perfecXion

AI Agent Security Framework

AI Agent Security Framework

A comprehensive security framework for protecting autonomous AI agents and multi-agent systems.

Framework Components

Agent Architecture Security

- Secure agent communication protocols
- Authentication and authorization mechanisms
- Resource access controls and limits

Behavioral Monitoring

- Real-time agent behavior analysis
- Anomaly detection and alerting
- Performance and security metrics tracking

Risk Assessment

- Agent capability analysis
- Threat modeling for autonomous systems
- Impact assessment for agent actions

Security Controls

Access Controls

- Implement least privilege principles
- Use role-based access control (RBAC)
- Monitor and log all agent activities

Communication Security

- Encrypt agent-to-agent communications
- Implement secure API authentication
- Use secure protocols for data exchange

Resource Management

- Limit agent resource consumption
- Implement rate limiting and quotas
- Monitor resource usage patterns

Incident Response

Develop specific response procedures for AI agent security incidents, including containment, investigation, and recovery processes.