information and privacy commission
new south wales

Guide December 2024

A guide to undertaking Privacy Impact Assessments on AI systems and projects

Who is this information for?	NSW public sector staff working with AI
Why is this information important to them?	This guide is intended to support agencies in understanding, assessing and reducing privacy risks in relation to the use of AI systems and projects when undertaking Privacy Impact Assessments (PIAs)

Introduction

Artificial Intelligence (**AI**) solutions offer a range of opportunities for agencies, including productivity, safety and accessibility benefits. To responsibly harness the benefits of AI however, privacy risks must be considered and appropriately managed.

This guide is intended to support agencies in understanding, assessing and reducing privacy risks in relation to the use of AI systems and projects when undertaking Privacy Impact Assessments (**PIAs**). This guide will also support agencies in undertaking privacy related assessments under the NSW AI Assessment Framework (**AIAF**) and the National framework for the assurance of artificial intelligence in government.

This guide builds on and is complementary to the Guide to Privacy Impact Assessments in NSW, to provide more specific guidance on AI-related privacy risks.

When to use this guideline

- When determining if a PIA is necessary
- When determining the likely scope and scale of a PIA
- When undertaking an AI-related PIA
- When periodically reviewing AI-related privacy risks
- When using the AIAF in the course of designing, developing, deploying, procuring, or using systems containing AI components.

What is AI?

The AIAF defines AI as the ability of a computer system to perform tasks that would normally require human intelligence, such as learning, reasoning, and making decisions. AI encompasses various specialised domains that focus on different tasks and includes automation.

Generative AI, Machine Learning (**ML**), Natural Language Processing (**NLP**) and Computer Vision (**CV**) are all kinds of AI, and each can have unique privacy impacts. More information on the kinds of AI and commonly used terms can be found on the Digital NSW website: A common understanding: simplified AI definitions from leading standards.

Examples of ways government agencies might seek to use AI include:

- an AI-powered chatbot on an agency website that visitors can interact with
- a traffic management system that collects or uses vehicle registration data from CCTV footage and/or toll collection systems to automatically issue fines
- a piece of software which uses large amounts of data held by the agency to predict or determine who is eligible for a subsidy and/or to calculate the subsidy they are entitled to

- a piece of software that uses agency records to predict which individuals or businesses are more likely to be non-compliant with certain obligations
- a technology that analyses crowd sentiment in a stadium by using CCTV footage combined with social media data and environmental system data to alert the stadium management to changes in customer sentiment during crowded events.

There are also different methods of deploying AI systems and solutions within an agency, such as by:

- procuring or subscribing to third party AI systems that could be hosted internally or externally and could be off the shelf or modified for the agency's needs
- internally developing a custom system.

Whichever way your agency implements AI, privacy risks must be considered.

The role of the IPC

The IPC provides advice and assistance to agencies about compliance with privacy and access to government information laws. As an independent oversight agency, the IPC and the Privacy Commissioner must maintain independence from government policy decisions and projects.

While it can give general advice, the IPC cannot:

- endorse an initiative or project as privacy compliant
- conduct or write a privacy impact assessment for an agency or entity
- comment on the policy objectives of the program or project.

For more information, see Fact Sheet - The Role of the Privacy Commissioner: Consulting the IPC on Initiatives and Projects.

PIA considerations when assessing AI systems and projects

The PIA process when assessing AI systems and projects

When considering a PIA, you will need to refer to NSW privacy legislation, which includes:

- the Privacy and Personal Information Protection Act 1998 (**PPIP Act**)
- the Health Records and Information Privacy Act 2002 (**HRIP Act**).

What is a PIA?	A PIA is a structured methodology that can help you to identify and minimise privacy risks, and realise privacy improvements, when you are starting a new project or making changes to existing initiatives. A PIA is one way to implement 'privacy by design' in your organisation's practices. It can help you to understand the collection and flows of personal information while building and demonstrating compliance with privacy laws. Refer to the Guide to Privacy Impact Assessments in NSW for more information about PIAs.

The first question to ask when assessing whether a PIA is needed is, "Will any personal or health information be collected, stored, used or disclosed in the project?" This question is equally important when assessing whether a PIA is needed on an AI system. In fact, the use of an AI system could mean there is an elevated risk associated with the collection, storage, use or disclosure of personal or health information.

If an AI system or project involves handling personal information, a PIA will typically be required. However, whether a PIA is needed is ultimately a risk-based decision that should be made on a case-by-case basis. The cost or size of a project or system is not a reliable indicator of whether a PIA should be conducted, as even low-cost or small-scale projects may have privacy impacts. Agencies should not consider a PIA to be a compliance checklist, but rather an opportunity to consider and assess the way a system or project could affect the privacy of individuals.

The PPIP Act and the HRIP Act describe the principles that govern how agencies must handle personal information and health information – the PPIP Act sets out 12 Information Protection Principles (**IPPs**) and the HRIP Act sets out 15 Health Privacy Principles (**HPPs**). The IPPs and the HPPs are drafted in a technology neutral way, which means that they can be applied to physical and electronic records as well as existing and emerging technologies.

The PIA process helps to identify and manage the privacy risks that may arise from using AI systems and projects that involve personal and health information. The IPC considers that the PIA process is suitable for this purpose, since the PPIP Act and HRIP Act already set out the rules for how to handle such information.

Below, we have provided additional considerations which could be relevant to your PIA when assessing AI systems and projects.

A PIA is a dynamic tool that should be regularly reviewed and updated throughout the development and implementation of an AI system or project. This is because the privacy risks may change as the AI system or project changes its inputs, outputs and impacts. Frequent reviews and updates to PIAs on AI systems and projects will help to keep track of, and address, these changes. It is advisable to conduct fresh PIAs (or PIA updates) at decision points such as before a purchase, before a system is launched, when a system is being reconfigured, when a new use case is being considered, or when a vendor contract is due for renewal. Agencies should ensure each use case being contemplated is assessed on its own merits against the IPPs and HPPs.

The PIA is also an opportunity to engage with key stakeholders that may include the system or project's users, the wider public, and people across both government and non-government sectors and those who have expertise in privacy and human rights.

Considerations when assessing AI systems and projects

What is personal information?

'Personal information' is information or an opinion (which can be part of a database) about an individual whose identity is apparent or can be reasonably ascertained from the information or opinion. Personal information can include fingerprints, retina prints, body samples or genetic characteristics.

What is health information?

'Health information' is a specific type of 'personal information' which may include information about physical or mental health or disability. It includes: personal information an individual provides to any health organisation, personal information about a health service requested by or provided to an individual, some personal information for organ donation, and some genetic information.

Is AI-related data really personal information?

AI systems and projects could involve data that, at first glance, may not appear to be personal information, such as randomly assigned identifiers that distinguish individuals from each other but do not include attributes such as a name, address or driver licence number. Agencies should be aware that data of this kind could still be considered personal information if a person's identity can be reasonably ascertained by referring to other data sources – even if the agency doesn't have any specific intention to do such an identification.

Because personal information can include an opinion about an individual, AI-generated inferences about individuals are also considered personal information, even if they are incorrect.

See: Fact Sheet - Reasonably Ascertainable Identity and Fact Sheet - De-identification of personal information

What makes AI different from other privacy impacting technologies?	The IPPs and the HPPs are applicable to the activities of agencies involving personal information, whether the activity involves AI or not. AI is a broad label and there can be differing views on what constitutes AI. So the question might arise: “Why does AI require special treatment when it comes to privacy?” For the purposes of privacy, what is important is that substantive privacy impacts on the community are assessed irrespective of the label applied to the technology. However, in practice AI can enable the processing and production of vast quantities of data on a scale which is uncommon among other technology solutions. AI can also be readily used to make or guide decision-making processes which have a profound impact on individuals. For these reasons, the impact and likelihood of traditional privacy risks can be greatly amplified by the use of AI technology. AI also intersects with other emerging technologies to create novel risks and challenges. These AI privacy risks are part of a range of ethical concerns arising from the technology. Ethical concerns include bias. fairness, transparency and accountability. They will often require careful consideration and for agencies to go beyond merely complying with privacy laws.

Additional considerations when assessing AI systems and projects

In addition to the steps set out in the Guide to Privacy Impact Assessments in NSW, consider the following IPPs or HPPs, questions to ask and some relevant considerations and illustrative examples.

Note there are exceptions throughout the IPP/HPP requirements which are not fully considered here, and the legislation should be consulted further. (In addition to the IPPs/HPPs, see Statutory Guidelines on exemptions)

IPP/HPP reference	Questions to ask	Considerations and impacts
IPP/HPP 1 - Lawful	• Is the personal information that is being collected for the AI system or project directly related to the agency's function or activities and necessary for that purpose? • Does the AI system or project collect personal information beyond what is reasonably necessary for the agency's function or activities? ○ Be sure to review each input. ○ Consider also that the outputs of the AI system or project could be considered a "collection through creation" of personal information. Consider whether those outputs are all reasonably necessary for the agency's functions or activities. • Does the AI system or project clearly comply with the relevant legislation?	AI technology enables both large-scale and novel kinds of data collection. Agencies should ensure the collection of personal information is reasonably necessary and directly related to the agency's functions or activities. AI systems can also create personal information through new insights about individuals. The creation of these new insights can be viewed as a collection of personal information. An example is a mobile fitness device and app that regularly 'creates' personal information about individuals through the monitoring of heart rate, pulse, walking or sleeping patterns. Over a period of time, new insights are created about an individual's likely health outcomes, including the detection and prediction of disease. Accordingly, agencies should ensure any new personal information created is also reasonably necessary and directly related to the function or activity of the agency. If an AI solution can collect personal information that is beyond the scope of the agency's activities, that capability should not be deployed. Configurations need to be carefully considered. Features that are merely interesting or fun should not be used if they involve collecting personal information that is not directly related to the agency's activities. AI systems which are not clearly explainable, or which do not transparently and comprehensively document the basis and evidence for their output, should not be relied upon for automated decision-making processes that affect individuals. Agencies that are seeking to use AI systems to make decisions that will impact individuals should carefully consider the legal basis underpinning the activity before proceeding and should ensure that there are appropriate opportunities for human review throughout the decision-making process.

IPP/HPP reference	Questions to ask	Considerations and impacts
IPP 2/HPP 3 – Direct	• Does the AI system or project collect personal information directly from the individual or seek their authorisation to collect it from other sources? • Consider personal information created through the use of AI systems and projects.	AI technology can involve collecting data from multiple sources – whether public or held by other organisations. Agencies should not collect personal information about individuals from other sources without consent from the individual. Agencies must consider whether the AI system creates personal information, and whether this "collection through creation" is without the user's involvement or with the consent of the individual. Some common ways that AI can indirectly collect personal information is collection of facial images or biometric data from cameras or sensors, or inferences about a person's emotional state derived from their movements, their voice or other physical characteristics. Collecting data from other sources including from public sources like social media posts, or semi-public sources like the ASIC company register which charges a fee, is also an indirect collection of personal information which can be unexpected, unfair or intrusive, and could also result in the collection of unsolicited information. When agencies enter into agreements to collect data from third party organisations, agencies should satisfy themselves that the data was collected from individuals lawfully and that the sharing to the agency is aligned with the original purpose of collection or individuals have consented. In addition, commonly used software applications are integrating AI features to collect user data 'behind the scenes' in a way which is often not obvious to users. Agencies should satisfy themselves when deploying software applications as to whether this is a feature of the software.

IPP/HPP reference	Questions to ask	Considerations and impacts
IPP 3/HPP 4 - Open	• Are individuals notified that their personal information is being collected, the purposes for the collection and the intended recipients of their personal information? • Consider third party providers managing the AI system. • Consider whether the individual is notified of any personal information created through the use of the AI system. • Consider whether personal information is being used to train the AI system, and whether that training is restricted to use on the agency's system or also benefiting the third party provider, and whether that purpose should be notified to the individual at the time of collection. • Is it clear whether the collection is required by law or is voluntary, and any consequences if the individual chooses not to provide the information?	AI technology may involve data collection that is covert or not obvious such as recording behaviour of users on a website or capturing data from a person's voice or movements to ascertain their emotional state. Agencies should ensure individuals are informed in the manner that is most effective in the context of the collection. Third party providers managing AI systems could be recipients of an individual's personal information and should be mentioned in privacy notices in such circumstances. If an agency is using an AI system that will use personal information solely for enhancing the agency's service, a privacy statement or notice should include this. If the personal information is used beyond this, for example to allow the AI vendor to enhance their product, seek consent from affected individuals. Generally, using personal information to train AI systems outside of an agency use case is high risk and should be avoided wherever possible. Agencies need to consider how to effectively communicate to individuals about novel forms of personal information collection. For example, where an AI-powered chatbot is engaging in a human-like conversation to elicit information, it is preferable to have the required notification communicated by the chatbot at the time the conversation is occurring, not posted on a web page which the user may not see.

IPP/HPP reference	Questions to ask	Considerations and impacts
IPP 4/HPP 2 - Relevant	• Is all the personal information collected for the AI system relevant to the purpose of collection? ○ Review each data point/input for relevance. • How are you ensuring the personal information being collected is accurate, complete, up-to-date and not excessive? ○ Consider the source of the personal information being collected. Can it be relied on? ○ Are there measures in place to ensure all required personal information is provided? What are the impacts if the individual provides only some of the required personal information? ○ Will the individual be prompted to provide all the required information? ○ Will the individual be prompted to provide the personal information in an appropriate format? ○ Is old information that may be out of date being used or relied upon?	Using poor quality information such as agency records which are out-of-date could result in unfair or wrong decisions. This is particularly concerning where it could limit an individual's access to a service, opportunity or benefit, or result in a penalty. Some AI tools, such as chatbots, may need to prompt an individual to provide personal information. It is important that this prompting occurs in a way that elicits relevant, accurate, complete and up to date information. When using AI technology, it can be easy to collect more information than is required and automated collection may not be subject to the same quality assurance that applies to personal information collected through other means. Agencies should guard against collecting excessive amount of data just because it is there or easy to do with the use of technology. If an AI tool needs to be fed data held by the agency, it should be limited to what is strictly required for a specific outcome. Agencies should consider whether the use of free text fields is necessary, as there can be a heightened risk of collecting unsolicited personal information. There is less risk in using structured data fields.

IPP/HPP reference	Questions to ask	Considerations and impacts
IPP/HPP 5 - Secure	- Is personal information being stored securely within the AI system? - Who has access to the personal information handled within the AI system? - Is there a risk that an individual may see personal information of another individual when using the AI system? - Have retention periods and appropriate data disposal methods been defined and implemented? - If a third party is handling personal information, what due diligence has been undertaken to ensure the personal information is protected from unauthorised use or disclosure?	Data in an AI tool will need to be securely managed as it would in any other platform or system, on-premise or cloud-hosted. Data retention policies will need to incorporate mandatory minimum retention periods as required by the State Records Act or any other legislative requirements. Contracts should be in place with third-party providers containing binding clauses on data security and compliance with the PPIP Act and the HRIP Act. Contracts should consider and account for obligations under the Mandatory Notification of Data Breach Scheme. Use caution in using AI tools provided by third parties on the basis of standard terms. These terms are unlikely to require third parties to handle personal information in accordance with the PPIP Act and HRIP Act.
IPP/HPP 6 – Transparent	- How is the agency explaining to the person what personal information about them is being stored, why it is being used and any rights they have to access it? - Can you describe what, how and why an individual’s personal information is being used in relation to the AI system?	Agencies should ensure clear information is made obviously and prominently available in a way that is appropriate to the situation, whether it is a website, in a live chat, on a sign in a public place or in a personal email addressed to the individual. Privacy Management Plans should be updated to include information about the use of AI. Individuals should not have to undertake onerous investigations or chase down information to know whether personal information about them is being stored and their rights in relation to that personal information.
IPP/HPP 7 – Accessible	- Are you prepared to meet requests for access to personal information, including insights or inferences, derived from the use of AI systems. - Are you allowing people to access their personal information without excessive delay or expense? o Consider whether the AI system will enable or delay an individual’s access to their personal information.	Agencies should consider whether arrangements with platforms or vendors and the data types and formats they use would allow for an extract of personal information to be provided in the event such a request was received.

IPP/HPP reference	Questions to ask	Considerations and impacts
IPP/HPP 8 - Correct	• How can an individual update, correct or amend their personal information where necessary in relation to the AI system or project? • Can you ensure the AI system or project will make decisions based on the updated personal information?	Agencies should consider whether arrangements with platforms or vendors and the data types and formats would allow for correcting of personal information in the event such a request was received.
IPP/HPP 9 – Accurate	• What measures are in place to ensure the personal information is relevant, accurate, up to date and complete before being used by the agency? ○ Consider whether additional checks are required if relying on AI generated outputs. ○ Should AI generated outputs be limited to specific, low risk use by the agency?	AI technology such as automated decision-making can be used to make decisions or to recommend decisions to agency staff which will affect an individual's access to a service, opportunity or benefit, or result in a penalty. AI may produce outputs that are sufficiently reliable for some purposes (such as recommending a service), but not for other purposes (such as approving an application). Consider the accuracy of the information when deciding how it will be used. Poor system design, or the use of poor-quality information such as historical agency records could result in biased, unfair or wrong decisions. Agencies must ensure training data and systems are reviewed with these risks in mind, especially where the AI system informs or makes decisions. Agencies should ensure that decisions made by an AI tool are explainable. If there is uncertainty about the reasons why the technology is making certain decisions or recommendations, it should not be used. Agencies should ensure that there is human validation of any AI process that uses personal information.

IPP/HPP reference	Questions to ask	Considerations and impacts
IPP/HPP 10 - Limited	• Are you only using the personal information for the purpose it was collected unless the person has given their consent, or the purpose of use is directly related to the purpose for which it was collected, or to prevent or lessen a serious or imminent threat to any person's health or safety? • If using a third-party provider, have you ensured an individual's personal information is not being used by that third party for their own purposes? • When it comes to training AI models have you considered ways this can be done without using personal information?	AI systems can make it easy for agencies to use data for multiple purposes. Agencies should ensure each use case is assessed on its own merits against the IPPs and HPPs. For example, a database of facial images for security passes should not be used for research without a separate assessment. Agencies should refrain from using personal information collected for a specific purpose for a different purpose unless individuals give their consent or an exemption applies. It is also common that AI system providers seek to use their customer's data for their own purposes, such as training their AI models. This is distinct from an agency using personal information for training an AI system exclusively for internal purposes. Agencies should ensure external use of personal information does not occur unless affected individuals give consent or an exemption applies. Generally, using personal information to train AI systems outside of an agency use case is high risk and should be avoided wherever possible. (In addition to the IPPs and HPPs, see Statutory Guidelines on exemptions)

IPP/HPP reference	Questions to ask	Considerations and impacts
IPP 11 – Restricted **HPP 11 - Limited**	- Does the AI system or project disclose personal information to another person or organisation? - If so, ○ Is the disclosure directly related to the purpose the information was collected, and the individual is unlikely to object to the disclosure? or ○ Is the individual reasonably likely to have been aware of this disclosure? or ○ Is the disclosure necessary to prevent or lessen a serious and imminent threat to the life or health of an individual?	Where AI technology vendors, platform providers, or any other external organisation, will have access to personal information being processed by the technology, this could be considered a disclosure, and the agency should ensure it is able to satisfy one or more of the criteria in IPP 11 to permit the disclosure. Where agencies provide their contracted service providers, including AI system vendors, with access to personal information for the sole purpose of carrying out a contracted service on behalf of the agency, this is likely to be a 'use' (see IPP/HPP 10). Agencies should also consider whether staff or users of a system may be able to view information of other individuals and implement measures to prevent this or ensure it occurs only with authorisation from the individual. Agencies should consider what information needs to be provided to individuals who are users of a system and what consents might be needed if they are partnering with a third-party vendor. In the PIA, agencies should document the disclosures which are intended and the rationale for how they comply with IPP 11. For example, AI vendor tech support is provided by a team of five staff who will have access to all data including personal information. This is a directly related purpose to which it may be assumed there is no objection if appropriate contractual obligations, access controls and security measures are in place.

IPP/HPP reference	Questions to ask	Considerations and impacts
IPP 12 – Safeguarded	• Have you ensured personal information relating to an individual's ethnic or racial origin, political opinions, religious or philosophical belief, trade union membership or sexual activities are not being disclosed unless there is a serious and imminent threat to the life or health of an individual? • Is the personal information being disclosed to a Commonwealth agency or a person or body outside of New South Wales? ○ Consider the location of any third-party providers engaged to support the AI system or project.	Agencies should carefully consider whether data they disclose could fall under the categories mentioned in IPP 12. Agencies should refrain from disclosing datasets automatically and should ensure there is human review of data before it is disclosed. Agencies should design systems and processes to remove sensitive information if it is included in chatbot free text fields. Agencies may also consider clear user guidance or notifications to deter users from entering sensitive information that is not required. Facial images may be considered 'sensitive information' because it is potentially personal information about racial or ethnic origin or can indicate a person's religion. Sensitive information points can be inferred from other personal information.
HPP 12 – Not identified	• In the event the AI system or project involves health information, can you achieve your objectives without using unique identifiers?	Agencies should only use unique identifiers such as a Medicare number or a patient number if it is reasonably necessary to carry out the activity efficiently. If a unique identifier is required, consider a randomly assigned number which is used temporarily for a specific purpose and then deleted.
HPP 14 - Controlled	• In the event the AI system or project involves health information, and you are transferring data outside NSW, have you considered if the recipient is bound to standards which are equivalent to the NSW HPPs or if you can obtain individuals' consent?	Agencies should ensure they are transferring health information only in accordance with the criteria in HPP 14. Key considerations are: • whether the recipient is bound to standards which are equivalent to the NSW HPPs, • whether the individual has given consent, or • whether there is a contract which would benefit the individual, it is impracticable to obtain consent, and the individual would be likely to consent.

AI privacy risks and mitigations

Common AI privacy risks

AI has the potential to amplify privacy concerns in its consumption and analysis of personal information. It is important to be aware of the common privacy risks associated with implementing AI systems and ensure these have been considered.

The table below describes some of the most common privacy risks associated with the use of AI and lists common techniques for mitigating those risks. For more information on the mitigation measures, see the following section.

This is a non-exhaustive list of risks and mitigations/controls for reducing those risks.

Risk	Risk Description	Common mitigations
Inaccurate, incomplete or out-of-date information	AI, particularly generative AI, can produce inaccurate, incomplete or out-of-date information. This can arise because of low quality training data. Inaccurate, incomplete or out-of-date information in training data or in outputs may lead to the misrepresentation of an individual. Many AI systems use factual and/or inferred data to generate outcomes or decisions. It is therefore important that data quality is maintained to avoid defective decisions.	• Data minimisation • Expert and legal advice • Automated destruction • Internal policies • Staff training • Human validation • Technical controls • Use of synthetic data • Data quality controls for training data
Bias and discrimination	AI bias and discrimination can occur where the AI is trained on poor quality data, or the AI is learning from and perpetuating systematic social inequalities (such as racial profiling or gender biases). It may not become apparent until a pattern is replicated. An example of this would be an AI system implemented for expediting processes for hiring staff which recommends male applicants over equally qualified female applicants because it was trained using data that wasn't gender balanced. Automated decisions made by AI should be carefully considered, and human oversight and intervention is often needed to monitor for bias and discrimination.	• Expert and legal advice • Internal policies • Staff training • Human validation • Technical controls • Use of synthetic data
Lack of transparency	AI systems can be complex and difficult to understand. Transparency and explainability are important principles to help prevent the risk of bias and discrimination. A decision made by AI and how it was made should be understandable by humans. In addition to transparency about how decisions are made, there should be transparency around how personal information is collected and stored. It can be difficult to obtain informed consent if there is not proper understanding about how the AI system will use the information being collected. If the information being collected is going to be used to further train the AI, consideration should be given as to whether it is really possible to obtain meaningful consent, particularly if the complexities of the AI system make it unclear what the information is being used for.	• Transparency and consent • Expert and legal advice • Internal policies • Use of synthetic data

Risk	Risk Description	Common mitigations
Secondary use of personal information	The potential for repurposing collected information for secondary uses is a significant risk when using AI systems as capabilities improve and the range of AI products and services expand. In addition, AI systems often need large amounts of data for training. It can be tempting to use existing data for this purpose; however, this can result in a misuse of personal information if that personal information was collected for another purpose. Additionally, third party providers may seek to use data collected through the AI system for their own purposes, which may be inconsistent with the initial purpose of collection by an agency. Generally, using personal information to train AI systems outside of an agency use case is high risk and should be avoided wherever possible.	• Staff training • Internal policies • Privacy policies and notices with clear information about AI system uses • Obtaining consent from affected individuals • Data minimisation • Third party due diligence • Contractual controls • Expert and legal advice • Automated destruction • Access restrictions • Technical controls • Use of synthetic data
Unauthorised disclosure	Some AI systems are managed by third party providers, who may gain access to personal information through the provision of their services. Depending on how the personal information is being handled, this may be a disclosure of personal information. There is a risk that these disclosures could be unauthorised where the disclosure isn't directly related to the purpose the personal information was collected or the individual isn't likely to have been aware of the disclosure.	• Data minimisation • Third party due diligence • Contractual controls • Expert and legal advice • Automated destruction • Internal policies • Staff training • Access restrictions • Technical controls • Use of synthetic data
Retention and disposal	AI systems often require significant volumes of data to improve the effectiveness of the system. This can result in extended retention and delayed disposal of personal information, expanding the privacy risk for an agency. Consideration should be given to if and how data can be removed or disposed of if it is held in an AI system.	• Data minimisation • Third party due diligence • Contractual controls • Expert and legal advice • Automated destruction • Use of synthetic data
Re-identification	AI systems can infer information about an individual, and may be able to identify an individual from seemingly de-identified data through the creation of new inferences or attributes. This can also occur where an AI system is able to combine a variety of data from multiple sources, and stitch information about an individual together in a way that makes them identifiable.	• Data minimisation • Third party due diligence • Contractual controls • Expert and legal advice • Automated destruction • Internal policies • Staff training • Access restrictions • Technical controls • Use of synthetic data

Risk	Risk Description	Common mitigations
Third party risk	Whenever a third party handles personal information, there is a risk that the third party will not handle personal information consistently with the requirements of NSW privacy laws and expectations of your agency. This risk can be significant when the third party is an AI system provider with access to significant volumes of personal information.	• Data minimisation • Third party due diligence • Contractual controls • Expert and legal advice • Security reviews • Access restrictions • Use of synthetic data
Data breaches	AI systems often rely on significant amounts of data to train and operate effectively. This can make them targets for malicious actors seeking to gain access to agency data. Additionally, poor system controls and human error can lead to non-malicious data breaches, where data relating to one individual becomes visible to another individual.	• Data minimisation • Third party due diligence • Contractual controls • Expert and legal advice • Security reviews • Automated destruction • Staff training • Access restrictions • Technical controls • Use of synthetic data
Cross border transfers	Many AI system providers operate in jurisdictions outside of NSW and may be subject to different privacy and data protection laws, which could offer less protection than the laws of NSW.	• Third party due diligence • Contractual controls • Expert and legal advice • Transparency and consent • Compliance with all legislative obligations including State Records Act.

Common AI privacy risk mitigations

One of the key challenges for developing and using AI systems in a responsible manner is to identify and mitigate the potential privacy risks that may arise from the use of AI technology. This section provides some examples of common AI privacy risk mitigations that can be applied at different stages of the AI lifecycle.

AI Governance

AI Governance policies, processes and frameworks as part of agencies' overall data governance arrangements establish guardrails for the safe and lawful use of Al and data, and also assign roles, responsibilities and accountability for decision-making.

Data minimisation

At all stages of the AI lifecycle, personal information should be limited to the minimum amount necessary to achieve the required purpose. This includes minimising the personal information used for training the AI system, as well as minimising the potential collection of personal information of users of the AI system, such as through the use of pick lists rather than free text fields where appropriate. Guidance notes encouraging users to limit the personal information provided when interacting with the AI system can also support data minimisation.

Third party due diligence

Conducting thorough privacy and security assessments of third parties can support you in understanding the privacy and security maturity of that third party and identify whether they meet the standards and expectations of your agency when handling personal information.

Contractual controls

Agencies should have clear contractual agreements with third parties, outlining how the third party will protect personal information, limitations on the use and disclosure of the personal information, processes in the event of a data breach, and mechanisms for auditing compliance with the agreement.

Expert and legal advice

Agencies should seek advice from their legal team and/or their privacy officers when developing AI projects and conducting a PIA.

Transparency and consent

Agencies should be transparent about the use of AI systems and disclosures to third parties supporting AI systems. Plain language explanations and just-in-time notifications should be provided to ensure individuals are appropriately informed of how and why their personal information is being handled. In some circumstances, agencies may need to seek the consent of individuals to use AI systems.

Security reviews

When implementing AI systems and projects, agencies should ensure security reviews have been undertaken and security risks are managed to protect the personal information from unauthorised access, disclosure and loss.

Routine destruction

Destruction (deletion) of personal information must occur in accordance with the *State Records Act*. Human oversight and approval should be factored into routine destruction.

Internal policies

Internal policies and procedures to manage the use of AI systems and AI system outputs should be defined, documented and implemented. These policies should consider the ways AI systems and their outputs should and should not be used, considering both the benefits and risks of the AI system. For example, outputs that are appropriate for general insights and analysis may not be appropriate for making decisions about an individual’s eligibility for a service.

Staff training

Staff should be trained on how to responsibly use AI systems, including the collection and use of data and the privacy impacts.

Access restrictions

As with any system, access controls should be implemented so only those with a need to use AI systems or view AI outputs are given those permissions. This includes programming regular review of access permissions and the removal of permissions when staff leave or change roles where the existing permissions are no longer appropriate or required for the role.

Human validation

While AI can be a helpful tool, human validation should be used where appropriate to reduce the likelihood of harms from unmonitored AI systems.

Technical controls

There are a variety of technical controls that can be implemented to reduce privacy risks, such as differential privacy, federated learning, and fully homomorphic encryption. These should be explored and implemented where appropriate to reduce risks.

Use of synthetic data

Synthetic data is artificially generated data that mimics real data and can be used as an alternative to real data. Synthetic data use can have its own risks, and expert advice should be sought to ensure the benefits outweigh the risks.

Monitoring and Assurance

Ongoing monitoring and periodic assurance or audit activities can help agencies detect if any known risks materialise, if mitigations put in place are not effectively managing risks, or if new risks emerge and need to be managed.

Additional resources

More information on how to identify and manage privacy risks related to AI technology is available and includes the following resources.

- NSW AI Assessment Framework: NSW Artificial Intelligence Assessment Framework | Digital NSW
- IPC PIA guidance: Guide to Privacy Impact Assessments in NSW
- Digital NSW AI resources: Artificial Intelligence | Digital NSW
- National framework for the assurance of AI in government: National framework for the assurance of artificial intelligence in government (finance.gov.au)
- Department of Industry, Science and Resources, Australia's AI Ethics Principles: National framework for the assurance of artificial intelligence in government (finance.gov.au)
- Fact Sheet - The Role of the Privacy Commissioner: Consulting the IPC on Initiatives and Projects (nsw.gov.au)
- Artificial Intelligence and Privacy – Issues and Challenges – Office of the Victorian Information Commissioner (ovic.vic.gov.au)
- Automated decision-making in the public sector (NSW Ombudsman)
- Administrative law and automated decision-making (NSW Ombudsman)
- Implementing automated decision-making (NSW Ombudsman)

For more information

Contact the Information and Privacy Commission NSW (IPC):

Freecall:	1800 472 679
Email:	ipcinfo@ipc.nsw.gov.au
Website:	http://www.ipc.nsw.gov.au/

NOTE: The information in this guideline is to be used as a guide only. Legal advice should be sought in relation to individual circumstances.