



Top five takeaways

Privacy considerations when developing or training generative AI models

- 1 Developers must take reasonable steps to ensure accuracy in generative AI models,** commensurate with the likely increased level of risk in an AI context, including through using high quality datasets and undertaking appropriate testing. The use of disclaimers to signal where AI models may require careful consideration and additional safeguards for certain high privacy risk uses may be appropriate.
- 2 Just because data is publicly available or otherwise accessible does not mean it can legally be used to train or fine-tune generative AI models or systems.** Developers must consider whether data they intend to use or collect (including publicly available data) contains personal information, and comply with their privacy obligations. Developers may need to take additional steps (e.g. deleting information) to ensure they are complying with their privacy obligations.
- 3 Developers must take particular care with sensitive information, which generally requires consent to be collected.** Many photographs or recordings of individuals (including artificially generated ones) contain sensitive information and therefore may not be able to be scraped from the web or collected from a third party dataset without establishing consent.
- 4 Where developers are seeking to use personal information that they already hold for the purpose of training an AI model, and this was not a primary purpose of collection, they need to carefully consider their privacy obligations.** If they do not have consent for a secondary, AI-related purpose, they must be able to establish that this secondary use would be reasonably expected by the individual, taking particular account of their expectations at the time of collection, and that it is related (or directly related, for sensitive information) to the primary purpose or purposes (or another exception applies).
- 5 Where a developer cannot clearly establish that a secondary use for an AI-related purpose was within reasonable expectations and related to a primary purpose, to avoid regulatory risk they should seek consent for that use and/or offer individuals a meaningful and informed ability to opt-out of such a use.**



This document is a summary only and should be considered together with the OAIC's Guidance on privacy and developing and training generative AI models