

Safe AI for Small Business

A plain-English security quickstart for teams adopting ChatGPT, Claude, copilots, agents, connectors, and automation without creating a new data leak or permissions problem.

FREE

Digital field guide

Practical. Vendor-neutral. Security-aware.

Educational material only. It does not replace legal, tax, accounting, financial, medical, cybersecurity, or other licensed professional advice. High-consequence decisions require qualified human review.

1. Start with a data rule

Before asking employees to “use AI,” decide what categories of information are allowed in which tools. Convenience is not a data classification policy.

Class	Examples	Default stance
Public	Published website copy, public product info	Generally lowest risk
Internal	Non-public procedures, ordinary internal notes	Use approved business tools
Confidential	Customer data, contracts, pricing, strategy	Restrict tools, access, retention
Highly sensitive	Credentials, secrets, regulated data, MNPI, private keys	Do not paste into unapproved systems; require specific controls

2. Ask these questions about every AI tool

- Is this a consumer or business/enterprise account?
- Can the provider use our inputs to train models?
- How long are prompts, files, and logs retained?
- Who inside our company can access the workspace?
- Can the tool connect to email, drives, calendars, code, CRM, or databases?
- What actions can an agent take without approval?
- How are API keys and secrets stored?
- Can we revoke access quickly?
- Can we audit what the system did?

3. Hallucinations are a control problem

If an AI answer can cause financial loss, a customer promise, a legal obligation, a security change, or an operational action, “the model sounded confident” is not a control. Define what evidence or human approval is required.

4. Connected AI changes the risk

A chatbot that can only draft text is different from an agent that can read email, modify records, run code, send messages, or call APIs. Every new permission expands the failure surface.

- Grant the minimum permissions required.
- Separate test and production environments.
- Require approval for destructive or externally visible actions.
- Treat web pages, emails, documents, and retrieved content as untrusted input.
- Log important actions and keep rollback paths.
- Review generated code before deployment.

5. Starter acceptable-use policy

Employees may: use approved AI tools for drafting, summarization, brainstorming, formatting, research support, and other low-risk work when outputs are reviewed before use.

Employees may not: place credentials, private keys, highly sensitive customer data, protected regulated information, or confidential material into unapproved AI tools.

Human responsibility: employees remain responsible for verifying factual claims, professional advice, financial figures, commitments to customers, security changes, code, and any externally published material.

Escalation: when a use case touches sensitive data, connected actions, legal/compliance obligations, or high-consequence decisions, obtain the appropriate internal approval before proceeding.

6. The five-minute safety check

- What data goes in?
- Where does it go?
- Who can see it?
- What can the AI do?
- How will a human detect a wrong answer or bad action?

If you cannot answer all five, pause the automation and map the controls first.