Suhas Bhairav

AI Phishing Report Triage Guide

A practical guide to classifying, prioritizing, and automating phishing report triage.

An actionable framework to automate phishing report triage while retaining human judgment for high-risk cases.

Suhas Bhairav

suhasbhairav.com

Phishing Report Triage Basics

This page defines the scope of AI-assisted phishing report triage for IT and Security teams. It outlines what to automate, what to keep human-reviewed, and how progress will be measured.

Key move **Automate safely and preserve human judgment for edge cases.**

PRACTICAL CHECKLIST

01 Intake normalization: standardize reports into fields (sender, subject, body, attachments).

02 Indicator extraction: parse URLs, domains, hashes, attachments, and known-bad indicators.

03 Auto-classification: assign risk levels and flag ambiguity for human review.

04 Human-in-the-loop review: escalate high-risk or unclear cases to the security lead.

05 Context enrichment: attach threat intel, user context, and incident history to each report.

06 Escalation routing: auto-queue to IR or security comms; notify owners as needed.

07 Compliance & audit: log decisions and preserve data per policy; enable audits.

Suhas Bhairav AI Resource

An actionable framework to automate phishing report triage while retaining human judgment for high-risk cases.

Automation Scope & Data Needs

Identify what to automate and what requires human validation in phishing report triage. Outline data, tools, and integration points.

Key move **Start with low-risk reports to validate automation first.**

PRACTICAL CHECKLIST

01	Data requirements: standardized report fields, email headers, attachments, and user context.
02	Tools and platforms: SIEM, EDR, email gateway, threat intel, RPA, and AI agents.
03	Automation targets: indicator extraction, auto-classification, enrichment, and low-risk user notifications.
04	Human-reviewed gates: high-risk, suspected compromise, policy exceptions, or indicators with low confidence.
05	Data quality & privacy controls: dedup, spoof checks, data minimization, and access controls.
06	AI prompts & governance: guardrails, prompt versioning, and review trails.
07	Weekly adoption plan: Week 1 intake standardization; Week 2 auto-extract indicators; Week 3 auto-classify; Week 4 alert and escalate.

Suhas Bhairav AI Resource

An actionable framework to automate phishing report triage while retaining human judgment for high-risk cases.

Risks, Roles, & Measurement

Define guardrails, approvals, and ownership for triage decisions. Establish metrics to gauge accuracy, speed, and user impact.

Key move **Document ownership and review thresholds; misclassification harms users.**

PRACTICAL CHECKLIST

01	Guardrails: require human review for high-risk or ambiguous cases before action.
02	Approvals & ownership: assign triage owners; document SLAs and change-control for automation scope.
03	Data privacy & retention: set retention windows, access controls, and anonymization where possible.
04	Risk scoring: define rubric (High/Medium/Low) with concrete indicators and examples.
05	Metrics & measurement: track false positives, MTTR, containment time, and user-reported incidents.
06	Workflow ownership: document process steps, ownership, and change-control for any automation adjustment.
07	User communication: use clear, non-alarmist messages; provide safe next steps and reporting channels.

Suhas Bhairav AI Resource

An actionable framework to automate phishing report triage while retaining human judgment for high-risk cases.

ABOUT THE CREATOR

Suhas Bhairav

I build AI systems that combine distributed architecture, retrieval, knowledge graphs, security, and practical workflow design. My focus is AI that can be used repeatedly by real teams: understandable, reliable, and grounded in the work people already do.

EXPERTISE

- Principal systems architecture and applied AI engineering
- Production-grade AI systems, agentic workflows, RAG, and knowledge graphs
- Full-stack workflow applications across AWS, Google Cloud, Azure, on-premise, and hybrid environments
- Practical AI implementation patterns for automotive, manufacturing, finance, logistics, cybersecurity, sales, and customer operations
- Security, governance, observability, approvals, and failure-mode thinking for operational AI systems

WEBSITE

suhasbhairav.com

EMAIL

suhasbhairav@gmail.com