AKYLADE®

AI Security Practitioner™ (A/AISP™) Exam Objectives

Exam Code: AIP-001

AKYLADE

EXAM DETAILS

EXAM NAME	AKYLADE® AI Security Practitioner™ (A/AISP™)
EXAM CODE	AIP-001
QUESTION ITEMS	50 scenario-based, multiple-choice questions
DURATION	120 minutes
PASSING SCORE	700 points on a scale of 100 to 900 points

	DOMAIN	EXAM COVERAGE
 1.0	Ethics and Human-AI Interaction	18%
 2.0	Advanced AI Risk Management	22%
 3.0	AI System Integration and Security	20%
 4.0	NIST AI RMF	24%
 5.0	AI Incident Response and Management	16%

IMPORTANT NOTE:

AKYLADE® continuously reviews and updates exam content to ensure that exams remain current, relevant, and secure. When necessary, AKYLADE® may publish updated exams based on the existing exam objectives, but all related exam preparation materials will remain valid during a given exam's lifecycle. Under each objective, a list of bulleted examples has been provided to aid candidates during their studies. These lists are not exhaustive and serve as examples of technologies, processes, regulations, or tasks related to the relevant objective. Other examples may appear on the exam, even if they are not listed directly under the objective in this document.

DOMAIN 1: Ethics and Human-AI Interaction (18%)

Candidates must be able to analyze ethical considerations, assess AI biases, and implement governance frameworks to ensure responsible AI interactions.

1.1 Given a scenario, develop and implement ethical guidelines for the adoption and use of AI systems.

- Governance
  - AI governance models
  - Accountability frameworks
  - Regulatory compliance
  - Compliance training programs
  - Ethical risk assessments
  - Ethical oversight mechanisms
  - Stakeholder engagement
- Best practices
  - Continuous improvement
  - Data privacy and security
  - Ethical AI design
  - Ethical AI deployment
  - Ethical AI policies
  - Ethical procurement
  - Fairness constraints
  - Sustainability considerations
  - Transparency mechanisms

1.2 Given a scenario, analyze ethical considerations in AI systems and traditional software systems.

- Accessibility
- Accountability
- Autonomy
- Environmental impact
- Human oversight
- Inclusivity
- Informed consent
- Privacy
- Safety
- Security
- Transparency
- Interpretation and resolution of ethical dilemmas

1.3 Given a scenario, assess and mitigate bias in AI systems.

- Identification
  - Automated bias detection tools
  - Manual bias assessment techniques
  - Bias detection and integration in development
- Impact assessment
  - Ethical implications
  - Context-specific analysis
  - Stakeholder impact evaluation
- Mitigation
  - Mitigation strategies
  - Bias correction
  - Mitigation documentation
- Data quality and representation
  - Diverse datasets
  - Representative datasets
  - Data pre-processing
  - Continuous dataset evaluation
- Algorithmic transparency
  - AI models
  - AI decisions
- Stakeholder engagement
  - Feedback on perceived biases
  - Insight integration into mitigation
- Training and education
  - Bias awareness training
  - Continuous learning and development
- Performance metrics
  - Fairness measurements
  - Bias measurements

1.4 Given a scenario, analyze human-AI interactions and their impacts.

- Collaboration strategies
  - Safe strategy development
  - Effective human-AI communication
  - User experience enhancement
- Impact assessment
  - AI decision impact evaluation
  - Short-term impact analysis
  - Long-term impact analysis
  - Benefit and risk balancing
- Human oversight
  - Human oversight integration
  - Human-in-the-loop systems
  - Escalation process for AI decisions
- Stakeholder impact assessment
  - Affected stakeholder identification
  - Stakeholder feedback gathering
  - Stakeholder feedback incorporation
- Trust and transparency
  - AI limitation communication
  - Trust building in AI systems
  - AI decision-making transparency
- User empowerment
  - User education on AI capabilities and limitations
  - Human control over AI interactions and decisions
- Usability and accessibility
  - User-friendly AI interfaces
  - Diverse user group accessibility
  - Usability testing and continuous improvement

DOMAIN 2: Advanced AI Risk Management (22%)

Candidates must be able to develop AI risk management strategies, conduct AI risk assessments, and implement risk mitigation measures.

2.1 Given a scenario, develop and implement an AI risk management strategy for an organization.

- AI risk management integration in organizational governance
  - Frameworks
  - Policies
  - Practices
  - Processes
  - Development guidelines
    - Risk by design
    - Impact on AI system development
  - Continual improvement
- Emerging guidance and standards
  - NIST Artificial Intelligence Risk Management Framework 1.0
  - ISO 27005/27090/27091/42001
  - OECD's AI Risk Evaluation Framework
  - UNESCO's Recommendations on AI Ethics
  - AI Act (Regulation EU 2024/1689)
  - Industry best practices
  - Regulatory requirements
- Governance roles and responsibilities
  - Board of directors
  - Chief executive officer (CEO)
  - Chief risk officer (CRO)
  - Chief information security officer (CISO)
  - Chief data officer (CDO)
  - AI governance committee
  - AI ethics officer
  - AI risk management team
  - Legal and compliance team
  - Internal audit team
  - Project managers
  - Business unit leaders

2.2 Given a scenario, assess the risk associated with an AI system and analyze the results of the risk assessment.

- Phases of an AI risk assessment
  - Risk identification
    - Define objective
    - Define scope
    - Identify risks
    - Conduct threat modeling
  - Risk analysis
    - Analyze risk scenarios
    - Interpret results for strategic decisions
    - Assess vulnerabilities
      - AI model
      - Data
      - Infrastructure
  - Risk evaluation
    - Prioritize risks
    - Compare against risk tolerance
  - Risk mitigation
    - Develop mitigation strategies
    - Deploy measures
    - Implement security controls
  - Documentation and reporting
    - Prepare risk assessment report
    - Communicate results of report
  - Monitor and review
    - Monitor risk environment
    - Consider emerging risks
    - Review and update assessment
- Other types of assessments
  - AI classification assessment
  - AI impact assessment
  - Algorithmic impact assessment
  - Bias assessment

2.3 Given a scenario, perform risk identification and prioritization for threats and vulnerabilities in an AI system.

- Risk identification
  - Identify potential security risks
    - Vulnerabilities in AI models, data, and infrastructure
    - Penetration test results
  - Identify ethical and social risks
    - Bias and fairness issues
    - Privacy and data protection issues
  - Identify operational risks
    - System performance
    - Reliability
    - System integrations
    - Deployments
- AI model risks
  - Bias
  - Hallucinations
  - Model poisoning
- AI prompt usage risks
  - Prompt injection
  - Prompt denial of service
  - Training data exfiltration
- Other AI-specific risks
  - Sensitive data exposure
  - Data leaks into training dataset
  - Non-regulatory compliance
- Threat modeling of AI systems
  - Throughout the AI lifecycle
  - Threat landscape
  - Regulatory changes

2.4 Given a scenario, analyze the impact of AI risks on an organization's assets and determine the appropriate risk response or risk treatment to implement.

- Risk acceptance
- Risk avoidance
- Risk transference
- Risk mitigation
  - Implement security controls
    - Encryption
    - Access controls
    - Secure development
    - Periodic security assessments
    - Security training
  - Address bias and fairness
    - Bias detection
    - Bias mitigation
    - Diverse training data
    - Representative training data
  - Enhance transparency and explainability
    - Use of interpretable models
    - Explainability of AI decisions
    - Transparency in development
  - Strengthen privacy protections
    - Data anonymization
    - Pseudonymization
    - Data protection regulatory compliance
  - Ensure robustness and reliability
    - Rigorous testing
    - Validation of AI models
    - Redundancy implementation
    - Failover mechanism utilization
  - Develop risk management plans
    - Use of frameworks and policies
    - AI risk monitoring and response
  - Conduct continuous monitoring and improvement
    - AI system performance monitoring
    - Review and update of strategies and practices
    - Security control effectiveness evaluation

DOMAIN 3: AI System Integration and Security (20%)

Candidates must be able to evaluate security risks, implement AI security measures, and integrate AI systems securely within an organization's infrastructure.

3.1 Compare and contrast the security needs of AI systems and traditional software technologies.

- AI-specific threats
  - Model poisoning
  - Evasion attacks
  - Adversarial attacks on models
  - Algorithmic attacks
- Data protection
  - Confidentiality
  - Integrity
  - Availability
  - Non-repudiation
  - Authentication
- Vulnerability management
  - Model biases and errors
  - Security flaws in algorithms
- Access controls
  - Role-based access for training and usage
  - Model parameter protection
  - Training data protection
- Monitoring and response
  - AI-driven analytics and monitoring
  - Continuous monitoring of AI models
- Compliance and regulatory requirements
  - AI-specific regulations
  - Ethical guidelines
  - Data protection laws for AI datasets
  - Intellectual property concerns
- Data privacy
  - Differential privacy techniques for AI
  - Personal data in AI training datasets
- Document findings

3.2 Given a scenario, develop and implement a security integration plan for an AI system.

- Gather security requirements
- Conduct risk assessment
- Design the security architecture
  - Recommend security best practices
  - Integrate with existing infrastructure
  - Secure data flow mechanisms
    - Data-at-rest
    - Data-in-processing
    - Data-in-transit
  - Manage user access
  - Implement model access control
- Configure continuous monitoring
- Implement anomaly detection systems
- Develop incident response plans
  - Roles and responsibilities
  - Recovery procedures
    - Model rollback
    - Model retraining
- Comply with regulations
  - Periodic compliance reviews
  - Legal considerations
  - Ethical considerations
  - Responsible AI usage
- Develop integration governance
  - Guidelines
  - Policies
  - Procedures

3.3 Given a scenario, evaluate and mitigate security risks that may occur during the integration of an AI system within an organization.

- Threat modeling
- Deployment challenges
- Security gap analysis
- Impact analysis
- Risk prioritization
- Security control selection
- Organizational measures implementation
- Security monitoring
- Stakeholder involvement
  - Communication
  - Feedback
  - Improvements

3.4 Given a scenario, evaluate and implement strategies for conducting continuous monitoring and improvement of an AI system's security.

- Continuous monitoring
  - Monitoring tools
  - Anomaly detection
  - Security metrics
  - Periodic reporting
  - Incident detection
  - Incident response
- Continuous improvement
  - Gather stakeholder feedback
  - Refine policies and procedures
  - Adopt best practices
  - Adapt to changing regulations

DOMAIN 4: NIST AI RMF (24%)

Candidates must be able to apply the NIST AI Risk Management Framework (AI RMF) to govern, assess, measure, and manage AI-related risks.

4.1 Given a scenario, implement the Govern function's categories and subcategories to conduct AI risk management.

- GOVERN 1 – Policies, processes, procedures, and practices across the organization related to the mapping, measuring, and managing of AI risks are in place, transparent, and implemented effectively
- GOVERN 2 – Accountability structures are in place so that the appropriate teams and individuals are empowered, responsible, and trained for mapping, measuring, and managing AI risks
- GOVERN 3 – Workforce diversity, equity, inclusion, and accessibility are prioritized in the mapping, measuring, and managing of AI risks throughout the lifecycle
- GOVERN 4 – Organizational teams are committed to a culture that considers and communicates AI risks
- GOVERN 5 – Processes are in place for robust engagement with relevant AI actors
- GOVERN 6 – Policies and procedures are in place to address AI risks and benefits arising from third-party software and data, and other supply chain issues

4.2 Given a scenario, implement the Map function's categories and subcategories to conduct AI risk management.

- MAP 1 – Context is established and understood
- MAP 2 – Categorization of AI systems is performed
- MAP 3 – AI capabilities, targeted usage, goals, and expected benefits and costs with appropriate benchmarks are understood
- MAP 4 – Risks and benefits are mapped for all components of the AI systems, including third-party software and data
- MAP 5 – Impacts on individuals, groups, communities, organizations, and society are characterized

4.3 Given a scenario, implement the Measure function's categories and subcategories to conduct AI risk management.

- MEASURE 1 – Appropriate methods and metrics are identified and applied
- MEASURE 2 – AI systems are evaluated for trustworthy characteristics
- MEASURE 3 – Mechanisms for tracking identified AI risks over time are in place
- MEASURE 4 – Feedback about efficacy of measurement is gathered and assessed

4.4 Given a scenario, implement the Manage function's categories and subcategories to conduct AI risk management.

- MANAGE 1 – AI risks based on assessments and other analytical output from the MAP and MEASURE functions are prioritized, responded to, and managed
- MANAGE 2 – Strategies to maximize AI benefits and minimize negative impacts are planned, prepared, implemented, documented, and informed by input from relevant AI actors
- MANAGE 3 – AI risks and benefits from third-party entities are managed
- MANAGE 4 – Risk treatments, including response and recovery, and communication plans for the identified and measured AI risks, are documented and monitored regularly to ensure effectiveness and continual improvement

4.5 Given a scenario, manage AI risks within an organization by creating, implementing, updating, and/or tailoring NIST Artificial Intelligence (AI) Risk Management Framework (RMF) Profiles.

- NIST AI RMF Profiles
  - Creation of a profile
  - Implementation of a profile
  - Updating a profile
  - Tailoring a profile
  - Benefits of using profiles
  - Challenges with using profiles
- Components of an AI RMF Profile
  - Executive summary
  - Background
    - Objective of profile
    - Scope of profile
    - Type of profile
      - Current
      - Target
    - Definitions (optional)
  - Intended audience
  - AI system management and schedule
    - Roles and responsibilities
    - Decision-making authority
    - Lifecycle stages
      - Decision points at each stage
      - Decision authority for each stage
      - Resources for each stage
      - Documentation of each decision point
  - Prioritization methods of subcategories
    - Audiences
    - AI lifecycle stages
    - Mission objectives
    - Risk tolerances
    - Domains
    - Business drivers
    - Manufacturer priorities
    - Security environments
  - Profile use cases
  - Recommended best practices
  - Profile alignment with regulatory requirements

DOMAIN 5: AI Incident Response and Management (16%)

Candidates must be able to plan, execute, and evaluate AI-specific incident response strategies to minimize security risks and operational disruptions.

5.1 Given a scenario, evaluate and select an appropriate incident response methodology for an organization to implement to minimize gaps and challenges in its response efforts.

- NIST incident response framework
- SANS incident response framework
- ISO/IEC 27035 incident management
- FIRST incident response framework
- CISA Cybersecurity Incident and Vulnerability Response Playbook
- CISA AI Cybersecurity Collaboration Playbook
- ISO/IEC JTC 1/SC 42

5.2 Given a scenario, conduct strategic incident response planning and management.

- Align incident response plans with business objectives
- Integrate AI incident response into existing security strategy
- Establish governance and oversight for incident responses
- Develop an AI incident response plan
- Develop continuous improvement plans for incident responses
- Develop a roadmap for adopting AI incident response plans

5.3 Given a scenario, perform incident response actions for an AI system within an organization.

- Preparation
- Detection and analysis
  - Containment, eradication, and recovery
  - Post-incident activity

5.4 Given a scenario, evaluate the effectiveness of an organization's AI incident response efforts.

- Establish metrics for incident response actions
- Evaluate the effectiveness of an AI incident response plan
- Assess the impact of AI incident response actions
- Evaluate the role of AI incident response in organizational security
- Establish mentorship opportunities for junior consultants
- Develop a training program for AI incident response actions
- Implement best practices for managing AI incidents
- Align AI incident response efforts with regulatory requirements