INTELLIGENT SYSTEMS PRESS™
EXECUTIVE ASSESSMENT SERIES

EA-012: AI Security & Operational Resilience Assessment™

Can Your Enterprise Protect, Sustain, and Recover Its AI-Enabled Operations?

Purpose

A practical executive assessment for evaluating how effectively the enterprise protects AI systems and data, anticipates adversarial and operational threats, maintains trustworthy service, responds to incidents, recovers from disruption, and strengthens resilience across the full AI lifecycle.

Educational Executive Self-Assessment | Version 1.0

www.intelligentsystemspress.com

About This Assessment

AI-enabled operations create a broader and faster-moving attack surface. Models, data pipelines, prompts, agents, integrations, infrastructure, users, and third parties can all become points of compromise or failure. This assessment helps leaders determine whether AI security and operational resilience are designed as an integrated enterprise capability rather than handled as isolated technical controls.

How to Complete It

1. Rate each statement	Use the 1-5 scale based on current, observable enterprise practice - not procurement intent, contract language alone, vendor claims, or isolated reviews.
2. Score each domain	Add the five ratings in each domain. Every domain score ranges from 5 to 25.
3. Calculate the total	Add all eight domain scores. The overall AI security and operational resilience score ranges from 40 to 200.
4. Interpret the pattern	Review the total and the distribution. Any domain below 15 represents a material weakness that may expose the enterprise to compromise, disruption, unsafe operation, or prolonged recovery.
5. Select 90-day actions	Choose a limited number of executive actions that strengthen security architecture, threat management, access control, model protection, monitoring, incident response, continuity, recovery, or learning.

Rating Scale

1 Strongly Disagree	Practice is absent or contradicted by current evidence.
2 Disagree	Practice is inconsistent, informal, or limited to isolated areas.
3 Partly Agree	Practice exists in several areas but is not yet enterprise-wide.
4 Agree	Practice is generally reliable, measured, and consistently applied.
5 Strongly Agree	Practice is integrated, evidence-driven, and continuously improved.

Recommended use: Complete individually first, then compare results across business, cybersecurity, technology, data, privacy, risk, compliance, operations, continuity, and internal audit leaders. Differences in ratings are evidence of security or resilience misalignment and should be discussed before averaging scores.

Domain 1: Security Governance & Accountability

Rate each statement from 1 (Strongly Disagree) to 5 (Strongly Agree). Base the rating on repeatable enterprise behavior and current evidence.

Executive leadership has defined accountability for AI security, operational resilience, risk acceptance, incident decisions, and recovery across business and technical functions.					
AI security requirements are integrated into enterprise policies, architecture standards, risk management, procurement, development, deployment, and change-control processes.					
AI systems are classified according to business criticality, data sensitivity, decision impact, autonomy, external exposure, and potential harm from compromise or failure.					
Security and resilience responsibilities are clearly assigned across system owners, model owners, data owners, platform teams, vendors, users, cybersecurity, operations, and continuity leaders.					
Leadership receives regular evidence on material AI threats, vulnerabilities, incidents, control effectiveness, service resilience, recovery readiness, and unresolved risk.					

Domain score: _____ / 25

Evidence or observations:

Domain 2: Threat Modeling & Attack Surface Management

Rate each statement from 1 (Strongly Disagree) to 5 (Strongly Agree). Base the rating on repeatable enterprise behavior and current evidence.

Threat modeling is performed for material AI use cases before deployment and is updated when models, data, agents, integrations, users, or operating conditions change.					
The enterprise evaluates AI-specific threats including prompt injection, data poisoning, model theft, evasion, adversarial inputs, unsafe tool use, excessive agency, and sensitive-data leakage.					
Attack-surface inventories include models, APIs, prompts, plugins, agents, vector stores, data pipelines, infrastructure, identities, endpoints, third parties, and human interaction points.					
Plausible abuse cases, insider threats, supply-chain threats, cascading failures, and misuse of AI-generated content are assessed with business and technical stakeholders.					
Threat findings are translated into prioritized controls, test scenarios, monitoring requirements, response playbooks, architectural changes, and documented risk decisions.					

Domain score: _____ / 25

Evidence or observations:

Domain 3: Identity, Access & Privileged Control

Rate each statement from 1 (Strongly Disagree) to 5 (Strongly Agree). Base the rating on repeatable enterprise behavior and current evidence.

Access to AI models, data, prompts, agents, administrative functions, tools, and production environments follows least-privilege and need-to-know principles.					
Privileged access is strongly authenticated, time-bound where appropriate, monitored, reviewed, and separated from ordinary user access.					
Machine identities, service accounts, API keys, tokens, secrets, and agent credentials are inventoried, protected, rotated, and revoked through controlled processes.					
Authorization boundaries prevent AI systems or agents from accessing data, executing tools, changing configurations, or taking actions beyond approved scope.					
Access rights are promptly updated when roles, employment, vendors, projects, risk classifications, or system responsibilities change.					

Domain score: _____ / 25

Evidence or observations:

Domain 4: Model, Data & Pipeline Protection

Rate each statement from 1 (Strongly Disagree) to 5 (Strongly Agree). Base the rating on repeatable enterprise behavior and current evidence.

Training, tuning, evaluation, retrieval, and production data are protected against unauthorized access, tampering, poisoning, leakage, and inappropriate reuse.					
Model artifacts, prompts, configurations, embeddings, code, dependencies, registries, and deployment packages are protected with integrity, provenance, and version controls.					
Data and model pipelines use secure transfer, validation, scanning, segregation, approval, and rollback mechanisms appropriate to risk.					
Sensitive inputs and outputs are minimized, filtered, masked, encrypted, retained, and deleted according to approved policy and business need.					
Changes to models, data sources, prompts, tools, dependencies, and security settings are traceable and subject to controlled review before production release.					

Domain score: _____ / 25

Evidence or observations:

Domain 5: Secure Engineering, Testing & Validation

Rate each statement from 1 (Strongly Disagree) to 5 (Strongly Agree). Base the rating on repeatable enterprise behavior and current evidence.

Security and resilience requirements are defined during AI solution design and verified throughout development, integration, testing, deployment, and maintenance.					
AI systems undergo risk-based security testing that includes misuse testing, adversarial testing, abuse-case testing, vulnerability assessment, and validation of control boundaries.					
Testing evaluates failure modes such as unsafe outputs, unauthorized actions, data exposure, degraded dependencies, manipulated inputs, unavailable services, and incorrect fallback behavior.					
Release decisions require evidence that critical vulnerabilities and resilience gaps are resolved, mitigated, formally accepted, or blocked from production.					
Independent review or red-team activity is used for high-impact AI systems, with findings tracked to verified remediation and executive visibility.					

Domain score: _____ / 25

Evidence or observations:

Domain 6: Monitoring, Detection & Service Integrity

Rate each statement from 1 (Strongly Disagree) to 5 (Strongly Agree). Base the rating on repeatable enterprise behavior and current evidence.

Production monitoring covers security events, anomalous access, suspicious prompts, unusual agent behavior, data leakage, model drift, service degradation, and control failures.					
Logs and telemetry are sufficiently complete, protected, time-synchronized, retained, and correlated to support detection, investigation, accountability, and recovery.					
Detection thresholds and alerting are tuned to business criticality, model behavior, user patterns, threat intelligence, and acceptable operating boundaries.					
Human operators can identify when AI outputs, actions, integrations, or dependencies are behaving unexpectedly and can intervene before harm expands.					
Monitoring results drive timely containment, remediation, configuration changes, model updates, access changes, control improvements, and risk reassessment.					

Domain score: _____ / 25

Evidence or observations:

Domain 7: Incident Response & Crisis Coordination

Rate each statement from 1 (Strongly Disagree) to 5 (Strongly Agree). Base the rating on repeatable enterprise behavior and current evidence.

AI-specific incident response playbooks address compromise, data leakage, malicious use, unsafe autonomous behavior, corrupted models, service outages, and loss of critical dependencies.					
Escalation paths define who can isolate systems, disable models or agents, revoke access, halt automated decisions, activate manual operations, and communicate externally.					
Incident processes preserve evidence, support root-cause analysis, meet legal and regulatory obligations, and coordinate business, cybersecurity, privacy, technology, vendor, and communications teams.					
Response exercises include realistic AI scenarios and test decision authority, technical containment, stakeholder notification, operational workarounds, and executive crisis management.					
Lessons from incidents, near misses, exercises, threat intelligence, and industry events are incorporated into architecture, controls, training, monitoring, and response plans.					

Domain score: _____ / 25

Evidence or observations:

Domain 8: Continuity, Recovery & Adaptive Resilience

Rate each statement from 1 (Strongly Disagree) to 5 (Strongly Agree). Base the rating on repeatable enterprise behavior and current evidence.

Business continuity plans identify AI-enabled processes whose failure, degradation, manipulation, or unavailability could materially disrupt operations or stakeholder outcomes.					
Critical AI services have tested fallback procedures, manual alternatives, redundancy, capacity protections, dependency mapping, and clearly defined recovery priorities.					
Recovery plans address restoration of trusted models, data, prompts, configurations, integrations, credentials, logs, and operating environments after compromise or disruption.					
Recovery exercises verify recovery time, recovery point, data integrity, model integrity, service safety, communication, and return-to-operation criteria.					
Resilience metrics, exercise results, incidents, changing threats, architecture changes, and business priorities are used to continuously improve preparedness and reduce systemic fragility.					

Domain score: _____ / 25

Evidence or observations:

Scoring & Security Resilience Interpretation

Record each domain score. Treat any domain below 15 as a priority security or resilience weakness even when the overall score is higher.

1. Security Governance & Accountability	____ / 25
2. Threat Modeling & Attack Surface Management	____ / 25
3. Identity, Access & Privileged Control	____ / 25
4. Model, Data & Pipeline Protection	____ / 25
5. Secure Engineering, Testing & Validation	____ / 25
6. Monitoring, Detection & Service Integrity	____ / 25
7. Incident Response & Crisis Coordination	____ / 25
8. Continuity, Recovery & Adaptive Resilience	____ / 25
Total AI security and operational resilience score	**____ / 200**

Overall Third-Party Risk Capability Levels

180-200	**Resilient by Design**	Security, monitoring, response, continuity, recovery, and learning operate as an integrated, evidence-driven, and continuously improved enterprise capability.
150-179	**Well Protected**	Strong security and resilience practices are established, with targeted gaps in threat coverage, testing, detection, coordination, recovery, or adaptive improvement.
110-149	**Developing**	Important controls exist, but security engineering, monitoring, incident readiness, continuity, or recovery remains uneven across AI systems and business areas.
70-109	**Fragmented**	Protection depends on isolated technical controls or reactive response, with material weaknesses in accountability, visibility, testing, coordination, or recovery.
40-69	**Highly Exposed**	Foundational governance, threat management, access control, system protection, detection, response, continuity, and recovery capabilities are insufficient for dependable AI operations.

Important: This assessment is a discussion and prioritization tool, not a certification, penetration test, security audit, or substitute for formal continuity planning. Results should be validated using current inventories, threat models, architecture records, access reviews, test evidence, vulnerability findings, monitoring data, incident records, continuity plans, recovery exercises, and risk decisions.

Executive Reflection & 90-Day Action Plan

Use the scoring pattern to identify the few leadership actions that will most reduce the likelihood, impact, and duration of AI-related compromise, unsafe operation, or service disruption.

Security and resilience strengths

Which governance, threat-management, access-control, engineering, monitoring, response, continuity, or recovery practices currently provide the strongest protection?

Critical security and resilience exposures

Which AI systems, data flows, identities, agents, dependencies, vulnerabilities, failure modes, or recovery gaps create the greatest enterprise exposure?

Decisions required from leadership

What investment, architecture, risk-acceptance, remediation, isolation, shutdown, continuity, recovery, or operating-model decisions cannot be delegated?

90-Day Priority Actions

Executive commitment: We will review progress by _____ and reassess affected domains using current security, incident, continuity, and recovery evidence.

Continue Your AI Security & Resilience Journey

Trustworthy AI requires more than preventing intrusion. Sustainable value depends on secure design, controlled access, protected data and models, continuous detection, decisive response, tested continuity, reliable recovery, and learning that strengthens the system after every change or disruption.

Explore the Intelligent Systems Press collection and Free Resources library
www.intelligentsystemspress.com