AI Security — Quick Reference Card

Three rules. Five threats. Print it, pin it.

THE DEFENDER'S THREE RULES

1 Never give AI more access than a new hire on Day 1.

Least privilege applies to machines too. Start with read-only. Add permissions only when there is a documented business reason.

2 If you wouldn't email it to a stranger, don't paste it into a prompt.

Treat every AI interaction as semi-public. Customer names, financial data, health records, passwords, API keys — keep them out of prompts.

3 Automate the boring, protect the critical.

AI is brilliant at repetitive tasks. Keep humans in the loop for anything that touches money, health, legal, or reputation. Every AI output that matters should have a human review step.

COMMON THREATS IN PLAIN LANGUAGE

Threat	What It Means	What To Do
Shadow AI	Employees using AI tools that IT doesn't know about	Create an approved tool list. Communicate it.
Data Leakage	Sensitive business data sent to AI providers and stored or used for training	Check every tool's data retention and training policy before use.
Prompt Injection	Malicious input that tricks AI into doing something unintended	Never let AI tools execute actions without human approval.
Credential Exposure	Accidentally pasting passwords, API keys, or tokens into AI chats	Treat every AI prompt like a public post.
Model Hallucination	AI generating confident but incorrect information	Always verify AI outputs against source data before acting.