BASIC AI SECURITY POLICY

A sensible baseline **AI Security Policy** usually covers these areas:

Data protection

- Classify what can and can't go into AI tools. Customer PII, financials, credentials, and trade secrets stay out of public tools unless approved.
- Use enterprise or business tiers that contractually don't train on your data, with defined retention and deletion terms.
- Redact sensitive data or use data-loss-prevention (DLP) tools to catch it before it reaches a prompt.

Access and accounts

- Allow only approved, company-managed AI tools, with single sign-on and multi-factor authentication. Personal accounts shouldn't be used for company work.
- Give people the least access they need. This matters most for AI agents or connectors that can read email, files, or CRMs.
- Keep an inventory of which AI tools are in use, to surface "shadow AI."

Vendor vetting

- Review AI vendors the way you'd review any SaaS vendor: SOC 2 or ISO 27001 reports, where data is stored, subprocessors, and breach-notification terms.

Output and human oversight

- Require human review before AI output goes to customers, into contracts, into code that ships, or into decisions about people such as hiring or credit.
- Treat AI output as unverified. Check facts, citations, and code for errors or vulnerabilities.

AI-specific threats

- Guard against prompt injection, where hidden instructions in documents, emails, or web pages hijack an AI tool. This is especially important for agents that can take actions.
- Don't let AI tools act autonomously on high-risk steps like payments, deletions, or sending external messages without confirmation.

Logging and response

- Log AI usage where possible, and add AI misuse and data leaks to the incident response plan.

People and governance

- Publish a short acceptable-use policy and train staff on it, including AI-enabled phishing and voice or video deepfakes.
- Name an owner for AI policy and review it at least annually. Frameworks like the NIST AI Risk Management Framework and ISO/IEC 42001 are good anchors.

AI Security Checklist

Use this checklist to confirm your organization has the minimum safeguards in place before staff use AI tools for company work. Any unchecked item is a gap to close.

1. Policy, ownership and tools

- A written AI acceptable-use policy is published, and every employee has signed it
- One named person owns AI policy and reviews it at least once a year
- There is a list of approved AI tools, with a set process for requesting new ones
- Staff use company accounts with single sign-on and multi-factor authentication, never personal accounts
- There is an inventory of the AI tools actually in use, including unapproved "shadow AI"

2. Data protection and vendors

- Staff know which data must never go into AI tools: passwords, account numbers, personal data, trade secrets
- Approved tools run on business or enterprise plans that do not train on company data
- Data retention and deletion terms are confirmed in writing for each AI vendor
- Data-loss-prevention (DLP) or redaction controls catch sensitive data before it reaches a prompt
- Each AI vendor has been reviewed for SOC 2 or ISO 27001, data location, subprocessors and breach-notice terms

3. Human review and AI agents

- A person reviews AI output before it reaches customers, contracts, public content or production code
- AI does not make decisions about people (hiring, discipline, pricing, credit) without human review
- AI tools connected to email, files or business systems have only the access each task requires
- AI cannot send external messages, make payments, change permissions or delete data without a person confirming
- Staff are trained to spot prompt injection, where hidden instructions in a document or web page hijack an AI tool

4. Monitoring, incidents and training

- AI tool use on company accounts is logged and reviewed periodically
- The incident response plan covers AI data leaks and misuse, with a named contact and a 24-hour reporting window
- Unusual requests that arrive by voice, video or email are verified through a known channel, as a guard against deepfakes
- All staff complete AI security training when hired and once a year after that
- Safeguards are mapped to a recognized framework, such as the NIST AI Risk Management Framework or ISO/IEC 42001

Reviewed by: _____ **Date:** _____ **Items open:** ____ of 20

These guidelines and checklist were developed with assistance from Claude, an AI model by Anthropic." Reviewed and edited by Richard Morahan Associates.

BASIC AI SECURITY POLICY

These are guidelines. Consult your legal team to accommodate your business environment.