AI RISK MANAGEMENT FRAMEWORK ALIGNMENT STATEMENT

NIST AI RMF 1.0 | IntelliThreat AI Platform & Autonomous Tier | Version 1.1 | Last updated: May 28, 2026

PURPOSE & SCOPE

This AI Risk Management Framework Alignment Statement ("AI RMF Statement") describes Blueshift Cybersecurity, Inc.'s ("Blueshift's") internal, good-faith self-assessment of how the IntelliThreat AI Platform and IntelliThreat Autonomous tier (collectively, the "Covered AI Systems") align with the **NIST AI Risk Management Framework (AI RMF 1.0)** (NIST AI 100-1). It also addresses applicable obligations under the **EU AI Act** and Blueshift's use of **ISO/IEC 42001** as a maturity reference only.

This Statement is made available to enterprise customers upon request as referenced in the **INTELLITHREAT AUTONOMOUS ADDENDUM, Section 7.3,** and the **AI SAFETY & MODEL GOVERNANCE Statement**. It is informational only, reflects Blueshift's internal self-assessment as of the date above, and does not create independent contractual obligations. This Statement is not an audit report, attestation, certification, legal opinion, or independent third-party assessment of the Covered AI Systems, Blueshift's controls, or Blueshift's compliance with any law, regulation, framework, or standard. Contractual terms governing AI use, liability, and data processing are set forth in the PLATFORM MASTER SUBSCRIPTION AGREEMENT, the INTELLITHREAT AI PLATFORM ADDENDUM, the INTELLITHREAT AUTONOMOUS ADDENDUM, and the DATA PROCESSING ADDENDUM — INTELLITHREAT AI PLATFORM.

1. NIST AI RMF 1.0 — Framework Overview

The NIST AI RMF 1.0 organizes AI risk management into four core functions: **GOVERN**, **MAP**, **MEASURE**, and **MANAGE**. The following table summarizes Blueshift's self-assessed alignment across each function for the Covered AI Systems.

NIST AI RMF Function	Description	Blueshift Practice	Maturity
GOVERN	Policies, accountability, culture, and organizational practices for AI risk	AI Governance Program with executive oversight; cross-functional review; documented policies for model lifecycle, data use, and automation-tier gating	Established
MAP	Categorize AI use cases, identify context, assess risk	AI use case registry; risk classification by consent tier; threat model reviews for new features; customer environment impact assessments	Established

NIST AI RMF Function	Description	Blueshift Practice	Maturity
MEASURE	Quantify and track AI risks using metrics and evaluation	Accuracy/bias/safety testing for model updates; red-team simulated adversarial testing; false positive/ negative monitoring; 12-month tamper-evident audit logs for Autonomous Actions	Established
MANAGE	Prioritize and treat AI risks; monitor and adjust	Two-tier consent model; Kill-Switch (60-second target); tiered incident notification (2/4/24-hour targets); Customer override controls; model update validation gates	Established

2. GOVERN — Organizational AI Risk Culture & Accountability

2.1 AI Governance Program

Blueshift maintains an internal AI Governance Program that includes:

- **Executive Oversight:** Cross-functional review of AI features, risk posture, and policy changes at the executive level.
- **Model Lifecycle Management:** Documented processes for model selection, training, validation, deployment, monitoring, and deprecation.
- **Risk Assessments:** Structured risk assessments for new AI features, automation capabilities, and significant model updates, with particular scrutiny for IntelliThreat Autonomous tier features.
- **Policy Documentation:** Written policies governing training data use, Customer Data protection, autonomous action authorization, and incident response for AI-related incidents.
- **Annual Review:** The AI Governance Program, this AI RMF Statement, and associated policies are reviewed at least annually, and updated as Blueshift determines appropriate when material changes to the Covered AI Systems occur.

2.2 Roles & Accountability

Role	AI Risk Responsibilities
Executive Sponsor	Overall AI governance accountability; approves IntelliThreat Autonomous tier feature releases
Engineering / ML Team	Model development, validation, safety testing, and red-team exercises; maintains model registry
Product Team	AI use case classification; feature risk assessment; customer-facing documentation and disclosures

Role	AI Risk Responsibilities
Security Team	Threat modeling for AI features; autonomous action security review; incident response for AI incidents
Legal / Compliance	Regulatory monitoring (EU AI Act, CCPA, GDPR); contractual AI obligations; this Statement and related disclosures
Customer	Configuration of consent tiers; maintenance of human oversight where desired; designation and maintenance of an Autonomous Operations Contact where applicable; Kill-Switch activation; review of environment-level risk acceptance

3. MAP — AI Use Case Context & Risk Identification

3.1 Covered AI Systems Classification

Blueshift has identified and classified the following AI use cases within the Covered AI Systems:

AI Capability	Category	Automation Tier	Risk Level
Threat Detection & Alerting	Detection	Advisory (default)	**Low — informational output only**
AI Insights & Recommendations	Decision Support	Advisory (default)	**Low-Medium — human review required before action**
Automated Alert Suppression	Autonomous	Autonomous (opt-in)	**Medium — requires human confirmation gate**
Policy / Firewall Rule Modification	Autonomous	Autonomous (opt-in)	**High — explicit enablement; Kill-Switch required**
Account Suspension / Access Restriction	Autonomous	Autonomous (opt-in)	**High — explicit enablement; audit logged**
Cloud Instance Termination	Autonomous	Autonomous (enterprise consent)	**Very High — enterprise written consent required**
SOC Human-Confirmation Review	Human-in-Loop	Human-in-Loop (separately purchased SOC-monitored IntelliThreat, if applicable)	**Low — separately purchased supplemental human oversight layer through SOC-monitored IntelliThreat, if applicable**

3.2 Risk Context Factors

For each AI use case, Blueshift evaluates the following risk context factors during feature design and review:

- **Impact Scope:** Potential blast radius of an erroneous autonomous action (e.g., single asset vs. enterprise-wide).
- **Reversibility:** Whether an automated action can be reversed or mitigated within a reasonable timeframe.
- **Data Sensitivity:** Nature of Customer Data processed, including whether it contains PHI, PII, or other regulated data.
- **Customer Control Availability:** Whether kill-switch, human confirmation, or override controls are technically feasible for the use case.
- **Third-Party Dependencies:** Integration surfaces (M365, AWS, GCP, Azure, Salesforce, SAP) that expand autonomous action scope.

4. MEASURE — AI Risk Quantification & Evaluation

4.1 Model Evaluation

Blueshift's ML and engineering teams conduct the following evaluations for model updates and new AI features:

- **Accuracy Testing:** Detection accuracy metrics including true positive rate, false positive rate, and precision-recall analysis across representative threat scenarios.
- **Bias Assessment:** Review of model outputs for systematic biases that could result in differential treatment of assets, users, or environments.
- **Safety Validation:** Pre-deployment review of autonomous action logic to confirm actions are appropriately scoped and are designed to operate within Customer-authorized parameters.
- **Regression Testing:** Before any model update affecting Autonomous Actions, regression testing confirms no material expansion of action scope except in accordance with the applicable notice, incorporation, and version-control mechanics under the PLATFORM MASTER SUBSCRIPTION AGREEMENT and the applicable Addenda.

4.2 Red-Team & Adversarial Testing

Blueshift conducts simulated adversarial attack exercises designed to:

- Test whether AI detection logic can be evaded by known adversarial techniques.
- Validate that autonomous action guardrails cannot be bypassed by malformed telemetry or injection attacks.
- Confirm Kill-Switch and override controls function correctly under adversarial conditions.

Findings from red-team exercises feed directly into model updates, guardrail enhancements, and governance reviews.

4.3 Audit Logging & Continuous Monitoring

- **Autonomous Action Audit Logs:** All Autonomous Actions are recorded in tamper-evident audit logs retained for a minimum of 12 months. Logs include: timestamp, action type, triggering signal, Customer configuration state, and outcome.
- **False Positive / Negative Monitoring:** Ongoing monitoring of autonomous action outcomes to identify systematic errors for model feedback.
- **Availability Metrics:** Platform availability and incident data may be published on Blueshift's public Service Status Page for informational purposes.
- **Customer Access:** Audit logs for Autonomous Actions are accessible to Customer via the Platform interface or API, subject to the applicable purchased offering and then-current Platform functionality.

5. MANAGE — AI Risk Treatment, Controls & Response

5.1 Two-Tier Consent Model

Blueshift's primary risk treatment for autonomous execution is a two-tier consent model that gates access to progressively higher automation levels:

Tier	Mode	Execution Model	Consent Required
1 — Advisory	Advisory Mode	AI recommendations only; no automated execution	None (default; all customers)
2 — Autonomous	Autonomous Mode	Full agentic execution; actions execute without per-action human review	Enterprise explicit written consent; activated on an applicable Order and subject to a signed Autonomous Mode Authorization on the applicable Order Form or Customer Authorization & Acknowledgment (with the framework set forth in the Autonomous Activation Record posted at https://www.blueshiftcyber.com/legal/autonomous-activation-record) or other express assent permitted by the PLATFORM MASTER SUBSCRIPTION AGREEMENT

5.2 Kill-Switch

The Kill-Switch is Blueshift's primary real-time risk treatment for Autonomous Mode:

- **Activation:** Customer may activate at any time via the Platform; no Blueshift authorization required.
- **Effect:** Future or pending Autonomous Mode actions are targeted for cessation within 60 seconds of activation, subject to then-current system availability and network conditions.

- **Scope:** Designed to apply to Autonomous Mode actions across Customer-authorized integration surfaces within the then-current Platform functionality.
- **Reactivation:** Requires deliberate Customer action; does not auto-reset.
- **Logging:** To the extent available in the then-current Platform, material Kill-Switch activations and mode changes are reflected in the Audit Log or comparable operational record.

5.3 Incident Response for AI Incidents

Blueshift maintains a documented AI incident response procedure that may include:

- **Detection & Classification:** AI incidents are classified by severity (1–3) based on scope of autonomous actions taken, potential customer impact, and whether Kill-Switch was activated.
- **Notification Targets:** Severity-1 (2 hours), Severity-2 (4 hours), Severity-3 (24 hours) from confirmation of incident, subject to the IntelliThreat Autonomous Addendum.
- **Containment:** Immediate suspension of affected autonomous action workflows; Kill-Switch activated on Blueshift's behalf if Customer has not already done so.
- **Root-Cause Analysis:** RCA may be initiated for material AI incidents in accordance with Blueshift's internal incident response procedures, and any related customer communications remain subject to the applicable governing agreements.
- **Customer Communication:** Blueshift may provide a summary report to Customer upon completion of root-cause analysis for Severity-1 and Severity-2 incidents, subject to the applicable governing agreements.

5.4 Model Update Risk Controls

- Model updates affecting Autonomous Action logic undergo safety validation and regression testing before deployment.
- Updates that materially expand Autonomous Action scope are subject to the applicable notice, incorporation, and version-control mechanics under the PLATFORM MASTER SUBSCRIPTION AGREEMENT and the applicable Addenda.
- Rollback capability is maintained for all model updates affecting autonomous execution.

6. EU AI Act — Regulatory Alignment

6.1 Current Classification Assessment

As of the effective date of this Statement, Blueshift's current assessment is:

BLUESHIFT CURRENT EU AI ACT POSITION
Blueshift does not currently believe the Covered AI Systems constitute **prohibited AI practices** under EU AI Act Article 5 (e.g., subliminal manipulation, social scoring, real-time biometric surveillance in public spaces).
With respect to **high-risk AI system** classification under Annex III: Cybersecurity AI systems assisting in national critical infrastructure protection may qualify as high-risk. Blueshift is monitoring evolving implementing regulations and sector-specific guidance.

6.2 High-Risk AI System Obligations (Upon Determination)

If any Covered AI Capability is determined to be a high-risk AI system under the EU AI Act, Blueshift will implement:

- **Conformity Assessment:** Documented technical documentation per Article 11; registration in the EU AI Act database where required.
- **Human Oversight:** Enhanced human oversight requirements beyond existing two-tier consent model, as required by implementing regulations.
- **Transparency:** Disclosure to deployers of high-risk classification, intended purpose, limitations, and required oversight measures.
- **Data Governance:** Training, validation, and testing data governance documentation per Article 10.
- **Logging Requirements:** Logging capabilities to ensure traceability per Article 12, supplementing existing 12-month audit log requirements.
- **Post-Market Monitoring:** Systematic post-market monitoring plan per Article 72.

Blueshift may update this Statement, the AI SAFETY & MODEL GOVERNANCE Statement, and, where appropriate, the applicable Addenda upon any determination of high-risk classification, subject to the applicable notice, incorporation, and version-control mechanics under the PLATFORM MASTER SUBSCRIPTION AGREEMENT and the applicable governing documents.

7. ISO/IEC 42001 — AI Management Systems Reference

Blueshift uses ISO/IEC 42001 as a maturity reference for its AI Governance Program. The following table maps key ISO 42001 clauses to Blueshift's current practices based on Blueshift's internal assessment and is not a statement of certification or third-party conformity assessment.

ISO 42001 Clause	Requirement	Blueshift Practice	Status
4 — Context	Understand internal/ external context; stakeholder requirements	AI use case registry; customer risk context; regulatory monitoring	**Aligned**
5 — Leadership	Executive accountability; AI policy	Executive AI governance sponsor; documented AI policies	**Aligned**
6 — Planning	Objectives; risk and opportunity assessment	Feature risk assessments; annual governance review	**Aligned**
7 — Support	Resources; competence; awareness; documentation	ML team training; model documentation; this Statement	**Aligned**

ISO 42001 Clause	Requirement	Blueshift Practice	Status
8 — Operation	Operational planning; AI system lifecycle	Model lifecycle management; validation gates; rollback capability	**Aligned**
9 — Performance	Monitoring, measurement, analysis	Accuracy/bias testing; audit logs; red-team exercises	**Aligned**
10 — Improvement	Nonconformity; continual improvement	Post-incident review; model update feedback loops	**Aligned**
Annex A — AI Controls	AI-specific controls (bias, transparency, safety, data governance)	Two-tier consent; Kill-Switch; audit logs; data de-identification; explainability where feasible	**Substantially Aligned**

Blueshift is evaluating formal ISO/IEC 42001 certification as part of its ongoing AI governance maturity roadmap and does not by this Statement represent that it is currently ISO/IEC 42001-certified.

8. Customer Responsibilities Under This Framework

This Statement describes Blueshift's AI risk management practices. Customers using the Covered AI Systems retain the following responsibilities, as further set forth in the INTELLITHREAT AUTONOMOUS ADDENDUM and the PLATFORM MASTER SUBSCRIPTION AGREEMENT:

- **Tier Selection:** Independently assessing whether Advisory or Autonomous Mode is appropriate for their environment and risk tolerance.
- **Configuration Validation:** Reviewing, testing, and validating automation settings and confirmation gates prior to production enablement.
- **Kill-Switch Readiness:** Ensuring designated personnel are trained and authorized to activate the Kill-Switch promptly if needed.
- **EU AI Act Assessment:** Conducting independent EU AI Act impact assessments for Customer's specific deployment context, particularly for critical infrastructure or regulated environments.
- **Human Oversight:** Maintaining appropriate human oversight and human-in-the-loop controls to the extent applicable to Customer's deployment, risk tolerance, and legal obligations.
- **Audit Log Review:** Periodically reviewing Autonomous Action audit logs for anomalies, unexpected patterns, or actions inconsistent with Customer configuration intent.

9. Availability & Updates

This AI RMF Statement is:

- Made available to enterprise customers upon request, as referenced in the INTELLITHREAT AUTONOMOUS ADDENDUM, Section 7.3.
- Reviewed periodically by Blueshift, including at least annually as a matter of current practice, and upon material changes to the Covered AI Systems or applicable regulatory requirements.
- May be updated upon any determination of EU AI Act high-risk classification, material expansion of Autonomous Action scope, or change in Blueshift's AI RMF maturity status.
- Not a contractual commitment — contractual terms governing AI are in the documents identified in the Purpose & Scope section above, and this Statement reflects Blueshift's internal self-assessment rather than any external audit, certification, or attestation.

For questions regarding this Statement, contact Blueshift at: legal@blueshiftcyber.com

DISCLAIMER

This AI RMF Statement is informational only. It describes Blueshift's practices as of the effective date and reflects Blueshift's internal self-assessment only; it does not constitute a warranty, guarantee, or contractual representation. This AI RMF Statement is not an audit report, attestation, certification, legal opinion, regulator-issued determination, or independent third-party verification of Blueshift's alignment with the NIST AI RMF, ISO/IEC 42001, the EU AI Act, or any other framework, standard, or law. Blueshift's actual contractual obligations regarding AI capabilities, data use, autonomous actions, liability, and incident response are exclusively governed by the PLATFORM MASTER SUBSCRIPTION AGREEMENT, the INTELLITHREAT AI PLATFORM ADDENDUM, the INTELLITHREAT AUTONOMOUS ADDENDUM (where applicable), and the DATA PROCESSING ADDENDUM — INTELLITHREAT AI PLATFORM. In the event of any conflict between this Statement and those agreements, the applicable agreement controls.