts or whole of a court-related procedure or process, it is possible that over time, the skills and knowledge to complete the process without AI will be lost, or the ability for staff to understand when an AI is in error will be compromised.

Organisations, including courts and tribunals, must develop business continuity and AI disengagement plans in the event of a cyber-attack, privacy or information security incident or other failures, to minimise disruption and ensure that the court or tribunal can continue to maintain services without the use of AI.

Processes should be put in place to enable humans to promptly take over from an AI system and manually complete tasks, to avoid situations where the courts are unable to undertake their functions. This will require court staff to continue to be trained on court processes, so that in the event that an AI system is turned off or compromised, the courts can continue to function with a human-led approach.

Human oversight

OVIC is of the view that human oversight of AI systems will be required. This is part of good AI governance. As the consultation paper notes, 'human oversight should be retained as a check on AI systems to address risks relating to bias, reliability and accuracy.'[37]

Humans are necessary to design, develop and procure AI tools, and it is essential that they understand how they work (noting there are some limitations to the extent we can understand this). Humans are also necessary to validate outputs. As discussed throughout this submission, AI makes predictions, but is not a source of truth. OVIC agrees with the sentiment in the consultation paper that 'AI systems and outputs should be evaluated and tested before use and monitored after implementation, with ongoing assessment.'[38]

[36] *Charter of Human Rights and Responsibilities Act 2006*, section 13.
[37] Consultation paper, page 80, paragraph 6.40.
[38] Consultation paper, page 80, paragraph 6.40.

Accountability and transparency

Courts and tribunals should be aware of, and transparent about, any AI projects and use cases to their stakeholders and the public. When it comes to collecting and using personal information, courts and tribunals must ensure that individuals are provided with timely, accessible and clear information about what information is held about them and how that information will be handled by AI systems.

The consultation paper seeks stakeholder views on whether courts and tribunals should consult with the public before using AI, and disclose to court users when and how AI is used.[39] OVIC strongly supports these transparency measures.

The accountability and responsibility of implementing, approving or managing AI systems should not fall solely on the IT department or equivalent. Given the breadth and scale of AI applications across the whole organisation, it is advisable to nominate the head of the agency as the responsible and accountable officer for the adoption of AI, with a whole-of-organisation approach taken to identifying and managing the risks involved.

Access to information

The community expects government organisations to be transparent and accountable, and to publicly report on their use of AI.[40]

The community should be able to access information about:

- why and how AI technologies are being used by courts and tribunals
- what information the AI algorithm can access, use and infer
- what the effects or outcomes are of the court or tribunal's use of AI technologies.

Courts and tribunals that use AI systems will need to ensure that appropriate documentation is created, to help current and future generations understand the use of these technologies. To do this well, will require robust information management processes and systems that appropriately capture information, and label, categorise, store and preserve it. For example, information about the logic involved in automated decision-making processes, including AI, needs to be documented in a way that is meaningful. If an individual requires a Master's degree to understand this information, then it is meaningful to only 1.2% of the population.[41]

Transparency can be difficult where AI systems are procured from commercial third parties. There have been situations in Victoria and New South Wales where information about a government agency's use of an AI system in the performance of government functions was not accessible under freedom of information laws, because the government agency did not have physical or constructive

[39] Consultation paper, page 108.
[40] See Information Access and Community Attitudes Study (2023), which notes that 82% of respondents agree that organisations should be required to publicly report on their use of AI. 84% agree that public access to government information improves transparency and accountability.
[41] See Australian Government Style Manual, Literacy and access, Reading levels in Australia.

possession of the relevant documents created and held by the private contractors and their sub-contractors.[42]

To address these issues, the Integrity and Oversight Committee of the Victorian Parliament, in its September 2024 report on the operation of the *Freedom of Information Act 1982* (Vic) (**IOC report**), recommended that agencies and Ministers be required to:

- record and proactively disclose and publish information relating to the use of AI and automated decision-making, including a statement of use and a description of the operation of the AI system[43]
- ensure the right of access under the FOI Act to information relating to the use of AI for government decision-making, where government contracted service providers are involved[44]
- include in their procurement contracts with private organisations, a clause prescribing that these agreements are subject to public scrutiny.[45]

OVIC endorses these same recommendations in the context of courts and tribunals. In particular, where courts and tribunals procure and use third party AI systems as part of their administrative functions, there will need to be strong contractual terms and assurance mechanisms to ensure appropriate documentation is created by the contractor and can be accessed by the courts, and in turn the public under the FOI Act.

Principles for responsible and fair use of AI in courts and tribunals

The consultation paper seeks stakeholder feedback on eight principles to guide the safe use of AI in Victoria's courts and tribunals:[46]

- impartiality and fairness
- accountability and independence
- transparency and open justice
- contestability and procedural fairness
- privacy and data security

[42] See the decisions in *EC3 and Department of Jobs, Precincts and Regions* [2022] VICmr 47 and *O'Brien v Secretary, Department Communities and Justice* [2022] NSWCATAD 100, discussed on pages 98-99 of OVIC's submission to the Integrity and Oversight Committee's Report on the Operation of the Freedom of Information Act 1982 (Vic).
[43] Recommendations 40 and 41, IOC report. This recommendation is similar to developments within the Australian Government, requiring organisations to make a public statement outlining their approach to AI adoption. See the Digital Transformation Agency's Policy on the responsible use of AI in government (last updated 10 October 2024).
[44] Recommendation 41, IOC report.
[45] Recommendation 19, IOC report.
[46] Consultation paper, page 75, paragraph 6.4.

- access to justice
- efficiency
- human oversight and monitoring.

OVIC is supportive of these principles. Many of these principles have been discussed in other parts of this submission, in relation to the risks involved with courts and tribunals' use of AI, and the steps that courts and tribunals should take to address those risks.

Guidelines for court and tribunal users

The consultation paper seeks feedback on whether guidelines for court and tribunal users are required in addition to the principles to guide the safe use of AI. OVIC recommends that guidelines are developed to accompany the principles.

Guidelines would be beneficial for a range of reasons, including to expand on the principles and provide further explanation of what is expected, including examples.[47] The guidelines can contain specific dos and don'ts, where there are critical measures courts and tribunals should take, or practices they should avoid. It will be important that guidelines are regularly reviewed, and examples updated to reflect current use cases.

OVIC suggests that many of the issues raised in this submission should be included in guidelines, such as:

- distinguishing between different types of AI (including publicly available vs enterprise Generative AI).
- the privacy and information security risks involved in using AI.
- other risks involved in using AI, such as bias, discrimination, transparency, explainability and other legal issues, including contractual obligations and legal professional privilege.
- how different users can or should not use AI tools (for example, whether it is appropriate to prohibit the use of publicly available Generative AI tools for some court users) the importance of self-represented litigants and legal professionals disclosing to courts and tribunals when AI has been, or is being, used, and the need for the courts to be transparent about their uses of AI to the public.

Summary of recommendations

1. Amend the PDP Act so courts and tribunals are not broadly exempt from the IPPs and VPDSS under section 10. Instead, repeal section 10 and:

[47] OVIC also recommended the publication of guidelines in its submission to the Department of Industry, Science and Resources on mandatory guardrails for AI in high-risk settings (page 2).

- create a new exemption in section 15 of the PDP Act to exempt courts and tribunals from complying with the IPPs when acting in their judicial capacity
- amend section 84 of the PDP Act to include courts and Court Services Victoria in the application of Part 4 for the handling of public sector information under the VPDSS.

2. When the use of an AI system or technology is categorised as high-risk, the use should be prohibited.
3. Consider a prohibition on the use of AI in decision-making by courts and tribunals.
4. Courts and tribunals should have business continuity and AI disengagement plans in place, so that humans can take over tasks from AI systems if required.
5. Education and training programs should be developed for court users, including legal professionals, court staff, judicial officers and self-represented litigants.
6. Any policies or guidance issued by courts and tribunals on the use of Generative AI should include a distinction between publicly available and enterprise tools. Careful consideration should be given to whether it is appropriate for publicly available Generative AI tools to be used in courts and tribunals at all.
7. Record and proactively disclose and publish information relating to the use of AI and automated decision-making, including a statement of use and a description of the operation of the AI system. These statements should include, at a minimum, information about:
    - why and how AI technologies are being used by courts and tribunals
    - what information the AI algorithm can access, use and infer
    - what the effects or outcomes are of the court or tribunal's use of AI technologies.
8. Ensure that courts and tribunals are able to access information about AI systems from third party vendors, contractors and subcontractors, to facilitate the proactive publication of this information.
9. Act consistently with Victorian Government and Commonwealth guidance on the use of AI, unless a deviation is demonstrated as necessary.