Adversarian Labs

Adversaia Resilience

Resilience Assessment

Fraud & FinCrime + Agentic Workflow Validation

Assessment Scope (Template)

Fraud & FinCrime Resilience • Agentic Workflow Resilience

PREPARED FOR

Customer / Prospect Name

PREPARED BY

Adversarian Labs

REPORT DATE

December 26, 2025

VERSION

1.0 (Template)

CONFIDENTIAL — For internal use by the recipient. Redistribution prohibited.

Confidentiality and Intended Use

This report is provided solely for the recipient’s internal evaluation of control effectiveness and resilience. All scenarios, personas, transactions, and evidence artifacts are synthetic and used for defensive validation only. No guidance in this report should be interpreted as instructions to commit wrongdoing.

How to read this report

Executive Summary provides headline outcomes (ARI score, band, and the most material gaps). Findings translate failed scenarios into concrete defensive patch patterns and retest guidance. Lineage provides decision-level traceability for each scenario result, supporting governance review and remediation verification.

Executive Summary

Adversaia executed synthetic adversarial scenario packs aligned to key financial decision workflows. Results below reflect the current defensive posture and highlight the highest-impact control gaps to remediate next.

OVERALL ARI
78/100
Amber band

PASS RATE
73%
Across scenarios

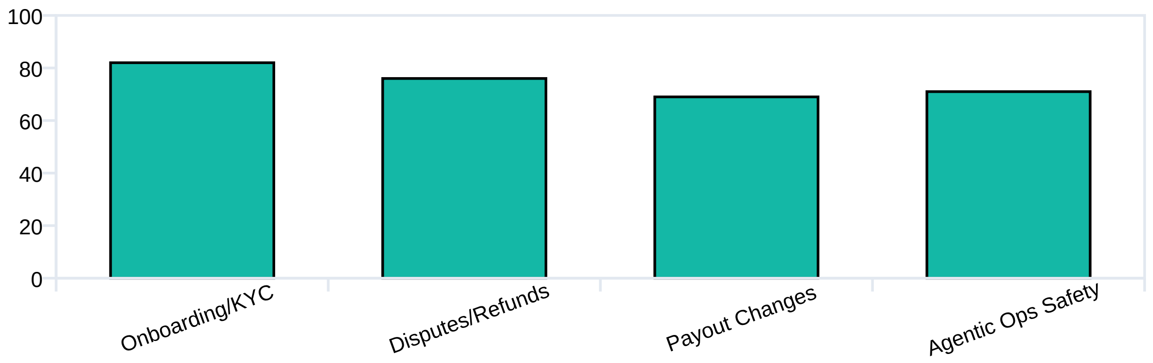
HIGH/CRIT FINDINGS
4
Require action

Top risks observed (sample)

• Insufficient cross-system correlation between support actions and payout/beneficiary changes.
• Weak step-up and cooling periods for high-risk account changes following recent authentication events.
• Incomplete agent/tool traceability and approval gating for sensitive operational actions (agentic workflows).
• Limited evidence packaging for repeatable governance review (exports and run comparisons).

Pack-level summary (sample)

Pack ARI Scores (Sample)



Key Findings

Findings are generated from failed scenarios and mapped to defensive remediation patterns. Each finding includes a control gap statement, a patch pattern, and retest guidance.

Severity	Finding	Control gap	Recommended patch pattern
CRIT	Support override + payout change not gated	No enforced approval gate when a support action precedes a payout destination change within a risk window.	Require step-up + second-line approval; enforce 24–72h cooling period; correlate support actions to payout changes.
HIGH	Insufficient device/session risk linkage	Device trust and session risk signals are not consistently required for beneficiary changes.	Require device posture + session binding for sensitive changes; trigger review if device/session anomalies exist.
HIGH	Agent tool scope too broad (AWR)	Agent tool calls can propose sensitive actions without narrow scopes and explicit policy checks.	Restrict tool scopes; enforce policy evaluation before execute; require approvals for high-risk tool actions; log every tool call.
MED	Dispute escalation lacks velocity constraints	Dispute and refund workflows do not consistently enforce velocity limits across linked entities.	Add velocity controls across account/device/entity graph; add friction for repeat dispute patterns; retest against dispute pack.

Retest workflow

After implementing patch patterns, re-run the same pack version to measure deltas. Target outcomes: reduced HIGH/CRIT findings, improved ARI score, and improved dimension scores for affected decision points.

Program Overview

Adversaia Resilience supports two complementary programs under one platform: **Fraud & FinCrime Resilience** (traditional controls and models) and **Agentic Workflow Resilience** (AI-in-the-loop operations: tool permissions, approvals, and policy enforcement).

Fraud & FinCrime Resilience

Validates decision workflows such as onboarding/KYC, disputes/refunds, and payout/beneficiary changes. Focus areas include signal completeness, correlation across systems, step-up/hold/review actions, and operational latency.

Agentic Workflow Resilience

Validates AI-assisted operations and automation layers. Focus areas include tool scope boundaries, approval gating, policy enforcement, and trace completeness (agent action → tool calls → checks → approvals → outcomes).

Methodology

Scenario packs

Scenario packs are curated suites of synthetic control-stressors mapped to workflow stages and decision points. Scenarios specify expected defensive controls and required signals; they never describe operational wrongdoing steps.

Scoring (ARI)

Each scenario produces dimension scores (0–100) (e.g., Identity, Behavior, Device, Network, Payments, Human Review, Policy). The overall ARI score summarizes defensive success across decision points and supports trending over time. Bands are used for executive readability: Green (strong), Amber (moderate), Red (weak).

Decision lineage

For every scenario, lineage events record what was evaluated, which checks were performed, and the outcome. Agentic workflows include simulated tool calls, policy checks, and approval gates for traceability.

Decision Contracts and Coverage

Decision contracts define what must always be true at a decision point (required signals, required checks, required approvals, and constraints). Coverage mapping shows which contracts are tested by which scenarios and identifies gaps to prioritize next.

Decision contract	Workflow	Coverage	Top gap (sample)
Payout destination change gating	payouts	Partial	Approval required but not consistently enforced.
Support override policy enforcement	support_ops	Weak	Support actions not correlated to high-risk changes.
Dispute velocity controls	disputes	Partial	Entity-level velocity constraints missing.
Agent tool scope boundaries	agentic_ops	Partial	Tool scopes too broad for sensitive actions.

Remediation Roadmap

Below is a practical remediation sequence aligned to the highest-impact findings. Each item is designed to be retestable using the same pack version to verify improvement.

Timeframe	Focus	Deliverables
0–30 days	Close CRIT gaps in payout changes and support correlation	Approval gating + cooling period; correlation rules; retest payouts pack and validate deltas.
31–60 days	Strengthen device/session linkage and dispute velocity controls	Signal requirements; entity-level velocity; playbooks; retest disputes and onboarding packs.
61–90 days	Agentic Workflow Resilience guardrails (tool scope + policy enforcement)	Tool scopes, policy-eval, approvals, trace completeness; run Agentic Ops Safety pack continuously.

Ongoing cadence (recommended)

Run core packs weekly (or on release), alert on regressions, and publish monthly executive summaries with trend deltas. Update packs and contracts as workflows evolve.

Appendix

Glossary

ARI — Attack Resilience Index — an interpretable 0–100 score summarizing defensive success across decision points.

Decision Contract — A definition of what must be true at a decision point (signals, checks, approvals, constraints).

Lineage — A timeline of events capturing what signals were evaluated and what decisions/actions occurred.

Patch pattern — A reusable defensive remediation approach (gating, step-up, cooldown, correlation, policy enforcement).

Shadow Mode — Non-invasive evaluation using exported signals to assess contract compliance without production integration.

Template notes

Replace sample scores, findings, and pack labels with customer-specific outputs from Adversaia. Maintain versioned runs and export evidence bundles for governance and repeatability.