

HRM RFP & Vendor Evaluation Kit

A Structured Procurement Framework for
Evaluating AI-Native Human Risk Platforms



Living Security helps organizations measure and reduce human cyber risk using AI-native Human Risk Management (HRM). This toolkit is designed to help security leaders build internal alignment, evaluate vendors, and operationalize measurable human risk reduction.

Procurement Overview

Human Risk Management is not an extension of Security Awareness Training. It is behavioral risk intelligence infrastructure.

This evaluation framework is designed to help procurement, security, and risk teams:

- Differentiate AI-native HRM platforms from enhanced SAT solutions
- Define non-negotiable architectural requirements
- Weight measurable exposure reduction over content features
- Maintain objective scoring discipline during vendor evaluation

Use this document during RFP issuance, vendor demos, and proof-of-concept evaluation.

Navigating the RFP Process

When you issue an RFP for Human Risk Management (HRM), evaluate more than training capabilities. HRM is not an extension of Security Awareness Training, it is behavioral risk intelligence infrastructure. The RFP process should prioritize architectural depth, measurable exposure reduction, and AI-native behavioral modeling.

Non-Negotiable Core Requirements (Must-Have)

The following capabilities define true HRM platforms and should be designated as **Mandatory (M)** in any RFP:

1. AI-Native Architecture

- Platform architected as a behavioral intelligence system — not analytics layered onto a training engine
- Machine-learning-driven dynamic recalibration of individual and cohort risk scores
- Real-time modeling of human and AI-assisted behavior

2. Behavioral Signal Coverage

- Aggregation and modeling of **100+ validated behavioral and technical signals**
- Cross-tool telemetry ingestion beyond phishing simulations (identity, endpoint, email, DLP, collaboration platforms)

- Ability to expand signal library as new threat patterns emerge

3. Exposure Duration Analytics

- Ability to measure, trend, and report **time spent in elevated risk states**
- Individual, cohort, and organizational exposure duration reporting
- Executive-level trend analysis tied to risk reduction metrics

4. Automated Intervention Engine

- Real-time, risk-triggered interventions
- Context-aware automation (training, nudges, policy enforcement)
- No manual assignment required for high-risk events

5. Parallel Deployment Capability

- Ability to operate alongside existing SAT and phishing vendors
- Telemetry ingestion from third-party training systems
- Support for phased transition without vendor disruption

Platforms unable to meet these criteria are unlikely to deliver continuous human risk governance.

Warning Signs:

Enhanced SAT vs. True HRM

During evaluation, be alert for indicators that a vendor is offering enhanced training rather than comprehensive HRM.

Red Flags include:

- Risk scores derived primarily from phishing simulation data
- Fewer than 100 behavioral signals tracked
- No measurable exposure duration analytics
- Static, rule-based scoring without ML recalibration

- Manual intervention workflows
- No cross-tool telemetry ingestion
- No visibility into AI tool usage or non-human identities
- Inability to operate in parallel with existing SAT vendors

If multiple red flags are present, the platform may lack the architectural depth required for continuous human risk governance.

Functional Evaluation Categories

Below is a structured RFP framework aligned to enterprise HRM evaluation.

A. Risk Visibility & Quantification

Category	Requirement	Requirement Example	Priority (M/Must, S/Should, O/Optional)
Risk Signal Aggregation	Continuous aggregation of behavioral and technical risk signals	Platform must ingest live telemetry from identity, email, endpoint, DLP, and collaboration systems and unify signals into a single risk model without manual uploads.	M
Behavioral Signal Depth	100+ validated behavior signals tracked and modeled	Vendor must provide documented list of 100+ discrete behavioral and technical signals used in risk modeling (e.g., MFA failures, anomalous logins, data downloads, phishing responses).	M
AI-Driven Risk Modeling	Machine-learning-based dynamic risk recalibration	Risk scores must automatically recalibrate using ML models as new behavioral data is ingested, without requiring static rule updates.	M
Exposure Duration Analytics	Exposure duration measurement and trend reporting	Platform must calculate and trend the time users remain in elevated risk states at individual, cohort, and enterprise levels.	M
Risk Segmentation	Risk segmentation by department, role, privilege level	System must support dynamic segmentation and filtering of risk scores by department, role, geography, manager, or privilege status.	S
Score Transparency	Explainable scoring with audit trail of changes	Platform must provide visibility into which behavioral signals influenced a score and maintain an audit trail of recalibration and intervention triggers.	M

B. Integrations & Extensibility

Category	Requirement	Requirement Example	Priority (M/Must, S/Should, O/Optional)
Security Stack Integration	Deep integration with IAM, endpoint, DLP, email, SIEM	Must provide prebuilt or API integrations with major platforms (e.g., Okta, Azure AD, M365, CrowdStrike, Proofpoint) and demonstrate live data ingestion.	M
Telemetry Architecture	Live telemetry ingestion (not batch exports)	Risk model must update in near real-time based on streaming or API-based telemetry, not CSV uploads or periodic batch sync.	M
SAT Interoperability	Integration with existing SAT vendors	Platform must ingest data from third-party training/phishing vendors and operate in parallel during contract periods.	M
Data Portability	API access for BI and data export	Provide documented API and support export of risk data into BI platforms (e.g., Power BI, Tableau) and GRC systems.	S
AI Governance Alignment	Support for AI governance tooling	Platform must detect AI tool usage signals and integrate with AI governance or DLP controls to monitor generative prompt exposure.	S

C. Integrations & Extensibility

Category	Requirement	Requirement Example	Priority (M/Must, S/Should, O/Optional)
Automated Risk Mitigation	Real-time automated interventions	System must automatically trigger targeted interventions (e.g., microlearning, nudges, policy enforcement) when risk thresholds are met.	M
Risk-Based Personalization	Risk-based training personalization	Training assignments must dynamically adjust based on individual risk score and behavior history.	M
Accountability Enablement	Manager and employee scorecards	Provide personalized dashboards for users and managers showing risk score, exposure duration, and behavior trends.	S
Policy Governance	Policy attestation workflows	Enable automated assignment and tracking of policy acknowledgments, particularly for elevated-risk users.	S
IR Workflow Alignment	Incident response integration	Integrate with incident response tools or workflows to trigger remediation actions when high-risk events occur.	S

D. Executive Reporting & Governance

Category	Requirement	Requirement Example	Priority (M/Must, S/Should, O/Optional)
Board Governance Reporting	Board-ready dashboards focused on risk reduction	Executive dashboards must quantify risk exposure, exposure duration trends, and high-risk user concentration with export-ready visuals.	M
Exposure Trend Reporting	Exposure duration trend reporting	Platform must provide longitudinal reporting of exposure duration reduction over time.	M
Risk Concentration Analysis	High-risk user concentration tracking	System must identify and report percentage of users driving disproportionate organizational risk.	M
Financial Impact Modeling	ROI and loss-exposure modeling tools	Provide modeling tools that estimate risk reduction impact and potential loss avoidance for CFO reporting.	S

E. Vendor Strength & Category Leadership

Category	Requirement	Requirement Example	Priority (M/Must, S/Should, O/Optional)
Category Leadership	Demonstrated category leadership in Human Risk Management	Vendor must demonstrate published HRM frameworks, research, and documented category positioning beyond SAT.	M
Market Validation	Analyst recognition or independent validation	Provide evidence of analyst coverage, industry recognition, or third-party validation in HRM or behavioral risk intelligence.	S
Thought Leadership	Published maturity framework and research contributions	Vendor must provide documented maturity model and research publications advancing HRM practices.	S
Enterprise Credibility	Enterprise customer references	Provide customer case studies and references from organizations of comparable size and complexity.	M

Vendor Evaluation Scorecard

To ensure objective comparison, use a weighted scoring framework during demos and vendor responses.

Category	Weight	Vendor A	Vendor B	Vendor C
AI-Native Architecture	20%			
Behavioral Signal Coverage (100+ signals)	15%			
Exposure Duration Analytics	15%			
Automated Intervention Engine	15%			
Cross-Tool Integrations	10%			
Parallel Deployment Capability	10%			
Executive Risk Reporting	10%			
Vendor Category Leadership	5%			

This structure ensures architectural depth and measurable exposure reduction are weighted more heavily than content libraries or phishing simulation features.

Questions That Reveal Platform Limitations

During vendor evaluation, ask:

- How many behavioral signals are actively tracked and modeled?
- Does the platform ingest live telemetry beyond phishing simulation data?
- How is exposure duration calculated and trended?
- Can risk scores dynamically recalibrate based on machine learning?
- Can interventions trigger automatically without manual assignment?
- Can the platform operate alongside our current SAT vendor?
- How does the system model AI-assisted and non-human identity behavior?

These questions quickly distinguish true HRM platforms from enhanced training solutions.

Guidance for Using This RFP Framework

1. Customize requirements to reflect your regulatory and risk posture.
2. Assign clear priorities (Must / Should / Optional).
3. Require vendors to provide evidence-based responses, including sample dashboards and scoring methodology documentation.
4. Use the weighted scorecard during demos to maintain evaluation discipline.
5. Conduct a proof of concept focused on exposure duration reduction and automated intervention workflows.

A digital, customizable spreadsheet version of this RFP framework is available to support your evaluation process.

Streamline Your HRM Evaluation

Access the 90-Day HRM Playbook for:

- A deployment approach optimized for speed-to-value
- Operational milestones you can use in procurement conversations
- Demo criteria and success indicators built into rollout planning

Download the 90-Day HRM Playbook

<https://www.livingsecurity.com/90-day-hrm-playbook>

PROOF POINT:

"Forrester flags vendors whose risk scoring relies primarily on training and phishing simulations rather than actual behavioral events, limiting HRM depth and effectiveness."

Source: Forrester Wave™ Human Risk Management, Q3 2024 <https://www.livingsecurity.com/forrester-wave-report-2024>

