ISO 27034 for Azure AI Foundry

Application Security Management Process (ASMP) with GitHub Integration
Aligned with OWASP LLM Top 10 2025 & Microsoft Ignite/Universe Security Innovations

Application Security Management Process (ASMP)

Step 1: Application Specification

Define AI agent business context and initial requirements

AI Agent Requirements

- Business objectives & use cases
- Agent personas & capabilities
- Data sources & integration
- Performance requirements

Azure AI Foundry Setup

- Model selection (GPT-4/Llama)
- Prompt flow design
- MCP server configuration
- Knowledge base definition

Regulatory Context

- EU AI Act classification
- GDPR data requirements
- Industry compliance (HIPAA, PCI)
- Geographic restrictions

Step 2: Risk Assessment

Identify threats using STRIDE model + OWASP LLM Top 10

STRIDE for AI Agents

S: Agent identity spoofing

T: Prompt/response tampering

R: Decision audit repudiation

I: Training data disclosure

D: Model availability attacks

E: Privilege escalation

AI-Specific Risk Factors

- Model poisoning & backdoors
- Hallucination impacts
- Prompt injection chains
- MCP server compromises
- Agent-to-agent attacks
- Supply chain vulnerabilities

Risk Scoring Matrix

Critical: Data exfiltration, compromise

High: Business logic bypass, PII

Medium: Performance degradation

Low: UI issues, non-sensitive

Step 3: ANF Creation & Control Selection

Map risks to ASC library controls based on Level of Trust

Levels of Trust (LoT) for AI Agents

Level 0: Untrusted - Public-facing agents
Max restrictions, full monitoring

Level 1: Minimal Trust - Authenticated users
Input validation, rate limiting

Level 2: Standard Trust - Internal employees
RBAC, audit logging, DLP

Level 3: High Trust - Privileged users
MFA, encryption, monitoring

Level 4: Maximum Trust - System agents
Zero-trust, continuous attestation

Control Formula: Base ONF + Risk ASCs + (5 - LoT) × Additional Controls

Step 4: Secure Implementation

Deploy controls via GitHub Actions & Azure AI Foundry

GitHub Security

- Advanced Security scan
- Copilot Autofix
- Secret scanning
- Dependency review
- SARIF integration

Azure AI Controls

- Prompt Shield
- Content Safety
- Responsible AI
- Model access
- Agent monitoring

CI/CD Pipeline

- Pre-commit hooks
- SAST/DAST
- Container scan
- Policy-as-code
- Deploy approvals

✅ Step 5: Verification & Audit

Continuous monitoring, red teaming with PyRIT, compliance validation

Security Monitoring	Red Team (PyRIT)	Compliance
• Sentinel integration	• Prompt injection	• Control metrics
• Defender alerts	• Model extraction	• Audit trails
• Monitor workbooks	• Data poisoning	• Risk register
• KQL queries	• Jailbreak detect	• Vuln SLAs
• Incident response	• Attack chains	• Quarterly review

📚 Application Security Life Cycle Reference Model (ASLCRM)

💾 Requirements	🌳 Design	💻 Development	📊 Testing	🖥️ Deployment	👁️ Operations
• Agent validation	• STRIDE+OWASP	• GitHub Copilot	• Adversarial	• Prompt Shield	• Sentinel
• Data classification	• Architecture	• Pre-commit	• Safety eval	• Content filters	• Incident resp
• Compliance	• Patterns	• SAST scan	• Performance	• Rate limits	• Improvement

⚡ 2025 Microsoft Security Innovations

👥 Agent Identity & Governance

Components: Entra Agent ID, Agent 365
Integration: Step 3 (ANF) - Auth controls

🌳 Build-to-Runtime Security

Components: GitHub Advanced, Defender
Integration: Step 4 - DevSecOps

⚡ AI-Powered Code Security

Components: Copilot Autofix, CodeQL
Coverage: 90%+ JS, TS, Java, Python

🛡️ Advanced Prompt Protection

Components: Prompt Shield, Purview DLP
Features: Real-time detection

📊 Autonomous Security Ops

Components: Security Copilot, Triage
Impact: 6.5x phishing detection

🌐 MCP Security Framework

Components: DevOps MCP, OAuth 2.1
Security: Sandboxed execution

⚠️ OWASP Top 10 for LLM Applications 2025

LLM01: Prompt Injection - Malicious input manipulation

MS: Prompt Shield, Content Safety, Input Filters
Controls: Validation, Template Hardening, Segregation
AI: Injection Detection, Prompt Protection

LLM02: Information Disclosure - Data exposure

MS: Purview DLP, Content Safety, Info Protection
Controls: Classification, Masking, Output Filter
AI: PII Detection, Sanitization, Redaction

LLM03: Supply Chain - Compromised components

MS: Model Catalog, GitHub Security, MCP Gov
Controls: Provenance, Sandboxing, Scanning
AI: Integrity Verification, Server Vetting

LLM04: Model Poisoning - Training manipulation

MS: Safety Evaluations, Defender AI, Drift Monitor
Controls: Validation, Anomaly Detection, Versioning
AI: Attack Detection, Quality Monitor

LLM05: Insecure Output - Insufficient validation

MS: Content Safety, Purview, API Management
Controls: Sanitization, Format Validation, Filtering
AI: Output Validation, XSS/SQLi Prevention

LLM06: Excessive Agency - Unintended actions

MS: Entra Agent ID, Agent 365, Azure Policy
Controls: Least Privilege, Approval Workflows
AI: Permission Management, Scope Limiting

ASMP Integration: Step 2 (Risk) → Step 3 (Controls) → Step 4 (Deploy) → Step 5 (PyRIT Test)

🎯 Microsoft AI Red Teaming & Continuous Verification

Verification by Trust Level:

- **L1-2:** PyRIT automated scans
- **L3:** Quarterly red team + PyRIT
- **L4:** Continuous red team + audit

AI Security Tools:

- **PyRIT:** AI red teaming toolkit
- **Copilot:** Autonomous triage
- **Defender:** Continuous monitor

ISO 27034 for Azure AI Foundry with GitHub

5-Step	6	0-4	10
ASMP Process	ASLCRM Layers	Levels of Trust	OWASP Threats

Framework: Risk-based control selection | Continuous verification | Living ANF documentation

ISO/IEC 27034-1:2011 Application Security Framework

ONF | ANF | ASMP | Azure AI Foundry with GitHub | OWASP LLM Top 10 2025

Version 1.0 | Last Updated: 22 November 2025