

AI Guardrails for Brokers | Dos and Don'ts



This guide provides best practices on how to use popular AI tools like OpenAI's ChatGPT, Anthropic's Claude, Google's Gemini, and Microsoft's Copilot in a safe and responsible way. It focuses on protecting client information, making sure AI is used properly, and following important rules and guidelines.

Protecting your data and your clients' data



- De-identify and depersonalise client information before inputting into AI.

Replace names, DOBs, addresses, and account numbers with tokens e.g. use **[CLIENT_NAME]** instead of the actual client's name, **[DOB]** for date of birth, or **[ACCOUNT_NO]** for account numbers. *See bad vs good example guide on the following page.*

- Check privacy policies of these services. It is highly recommended to use paid accounts — even better, enterprise or business versions of AI platforms (ChatGPT Enterprise/API, Claude Enterprise, Gemini for Workspace, Copilot for Microsoft 365). These tiers provide stronger privacy and data protections. For more information, consult this [AI Platform Privacy Comparison](#).
- Apply robust security controls: multi-factor authentication (MFA), single sign-on (SSO), role-based access, and audit logging.



- Don't paste raw loan documents, ID scans, or bank statements into AI tools — especially consumer/free accounts.
- Don't assume data is stored in Australia. Most AI platforms process data offshore unless specifically contracted otherwise.
- Don't rely on privacy policies alone. Contracts must cover storage region, retention, sub-processors, and breach notification obligations.
- Don't share AI account logins across your team. Each user should have their own controlled access.

[EXAMPLE] Giving AI tools client information



Good Example

Customer: [Client A]

Age: Late 30s

Location: Metropolitan VIC

Income: \$120,000 per year

Liabilities: \$20,000 unsecured debt

Assets: \$50,000 in savings

Loan Purpose: Investment property purchase

Loan Amount Requested: \$600,000.

Why this is better:

- All personal details removed (no name, DOB, address, bank names).
- The use of nickname ("Client A") maintains relevance to the customer for subsequent prompts
- Financial figures retained exactly, so assessment accuracy is not compromised.
- Age/location generalised just enough to preserve relevance while preventing re-identification.



Bad example

Customer Name: John Smith

Date of Birth: 12/03/1987

Address: 45 Park Street, Melbourne VIC 3000

Income: \$120,000 per year

Liabilities: \$20,000 credit card debt with ANZ

Assets: \$50,000 in Commonwealth Bank savings

Loan Purpose: Investment property in Brunswick

Loan Amount Requested: \$600,000

Why this is bad:

- Personal details (name, DOB, full address) are exposed.
- Unnecessary identifiers (bank names, suburb) increase re-identification risk.
- Personal details aren't needed for the AI's task of producing an assessment summary.

Using AI safely and responsibly



- Maintain human oversight. Always review and sign off on AI outputs before sharing with clients.
- Use AI for nonsensitive tasks (drafting marketing content, summarising internal notes) and keep client critical tasks under tight control. Remember to anonymise anything you input that relates to clients.
- Verify accuracy: AI can “hallucinate”. Double-check calculations, lender policy references, and compliance details.
- Train your team on AI risks — including data leakage or prompt injection:
 - **Data leakage:** sensitive client information can be shared by mistake with AI tools, risking privacy breaches.
 - **Prompt injection:** hidden instructions in text can manipulate AI to produce unintended outputs, possibly sharing confidential info or giving inaccurate advice.
- Look for fairness and inclusivity, don't ignore bias. AI outputs may be skewed and favour certain groups, profiles or products, affecting fairness in recommendations or assessments.
- Consider the implementation of an “AI acceptable use” (first principles) policy within your business.



- Don't let AI make final credit decisions, loan recommendations, or best interest assessments.
 - Ensure fairness in lending decisions. Regulators require evidence of human judgement.
- Don't upload third-party content (e.g. lender documents) without rights to do so.
- Don't connect unapproved plugins, apps, or extensions to your AI tools without risk/security review.

TL;DR | Key awareness points

- Free accounts often use your inputs to train models. Enterprise/business tiers generally do not.
- Data is often processed offshore (commonly in the US). Use providers that offer regional hosting if sovereignty is required.
- Depersonalisation/anonymisation is the safest default: if in doubt, strip it out.
- Don't treat AI as a "black box". You remain accountable for advice, content, and outcomes presented to clients.
- Regularly audit AI activity and vendor contracts to ensure ongoing compliance.
- Align AI use with the Privacy Act 1988 and the Australian Privacy Principles (APPs). Prepare for reforms removing the small business exemption.
- Follow OAIC guidance (Oct 2024): avoid entering personal information into public generative AI tools.
- The Privacy Act will be reformed likely in 2026 to bring stricter obligations and penalties.

