Australian Government
Office of the Australian Information Commissioner

Checklist

Privacy considerations when developing or training generative AI models

Who is this guidance for?

This guidance is intended for developers of generative artificial intelligence (AI) models or systems who are subject to the Privacy Act. A developer includes any organisation who designs, builds, trains, adapts or combines AI models and applications. This includes adapting through fine-tuning, which refers to modifying a trained AI model (developed by them or someone else) with a smaller, targeted fine-tuning dataset to adapt it to suit more specialised use cases.

Privacy issue

Can obligations to **implement practices, procedures and systems** to ensure APP compliance be met?

Considerations

- Has a privacy by design approach been adopted?
- Has a privacy impact assessment been completed?

Privacy issue

Have reasonable steps been taken to ensure **accuracy** at the planning stage?

Considerations

- Is the training data accurate, factual and up to date considering the purpose it may be used for?
- What impact will the accuracy of the data have on the model's outputs?
- Are the limitations of the model communicated clearly to potential users, for example through the use of appropriate disclaimers?
- Can the model be updated if training data becomes inaccurate or out-of-date?
- Have diverse testing methods or other review tools or controls been implemented to identify and remedy potential accuracy issues?
- Has a system to tag content been implemented (e.g. use of watermarks)?
- Have any other steps to address the risk of inaccuracy been taken?

Privacy issue

Can obligations in relation to **collection** be met?

Considerations

- Is there personal or sensitive information in the data to be collected?
- Could de-identified data be used, or other data minimisation techniques integrated to ensure only necessary information is collected or used?
- Once collected, has the data been reviewed and any unnecessary data been deleted?
- Is there valid consent for the collection of any sensitive information or has sensitive information been deleted?
- Is the means of data collection lawful and fair?
- Could the data be collected directly from individuals rather than from a third party?
- Where data is collected from a third party, has information been sought about the circumstances of the original collection (including notification and transparency measures) and have these circumstances been considered?
- Have assurances been sought from relevant third parties in relation to the provenance of the data and circumstances of collection?
- If scraped data has been used, have measures been undertaken to ensure this method of collection complies with privacy obligations?

Examples of personal information include a person's name, email address and images or videos where a person is identifiable.

Privacy issue

Can obligations in relation to **use and disclosure** be met?

Considerations

- What is the primary purpose the personal information was collected for?
- Where the AI-related use is a secondary purpose, what were the reasonable expectations of the individual at the time of collection?
- What was provided in the APP 5 notification (and privacy policy) to individuals at the time of collection? Have any new notifications been provided?
- Does consent need to be sought for the secondary AI related use, and/or could a meaningful and informed opt-out be provided as an option?

Privacy issue

Have **transparency** obligations been met?

Considerations

- Are the privacy policy and any APP 5 notifications clearly expressed and up-to-date? Do they clearly indicate and explain the use of the data for AI training purposes?
- Have steps been taken to notify affected individuals that their data has been collected?
- Where data scraping has been used and individual notification is not possible, has general notification been considered?

Learn more

The OAIC's **Guidance on privacy and developing and training generative AI models** includes practices that developers that are APP entities must follow in order to comply with their obligations under the Privacy Act as well as good privacy practices for developers when developing and training models.