CYPROGLOBAL

Vendor-Agnostic Cybersecurity & AI Security Consulting

AI Security Compliance Checklist

68 Actionable Controls Across 10 Critical Domains

Aligned to EU AI Act · NIST AI RMF · OWASP LLM Top 10 · MITRE ATLAS · ISO 42001

EU AI Act | NIST AI RMF | OWASP LLM Top 10 | MITRE ATLAS | ISO 42001

10	68	5
Security Domains	Total Controls	Frameworks

About This Checklist

This AI Security Compliance Checklist is produced by CyproGlobal — a vendor-agnostic cybersecurity and AI security consultancy. It provides 68 actionable security controls organized across 10 critical domains, aligned to the leading AI security frameworks in use in 2026.

Each item is assigned a priority level — Critical, High, or Medium — to help your security team triage and sequence remediation work. Use this checklist during security assessments, compliance reviews, or as an ongoing governance tool. It is designed to be vendor-agnostic: the controls apply regardless of which AI platforms, cloud providers, or security tools your organization uses.

Priority	Definition	Recommended Timeline
● CRITICAL	Poses immediate risk to data, systems, or regulatory compliance	Address within 30 days
● HIGH	Significant risk if unaddressed; impacts governance or security posture	Address within 90 days
● MEDIUM	Improves overall maturity; best practice implementation	Address within 6 months

How to Use This Checklist

For each item, check the box when implemented, mark 'In Progress' when work has started, or leave blank if not yet addressed. Prioritize Critical items first, then High, then Medium. Review quarterly — AI threats and regulations evolve rapidly.

1 AI Asset Discovery & Shadow AI Management

NIST AI RMF · OWASP LLM Top 10

☐	Connect workforce tools (M365/Google Workspace) to discover all OAuth-connected AI apps	CRITICAL
☐	Scan all endpoints for locally installed LLM runtimes, IDE AI extensions, and coding agents	CRITICAL
☐	Enumerate all SaaS platforms for embedded AI features used without IT approval	HIGH
☐	Build and maintain a formal AI Asset Register with owner, purpose, data access, and risk tier	CRITICAL
☐	Identify all shadow AI tools — AI used by employees without security team knowledge	CRITICAL
☐	Classify each AI asset: Vendor-provided / Custom-built / Embedded / Agentic	HIGH
☐	Validate that all discovered AI tools are covered by a Data Processing Agreement (DPA)	HIGH
☐	Schedule quarterly re-discovery scans to catch newly introduced AI tools	MEDIUM

02 LLM Runtime Protection

OWASP LLM-01 · OWASP LLM-02 · OWASP LLM-06

☐ Deploy real-time prompt inspection layer covering all production LLM applications	CRITICAL
☐ Implement jailbreak and prompt injection detection with block/warn/allow enforcement tiers	CRITICAL
☐ Enable PII detection and automatic redaction for all data entering or leaving LLMs	CRITICAL
☐ Configure denied topic policies — block LLMs from discussing restricted business areas	HIGH
☐ Enable hallucination detection using contextual grounding checks on high-stakes outputs	HIGH
☐ Implement output filtering: block harmful content, malicious links, and code injection in responses	HIGH
☐ Establish bi-directional guardrails: enforce policies on both user inputs AND model outputs	HIGH

03 Agentic & Multi-Agent Security

OWASP Agentic Top 10 · MITRE ATLAS

☐ Inventory all autonomous AI agents: SaaS-managed, cloud-hosted, and endpoint-based	CRITICAL
☐ Map agent-to-agent (A2A) communication paths and identify unmonitored interaction chains	CRITICAL
☐ Assess each agent's identity, permissions, and data access footprint (blast radius analysis)	CRITICAL
☐ Implement least-privilege access: agents should have only the permissions required per task	CRITICAL
☐ Deploy stateful threat detection that analyzes full agent interaction chains, not just single prompts	HIGH
☐ Establish human-in-the-loop checkpoints for all high-stakes agentic decisions	CRITICAL
☐ Monitor for agent goal hijacking, memory poisoning, and tool misuse in real time	HIGH
☐ Produce and maintain an AI Bill of Materials (AI BOM) covering all agent dependencies	HIGH

04 MCP Server Security

OWASP LLM-03 · NIST AI RMF

☐	Inventory all Model Context Protocol (MCP) servers — public, private, and shadow	CRITICAL
☐	Scan MCP servers for poisoned tools, schema injection, and malicious capabilities before use	CRITICAL
☐	Enforce transport-layer isolation for all MCP server connections	CRITICAL
☐	Audit MCP server permissions: flag any over-privileged tool capabilities	HIGH
☐	Implement cross-server shadowing detection to prevent tool poisoning across MCP registries	HIGH
☐	Integrate MCP server scanning into CI/CD pipelines before production deployment	MEDIUM

05 AI Supply Chain Security

OWASP LLM-03 · MITRE ATLAS · EU AI Act Art.13

☐	Scan all AI model files (35+ formats) for deserialization attacks and embedded backdoors	CRITICAL
☐	Request Model Cards from all AI vendors documenting training data, limitations, and behavior	HIGH
☐	Review and update all AI vendor contracts with right-to-audit and AI-specific DPA clauses	HIGH
☐	Implement automated vendor risk questionnaires covering training data transparency and security certs	HIGH
☐	Monitor third-party model repositories for tampered or malicious model versions	HIGH
☐	Establish an AI vendor risk tier: Tier 1 annual audit, Tier 2 annual questionnaire	MEDIUM

06 AI Red-Teaming & Adversarial Testing

MITRE ATLAS · OWASP LLM Top 10

☐ Conduct automated red-teaming covering 200+ threat categories before any AI deployment	CRITICAL
☐ Test for OWASP LLM Top 10: prompt injection, insecure output handling, training data poisoning	CRITICAL
☐ Test for OWASP Agentic Top 10: goal hijack, tool misuse, privilege escalation, memory poisoning	CRITICAL
☐ Integrate AI red-teaming into CI/CD pipelines — validate before every model update or release	HIGH
☐ Conduct quarterly manual red-team exercises with an external AI security specialist	HIGH
☐ Test RAG pipelines for poisoned document injection (PoisonedRAG attack patterns)	HIGH
☐ Perform annual full AI red-team exercise with an external firm specializing in AI security	MEDIUM

07 AI Compliance & Governance

EU AI Act · NIST AI RMF · ISO 42001

☐ Classify all AI systems by risk tier: Unacceptable / High-Risk / Limited-Risk / Minimal-Risk	CRITICAL
☐ Conduct Algorithmic Impact Assessments (AIA) for all High-Risk AI systems	CRITICAL
☐ Maintain technical documentation packages for all High-Risk AI systems (EU AI Act Art.11-13)	CRITICAL
☐ Ensure audit logs are immutable and retained for minimum 6 months (EU AI Act Art.12)	CRITICAL
☐ Establish an AI Governance Committee with defined charter, decision rights, and Board reporting	HIGH
☐ Implement mandatory AI Acceptable Use Policy with annual employee attestation	HIGH
☐ Map AI compliance posture across all applicable frameworks simultaneously	HIGH
☐ Deploy continuous compliance monitoring — replace point-in-time audits with 24/7 automated checks	HIGH

08 Cloud AI Security

AWS · Azure · GCP · NIST CSF

☐	Enable native AI security controls on each cloud platform (Bedrock Guardrails / Azure Content Safety / GCP Model Armor)	CRITICAL
☐	Enforce VPC isolation for all AI model inference endpoints — no public exposure	CRITICAL
☐	Apply customer-managed encryption keys (CMK) to all AI model data at rest and in transit	CRITICAL
☐	Configure cross-account AI governance policies to enforce controls across all cloud accounts	HIGH
☐	Review all API gateway configurations for AI endpoints: authentication, rate limiting, logging	HIGH
☐	Implement CSPM tool to continuously scan AI workload configurations for drift	HIGH

09 AI Incident Response

NIST CSF · EU AI Act · MITRE ATLAS

☐	Define AI-specific incident types: prompt injection breach, data exfiltration via LLM, model compromise	CRITICAL
☐	Establish a tiered AI incident severity classification: P1 (regulatory breach) to P4 (minor inaccuracy)	CRITICAL
☐	Automate SIEM alerts for AI events: anomalous query volumes, policy violations, new shadow AI	CRITICAL
☐	Write and test AI incident response runbooks including regulator notification timelines	HIGH
☐	Conduct quarterly tabletop exercises simulating AI-specific incidents	HIGH
☐	Document post-incident review process: root cause analysis and guardrail improvement cycle	MEDIUM

10 Data Privacy & AI

GDPR · CCPA · EU AI Act · HIPAA

☐	Map all personal data flows into and out of every AI system — document what enters and is processed	CRITICAL
☐	Verify that no sensitive data (health, financial, biometric) enters AI systems without authorization	CRITICAL
☐	Ensure Standard Contractual Clauses (SCCs) are in place for EU personal data sent to US AI vendors	CRITICAL
☐	Validate that users are informed when interacting with AI and have exercisable rights	HIGH
☐	Audit AI training data for residual PII — enforce data minimization and purpose limitation	HIGH
☐	Review children's data exposure: ensure COPPA/GDPR-K compliance for all AI systems	HIGH

Need Help Implementing These Controls?

CyproGlobal is a vendor-agnostic cybersecurity and AI security consultancy. Our team of seasoned cyber professionals specializes in implementing every control in this checklist — from AI asset discovery and shadow AI management through full EU AI Act compliance and agentic security governance. We work with any vendor, any cloud, and any stack. Our only allegiance is to your security.

- AI Asset Discovery & Shadow AI Management
- LLM Runtime Protection & Prompt Guardrails
- Agentic AI & Multi-Agent Security (A2A · MCP)
- AI Red-Teaming & Adversarial Testing (200+ categories)
- AI Compliance (EU AI Act · NIST AI RMF · ISO 42001)
- Cloud Security (AWS · Azure · GCP · Multi-Cloud)
- Penetration Testing & Zero-Trust Architecture
- Forward Deployment Security Engineer (FDSE) Engagements

Book a Free Security Assessment **cyproglobal.com**