Neural Profit Engines

OPERATING GUIDE

Secure AI for Small and Medium-Sized Businesses

How to select, install, govern, and operate LLM chatbots, AI agents, and connected AI applications while protecting company and client data

THE CENTRAL OPERATING RULE

An AI system should never receive more data, authority, access, or trust than is necessary for its specific task.

Written for owners, executives, financial professionals, operations leaders, and technology advisers in small and medium-sized businesses.

September 2026

Executive Summary

AI security is achieved by design. AI security isn't achieved by buying a well-known chatbot and telling employees to be careful. The business must control the environment surrounding the model: accounts, data, permissions, connected systems, automated actions, human review, logs, vendors, and incident response.

For most SMBs, the safest starting point is a company-managed business or enterprise AI service using company-owned identities, multifactor authentication, centralized administration, conservative retention, and no use of company content for model training unless explicitly authorized. Sensitive work must not be performed through personal or free consumer accounts.

RECOMMENDED ADOPTION SEQUENCE

Start with public information and read-only assistance. Add internal data only after permissions and retention are tested. Add write access or autonomous actions last, using narrow permissions, transaction limits, human approval, complete logging, and an emergency stop mechanism.

The six-control model

Function	Management question	Minimum SMB outcome
Govern	Who is accountable?	Approved-use policy, owners, inventory, risk decisions.
Identify	What can the AI reach?	Mapped data, users, systems, vendors, and consequences.
Protect	How is access constrained?	MFA, least privilege, encryption, minimization, secure configuration.
Detect	How will misuse be discovered?	Audit logs, alerts, tests, cost and behavior monitoring.
Respond	What happens when something fails?	Kill switch, revocation, evidence preservation, notification procedure.
Recover	How does the business resume safely?	Backups, rollback, tested restoration, non-AI fallback.

1. What the Business Is Actually Installing

"AI application" is an umbrella term. Risk levels change sharply as the system receives company data, connects to other software, or gains authority to act.

System type	Typical capability	Primary security concern
Standalone chatbot	Drafts, summarizes, researches, or analyzes uploaded material.	Unauthorized submission, retention, account compromise, false output, uncontrolled sharing.
Connected chatbot / RAG	Retrieves from company documents, databases, email, CRM, or accounting systems.	Permission errors, cross-client leakage, poisoned documents, indirect prompt injection.
AI agent	Uses tools to send, change, execute, transact, or communicate.	Excessive authority, credential misuse, destructive or fraudulent action, runaway loops.
Embedded AI	Appears inside existing productivity, finance, HR, meeting, or browser software.	Unnoticed activation, unclear data use, shadow integrations, inconsistent governance.
API-built application	Business code sends inputs to a model and consumes outputs.	Secret exposure, insecure output handling, application vulnerabilities, unexpected consumption.
Self-hosted model	Runs on company-controlled hardware or private cloud.	Patching, configuration, supply-chain integrity, monitoring, physical security, availability.

Self-hosting is not automatically safer. It reduces some provider exposure but transfers responsibility for operating-system security, model provenance, patching, authentication, logging, backups, and disaster recovery to the business.

2. Define and Classify the Use Case

Do not begin with the product. Begin with a written use case. Approval should identify the business problem, intended users, required data, systems accessed, outputs produced, actions permitted, accountable owner, expected benefit, foreseeable harm, and required human review.

Risk-tier the application

Tier	Examples	Default treatment
Low	Drafting from public material; formatting; brainstorming.	Approved tool; normal review; no confidential data.
Moderate	Internal summaries, meeting notes, forecasts, document search.	Business account; access control; minimization; logging; source verification.
High	Client advice, HR screening, accounting entries, production code, customer-specific answers.	Formal assessment, qualified review, testing, narrow access, approval gates.
Critical	Payments, legal commitments, safety control, final employment/credit/medical decisions.	Normally prohibit autonomous operation; require specialized controls and executive approval.

Classify the information

Class	Examples	Default rule
Public	Published website, brochure, press release.	Approved tools permitted.
Internal	Procedures, ordinary meeting notes, non-sensitive forecasts.	Company-managed business tools only.
Confidential	Client records, contracts, pricing, financial statements, source code.	Controlled environment; minimum necessary data; explicit use-case approval.
Restricted	Passwords, keys, card data, SSNs, bank credentials, PHI, highly sensitive HR or legal data.	Do not submit unless specifically approved under specialized regulatory and technical controls.

Removing a name does not necessarily anonymize data. Employer, location, job title, transaction details, health information, and unusual facts can identify a person in combination. Aggregation, masking, redaction, and synthetic data should be used whenever they permit the task to succeed.

3. Establish Governance Before Deployment

Assign accountability

Responsibility	Typical SMB owner
Executive accountability	Owner, CEO, COO, or designated executive.
Technical security	Internal IT lead or managed service provider.
Data authorization	Department or client-data owner.
Privacy and contracts	Qualified manager, attorney, or privacy adviser.
Performance and human review	Business-process owner.
Incident coordination	Named security incident lead.

Create an acceptable-use policy

- Identify approved products, accounts, plugins, connectors, and business purposes.
- Define what employees may and may not enter, upload, retrieve, or generate.
- Require verification and qualified human review for material output.
- Prohibit sharing credentials, bypassing controls, or using personal accounts for confidential work.
- Establish rules for customer-facing chatbots, generated code, automated decisions, copyright, and public claims.
- Give employees a fast way to request a safe new use case and report mistakes without concealment.

Maintain an AI system register

The register should record the provider, product, business and technical owners, users, model or version when known, data categories, connectors, permissions, hosting location, retention, training-use setting, contract dates, risk tier, tests, incidents, last review, and retirement status. Include pilots, browser extensions, meeting assistants, and AI features embedded in existing software.

4. Select the Deployment Model

Managed business or enterprise service

This is usually the appropriate SMB starting point. Security requires centralized administration, company identities, MFA, role-based access, user provisioning and deprovisioning, audit logs, configurable retention, clear training terms, encryption, deletion procedures, security documentation, incident notification, and export capability.

API-based application

An API gives the business more control over workflow and interface but also creates software-security obligations. Store keys in a secrets manager; separate development, testing, and production; authenticate every user; restrict network destinations; validate input and output; set rate and cost limits; rotate keys; and avoid recording content unnecessarily sensitive in logs.

Private cloud or self-hosting

Use self-hosting when a documented privacy, latency, customization, or regulatory requirements justifies the operational burden. Evaluate hardware security, model licensing, patching, supply chain, monitoring, availability, disaster recovery, staffing, and the realism of the proposed isolation.

5. Conduct Vendor and Contract Due Diligence

Marketing language such as “enterprise-grade security” is not sufficient. Obtain written answers, verify the configuration for the exact service tier, and place material commitments in the contract.

Questions about data

- Will prompts, files, outputs, feedback, metadata, or logs be used to train, improve, tune, evaluate, or review models?
- Do the rules differ between consumer, business, API, preview, and beta features?
- Can provider personnel or subprocessors access content, and for what purposes?
- How long are content, embeddings, caches, safety records, and backups retained?
- What does deletion mean, and how long until deletion reaches backups?
- Where is data stored and processed? Can the provider change regions or subprocessors?

Evidence to request

- SOC 2 Type II report or appropriate equivalent.
- ISO/IEC 27001 certification and relevant scope.
- Penetration-test or security-assessment summary.
- Secure-development, vulnerability-management, encryption, and incident-response descriptions.
- Subprocessor list, business-continuity information, and breach-notification commitment.

Contract provisions

- Ownership; confidentiality; permitted uses; training restrictions; retention; deletion; and return of data.
- Security obligations; subprocessors; data location; incident notice; cooperation; and audit evidence.
- Liability, indemnification, insurance, regulatory cooperation, portability, termination, and survival of confidentiality.

REGULATED INFORMATION

Health, payment, employment, financial, educational, children's, biometric, and government information may trigger sector-specific duties. For example, a provider handling electronic protected health information may require a HIPAA business associate agreement. Payment-card data should remain outside general-purpose AI tools and inside the applicable PCI DSS environment.

6. Build the Security Architecture

Identity and access

- Use individual company-owned identities; prohibit shared accounts.
- Use single sign-on where practical and require phishing-resistant MFA or passkeys when available.
- Separate administrator accounts from daily accounts.
- Automate or document onboarding, role changes, and immediate termination.
- Review users, administrators, service accounts, connectors, and privileges at least quarterly.

Least privilege and isolation

The AI must never increase a user's effective access. If an employee cannot open an HR folder directly, the retrieval layer must prevent the chatbot from retrieving that folder. A natural-language instruction telling the model not to disclose information is not an authorization control.

- Separate development, testing, and production.

- Separate each client or tenant; test that separation adversarially.
- Separate public knowledge from internal and restricted knowledge.
- Separate read, draft, approve, and execute permissions.
- Use allowlists for tools, commands, data sources, and external destinations.

Encryption and secrets

Use encryption in transit, at rest, on backups, and on company devices. Encryption protects stored and transmitted information, but it does not prevent an authorized provider, compromised account, or overprivileged application from accessing information after decryption.

Passwords, API keys, private keys, tokens, and connection strings must never appear in prompts, system instructions, RAG documents, source repositories, spreadsheets, email, or ordinary application logs. Use a secrets manager, scoped short-lived credentials, rotation, and rapid revocation.

7. Secure RAG and Company Knowledge

Control ingestion

- Verify source and ownership; confirm the business has the right to use the material.
- Scan files for malware and reject unsupported types.
- Classify documents; remove secrets and unnecessary personal information.
- Apply access permissions and record owner, version, effective date, and expiration.
- Detect hidden text, macros, links, or instructions that could manipulate the model.

Control retrieval

- Enforce authorization before retrieval, not after the model generates an answer.
- Limit the number and size of retrieved records and prevent broad wildcard access.
- Preserve client and tenant boundaries in storage, indexes, caches, and evaluation.
- Require source citations and expose document date or version when relevant.
- Ensure deletion and revoked permissions propagate to embeddings, indexes, and caches.

Embeddings are not harmless because they are not human-readable. Protect vectors and metadata according to the sensitivity of the source data.

8. Defend Against AI-Specific Threats

Threat	What can happen	Core controls
Prompt injection	Hostile input causes the model to disregard intended instructions.	Treat all external content as untrusted; isolate instructions; restrict tools; enforce policy outside the model.
Sensitive disclosure	The system exposes secrets, personal data, or another user's records.	Minimize context; authorize before retrieval; segregate tenants; test extraction; block secret patterns.
Improper output handling	Generated text becomes executable code, SQL, HTML, shell input, or a dangerous link.	Treat output as untrusted; validate, sanitize, parameterize, sandbox, and review.
Excessive agency	An agent has more authority than the task requires.	Read-only first; narrow tools; allowlists; transaction limits; human approval; kill switch.
Supply-chain compromise	Malicious model, package, plugin, dataset, or update enters the environment.	Inventory; verify provenance; pin and scan versions; patch; test changes.
Poisoning	Corrupted training or knowledge data changes behavior or decisions.	Controlled ingestion; provenance; versioning; anomaly review; rebuild capability.
Misinformation	Plausible but false output causes a business error.	Source grounding; independent verification; uncertainty behavior; qualified review.
Unbounded consumption	Runaway loops or abuse create outages and cost.	Rate, step, time, context, output, tool-call, and spending limits; anomaly alerts.

System prompts and content filters are useful layers, but neither is a complete security boundary. Authorization, transaction limits, destination restrictions, and approval rules must be implemented outside the model in deterministic software.

9. Install and Configure the System

1. Prepare the foundation. Patch devices and servers; encrypt endpoints; deploy endpoint protection; secure email and identity; configure backups; remove obsolete administrators; and establish incident reporting.
2. Create the company tenant. Register it in the company's name, use company-controlled administrators, configure SSO and MFA, disable public sharing, disable unapproved connectors, select retention and training settings, and enable logging.
3. Define roles. Establish AI administrator, security reviewer, department owner, standard user, restricted-data user, developer, and agent approver roles as needed.
4. Start with safe data. Use public, synthetic, or masked data first. Add internal repositories incrementally after access revocation, deletion, and cross-user isolation have been tested.
5. Configure tools. Document every tool's purpose, users, readable data, writable systems, permitted operations, destinations, transaction limits, approval rule, logs, and shutdown method.
6. Pilot. Use a small group, low-risk task, limited period, known-answer evaluation set, adversarial tests, and read-only access.
7. Approve production. Require evidence of security, accuracy, practical human oversight, logging, safe failure, vendor review, and rapid disablement.

10. Test Before Launch

Functional and business testing

- Correctness, completeness, calculations, sources, dates, names, and handling of missing information.
- Known-answer cases, edge cases, ambiguous requests, and intentionally false premises.
- Time saved, cost per task, correction rate, unsupported-claim rate, and customer complaints.
- Whether the required human review can actually be performed at operating volume.

Security and abuse testing

- Attempt to reveal system instructions, secrets, and another user's or client's records.
- Place malicious instructions in webpages, email, PDFs, spreadsheets, images, and retrieved documents.
- Attempt unauthorized tool use, external transmission, code execution, and privilege expansion.
- Verify that deleted documents and revoked permissions no longer produce results.
- Attempt to exceed budgets, step limits, transaction limits, and denied approvals.
- Confirm that failure is contained, logged, visible, reversible where possible, and recoverable.

11. Operate AI Safely

Make human review real

A review control is ineffective if the reviewer sees only the AI's conclusion or routinely clicks Approve. For a material decision, the reviewer should receive the source evidence, assumptions, proposed action, uncertainty, and consequences.

- Independently recalculate material financial figures.
- Verify important facts using authoritative sources.
- Review accounting entries, tax positions, contracts, legal statements, medical or safety content, pricing, payments, customer commitments, employment decisions, production code, and access changes.
- Record approval when the consequence or regulatory environment requires evidence.

Customer-facing chatbots

- Identify the system as automated and explain its purpose and limitations.
- Do not request unnecessary sensitive information.
- Authenticate users before providing account-specific information.
- Isolate sessions and customer records.
- Provide a clear transfer to a human.
- Rate-limit abuse and constrain company commitments.
- Monitor harmful, discriminatory, deceptive, or consistently inaccurate answers.

Train employees

Training should demonstrate data leakage, hallucinations, false citations, direct and indirect prompt injection, deepfake-enabled fraud, safe document handling, client confidentiality, review obligations, and incident reporting. Train at onboarding, at least annually, and after major changes.

12. Log and Monitor Without Creating a New Privacy Problem

Collect enough evidence to detect abuse and investigate incidents, but avoid retaining complete sensitive conversations when metadata, categories, hashes, or redacted excerpts will serve the purpose.

- Log authentication, administrative changes, model and application use, connectors, tools, repositories accessed, approvals, blocks, exports, sharing, and retention changes.

- Alert on mass retrieval, cross-client attempts, new administrators, new connectors, disabled logging, secret patterns, abnormal cost, large exports, unusual destinations, and out-of-hours actions.
- Encrypt logs, restrict access, synchronize time, protect integrity, define retention, and test alert delivery.

13. Respond to AI Incidents

AI incidents include confidential data entered into an unauthorized tool, compromised accounts, cross-client disclosure, harmful decisions, prompt-injection success, poisoned knowledge, exposed keys, runaway agents, provider breaches, deepfake fraud, and unauthorized public statements.

8. Stop the affected application, agent, or connector.
9. Revoke sessions, keys, tokens, and delegated access.
10. Preserve logs, prompts, files, approvals, transactions, and relevant configuration.
11. Determine affected users, clients, data, systems, actions, and time period.
12. Contain further retrieval, execution, communication, or spending.
13. Notify management, counsel, insurer, provider, clients, and regulators as applicable.
14. Reverse or reconcile harmful transactions and communications.
15. Remove poisoned content and rebuild affected indexes or environments.
16. Restore a known-safe configuration and test it before reactivation.
17. Document cause, impact, decisions, notifications, and corrective actions.

14. Recover, Maintain, and Retire

Business continuity

Do not make the chatbot the sole repository for business procedures, records, custom instructions, evaluation data, configurations, or customer communications. Maintain independent backups and a non-AI method for critical operations such as invoicing, payroll, payment approval, and emergency response.

Change management

Reassess the application when the provider, model, system instructions, data sources, permissions, connectors, automation level, retention, hosting region, subprocessors, or business purpose changes. A provider can materially change system behavior even when the business changes no code.

Secure retirement

- Disable users and service accounts; revoke keys and tokens.
- Remove connectors, webhooks, plugins, and delegated permissions.
- Export required records and preserve legal holds.
- Delete provider data, local caches, embeddings, indexes, and backups according to policy.
- Update the inventory, redirect dependent workflows, stop billing, and monitor residual access.

Point-by-Point Implementation Checklist

Use this checklist for initial approval, production launch, and recurring review. Mark an item complete only when evidence exists.

A. Governance and ownership

☐ An executive is accountable for AI risk.

☐ Every AI application has named business and technical owners.

☐ A written AI acceptable-use policy is approved and distributed.

☐ Approved and prohibited products, accounts, plugins, and uses are identified.

☐ Personal accounts are prohibited for confidential business work.

☐ An inventory covers chatbots, APIs, agents, embedded AI, browser extensions, and pilots.

☐ Each system has a documented purpose, risk tier, review date, and retirement status.

☐ High-impact and regulated uses receive management, legal, and security review.

☐ Shadow-AI discovery and remediation procedures exist.

B. Data protection

☐ Information is classified as public, internal, confidential, or restricted.

☐ Employees receive company-specific classification examples.

☐ Restricted information is blocked unless specifically authorized.

☐ Client contracts permit the proposed processing.

☐ Only the minimum necessary data is submitted.

☐ Synthetic, masked, aggregated, or redacted data is used where practical.

☐ Secrets are removed from prompts, files, code, knowledge bases, and logs.

☐ Client and tenant data is segregated.

☐ Retention and deletion rules cover prompts, files, outputs, embeddings, caches, logs, and backups.

- ☐ Privacy notices accurately describe actual practice.

C. Vendor and contract review

- ☐ The correct business or enterprise service tier is used.
- ☐ Data ownership and permitted uses are clear.
- ☐ Training, improvement, human review, and feedback uses are documented.
- ☐ Opt-out and retention settings are verified in the actual tenant.
- ☐ Subprocessors and data locations are reviewed.
- ☐ Security assessments, certifications, and incident processes are reviewed.
- ☐ Contract terms address security, confidentiality, deletion, incidents, portability, liability, and termination.
- ☐ Required BAAs, DPAs, or industry agreements are executed.
- ☐ Provider terms and material changes are monitored.

D. Identity and access

- ☐ All users have individual company-owned identities.
- ☐ Shared accounts are prohibited.
- ☐ MFA is required; phishing-resistant methods are preferred.
- ☐ SSO is configured where practical.
- ☐ Administrator accounts are separate and limited.
- ☐ Dormant users and unnecessary service accounts are disabled.
- ☐ Departing users are removed immediately.
- ☐ Access, administrators, connectors, and service accounts are reviewed quarterly.
- ☐ API credentials are scoped, stored in a secrets manager, rotated, and rapidly revocable.

E. Secure configuration

- ☐ Company data is excluded from model training unless expressly approved.
- ☐ Retention is set to the shortest operationally appropriate period.
- ☐ Public sharing is disabled by default.
- ☐ Unapproved plugins, extensions, and connectors are blocked.
- ☐ Audit logging and security alerts are enabled.
- ☐ Development, test, and production environments are separated.
- ☐ Devices and servers are patched, encrypted, protected, and backed up.

☐ Configuration is documented and changes require approval.

F. RAG and knowledge systems

☐ Every source has an owner, classification, provenance, version, and expiration.

☐ Files are scanned and suspicious hidden instructions are tested.

☐ The company has the right to use each source.

☐ Authorization is applied before retrieval.

☐ Cross-user and cross-client isolation is tested.

☐ Deletion and access revocation propagate to indexes and caches.

☐ Answers cite source documents and distinguish evidence from interpretation.

☐ Unsupported questions produce uncertainty rather than invention.

☐ Embeddings and metadata receive source-equivalent protection.

G. Agents and automation

☐ Each agent has one defined purpose.

☐ Tools, commands, repositories, and destinations are allowlisted.

☐ Read, draft, approve, and execute privileges are separated.

☐ The agent cannot increase its own authority or alter governing policy.

☐ Transaction and spending limits are enforced outside the model.

☐ Consequential actions require qualified human approval.

☐ Approvers see evidence, assumptions, and the exact proposed action.

☐ Loops have step, time, context, tool-call, and cost limits.

☐ All tool calls and approvals are logged.

☐ A tested kill switch exists.

☐ Actions can be reversed or reconciled where possible.

H. Application security

☐ Input length, file size, and file type are constrained.

☐ External content is treated as untrusted.

☐ Prompt-injection defenses are implemented beyond system prompts.

☐ Model output is treated as untrusted.

☐ Generated HTML is sanitized; code is sandboxed and reviewed.

☐ Database queries are parameterized and narrowly authorized.

- ☐ Generated links and commands are not followed automatically.
- ☐ Rate limits, quotas, and cost alerts are configured.
- ☐ Models, packages, plugins, datasets, and containers are inventoried and verified.

I. Testing and approval

- ☐ Known-answer, edge-case, and missing-information tests are complete.
- ☐ Calculations and material factual claims are independently verified.
- ☐ Hallucination, correction, and unsupported-claim rates are measured.
- ☐ Cross-user, cross-client, secret-extraction, and system-prompt attacks are tested.
- ☐ Direct and indirect prompt injection are tested.
- ☐ Unauthorized tools, destinations, code execution, and privilege changes are tested.
- ☐ Deleted data and revoked access are tested.
- ☐ Runaway consumption and denial-of-service behavior are tested.
- ☐ Failure is contained, visible, logged, and recoverable.
- ☐ Production approval and test evidence are retained.

J. Human oversight and customer protection

- ☐ Employees understand that confident output may still be wrong.
- ☐ Important facts and material calculations require verification.
- ☐ High-impact output receives qualified human review.
- ☐ The AI cannot autonomously make final legal, medical, safety, credit, employment, payment, or access decisions.
- ☐ Public claims are substantiated.
- ☐ Customer-facing AI identifies itself and explains limitations.
- ☐ Account-specific information requires authentication.
- ☐ Customer sessions are isolated and a human escalation path exists.
- ☐ Complaints, harmful outputs, and repeated corrections are monitored.

K. Monitoring and response

- ☐ Authentication, administrative, retrieval, connector, tool, approval, export, and sharing events are logged.
- ☐ Alerts cover mass retrieval, cross-client attempts, secret patterns, abnormal cost, new administrators, new connectors, and disabled logging.
- ☐ Logs are protected, minimized, time-synchronized, and retained according to policy.

- ☐ The incident plan explicitly includes AI systems and providers.
- ☐ Employees know how to report an AI incident.
- ☐ Sessions, credentials, agents, and connectors can be disabled quickly.
- ☐ Evidence preservation and notification responsibilities are documented.
- ☐ Poisoned data can be removed and indexes rebuilt.
- ☐ Erroneous transactions can be reversed or reconciled.
- ☐ Incident exercises and lessons-learned reviews are performed.

L. Continuity, change, and retirement

- ☐ Critical records and procedures exist outside the AI application.
- ☐ Backups and restoration are tested.
- ☐ Critical operations have a non-AI fallback.
- ☐ Provider export and shutdown procedures are tested.
- ☐ Material model, vendor, data, connector, permission, or purpose changes trigger reassessment.
- ☐ Retired systems have users, keys, connectors, data, indexes, and billing removed.
- ☐ Legal holds and required records are preserved.
- ☐ The inventory records final disposition and retirement date.

M. Recurring review schedule

- ☐ Monthly: review alerts, unusual activity, cost anomalies, incidents, and unauthorized connectors.
- ☐ Quarterly: review identities, administrators, service accounts, permissions, deletion, client isolation, performance, and kill switches.
- ☐ Annually and after material change: reassess risk, vendor terms, security evidence, legal requirements, adversarial testing, training, incident response, and recovery.

Implementation Record

Field	Entry
Application / vendor	
Business owner	
Technical owner	

Field	Entry
Approved use case	
Risk tier	
Data classes permitted	
Connected systems	
Production approval date	
Next review date	
Kill-switch owner and method	

Authoritative References

National Institute of Standards and Technology. Cybersecurity Framework 2.0: Small Business Quick-Start Guide. https://www.nist.gov/publications/nist-cybersecurity-framework-20-small-business-quick-start-guide

National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework: Generative AI Profile (NIST AI 600-1). https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=958388

Cybersecurity and Infrastructure Security Agency. Joint Guidance on Deploying AI Systems Securely. https://www.cisa.gov/news-events/alerts/2024/04/15/joint-guidance-deploying-ai-systems-securely

Cybersecurity and Infrastructure Security Agency. Careful Adoption of Agentic AI Services. https://www.cisa.gov/resources-tools/resources/careful-adoption-agentic-ai-services

OWASP GenAI Security Project. OWASP GenAI / LLM Top 10. https://genai.owasp.org/llm-top-10/

OWASP GenAI Security Project. Top 10 for Agentic Applications. https://genai.owasp.org/resource/owasp-top-10-for-agentic-applications-for-2026/

Federal Trade Commission. Privacy and Confidentiality Commitments for AI Companies. https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/01/ai-companies-uphold-your-privacy-confidentiality-commitments

U.S. Department of Health and Human Services. Guidance on HIPAA and Cloud Computing. https://www.hhs.gov/hipaa/for-professionals/special-topics/health-information-technology/cloud-computing/index.html

PCI Security Standards Council. AI and Payments: Pitfalls and Security Risks. https://blog.pcisecuritystandards.org/ai-and-payments-exploring-pitfalls-and-potential-security-risks

IMPORTANT NOTICE

This guide provides general cybersecurity and operational guidance. It is not legal, regulatory, tax, medical, or insurance advice. Requirements vary by jurisdiction, industry, contract, data type, and use case. Engage qualified counsel and security professionals for regulated or high-impact deployments.