
 FRAMEWORK · VERSION 1.0

The Continuous Architectural Resilience Framework

An operating model for regulated banks in the era of agentic exploitation

Prepared for: CISOs, CROs and Heads of Information Security

Status: a framework definition, not a certification standard, and not legal or regulatory advice

AUTHORS

John Pellew

john@kolega.ai

Faizan Raza

faizan@kolega.ai

ISSUED BY



Kolega Labs

COMPANION DOCUMENTS

Rings of Risk (working draft, 22 October 2026)
RealVuln Benchmark Comparative Report v3.1

ISSUED

23 October 2026
Framework definition

Executive summary

Most institutions still assure their application estate with the same basic loop: scan on a schedule, raise tickets, fix what surfaces, and attest that the cycle ran. That model made sense when vulnerabilities were found by inspection and exploited one at a time. **Neither condition still holds.**

Autonomous tooling combines small, individually unremarkable weaknesses into paths that cross several code boundaries. The scanners most institutions own cannot see the classes of flaw those paths are built from.

The evidence is no longer anecdotal. In the RealVuln v3.1 benchmark, rule-based scanners found none of 232 broken-access-control findings and none of 173 missing-authentication findings. Semgrep missed 3,728 of 4,138 labelled vulnerabilities (90.1%). In provisional large-application results, a leading general-purpose AI scanner's aggregate recall fell from 80.55% on small applications to 1.77% on systems of 263,000 to 816,000 lines, the size range of a bank's core platforms.

0/232	90.1%	1.77%	246
Broken-access-control findings detected by rule-based scanning	Of labelled vulnerabilities missed by Semgrep	Recall on systems of 263k-816k lines of code	Source-supported routes to B1/B2 outcomes across 15 applications

This framework replaces scan-and-attest with a continuous loop whose **unit of assurance is the reachability of protected outcomes rather than the number of findings**. It has five capabilities (Observe, Validate, Prioritise, Close, Attest) and three rules that separate it from a renamed scanning programme:

- a route is only closed when the exact path has been re-executed and fails;
- prioritisation follows outcomes and chokepoints, not severity scores;
- every assurance statement must cite evidence a third party could reproduce.

- | | | |
|-------|--------------------------------------|--|
| 1 | From findings → routes | A finding describes a weakness on its own. A route describes whether that weakness gives an attacker a position from which another weakness becomes reachable. Two scanners with the same finding recall can leave very different protected outcomes reachable through what they miss. |
| <hr/> | | |
| 2 | From periodic → continuous | A point-in-time scan produces a point-in-time artefact. An exploitable path has a lifetime: it opens when code, configuration or an identity condition changes, and it closes when someone breaks the path. Assurance has to measure that lifetime instead of the date of the last report. |
| <hr/> | | |
| 3 | From asserted → demonstrated closure | A resolved ticket records an intention. A path that has been re-executed and no longer succeeds is evidence. When an examiner asks how you know, the second is the one they can check. |
-

■ SECTION 1

The operating model being replaced

Four structural failures make the inherited model unsafe. Each belongs to the model, not to any vendor's product.

The instrument is blind where chaining happens

Pattern-based detection only covers flaws that reduce to a recognizable syntax. An authorization defect is a missing check across a call path, which gives a pattern matcher nothing to match. The benchmark corpus shows the result: zero detection across 232 access-control findings.

The queue is a false register

Remediation is managed by what was detected. If roughly nine in ten issues never enter the queue, the queue records the visible subset rather than exposure. Presented as the institution's vulnerability position, it is misleading.

Coverage claims are configuration claims

"The scanner is deployed across all repositories" describes installation. The benchmark shows a 69.5-point F3 gap between the leading general-purpose AI scanner and the leading rule-based tool on identical code.

Scale is unaddressed

Tools validated on small fixtures behave very differently on systems of hundreds of thousands of lines. Assurance built on sample-scale evidence is extrapolation presented as measurement.

WHAT THIS FRAMEWORK DOES NOT CLAIM

It does not claim that every path can be enumerated, that any tool finds everything, or that continuous validation replaces secure design, code review or threat modelling. The claim is narrower: **known routes to protected outcomes can be mapped, validated, prioritised and proven closed**. That is a materially better basis for assurance than finding counts and periodic reports.

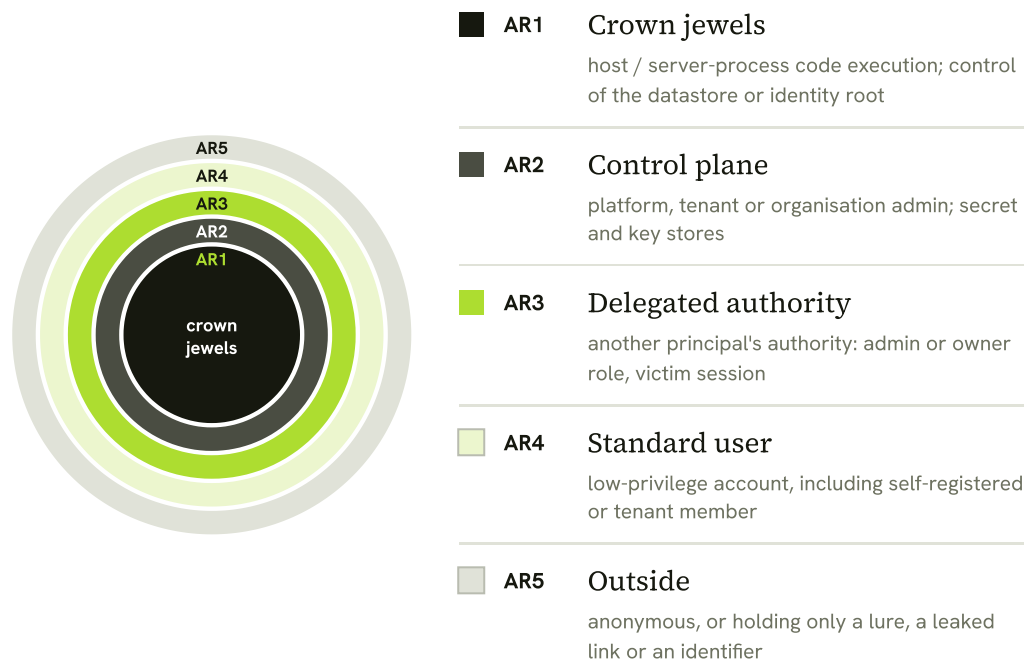
SECTION 2

The architectural vocabulary: Rings of Risk

The framework borrows its language from the Rings of Risk model, which describes how weaknesses in application code combine to reach protected outcomes. Three devices do the work.

Access Rings place an attacker, or a protected asset, by the privilege tier held or required. The rings are numbered from the core outward: AR1 Crown jewels, AR2 Control plane, AR3 Delegated authority, AR4 Standard user, AR5 Outside. A finding is an arrow. Its tail is the position the attacker occupies and its head is the position gained. The arrow's length is what the finding actually gives an adversary, which is often nothing and occasionally everything.

FIGURE 1 The five Access Rings



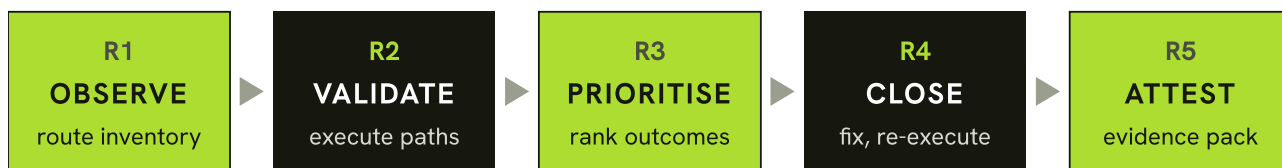
Ring assignment reflects the privilege a capability confers, not a CVSS score. A B1 Critical account takeover can end in AR3 while server-process execution ends in AR1: consequence and position are different measurements. Definitions follow the Rings of Risk working draft, §3.4.

Consequence bands grade the harm of an outcome. **Reach Rings** grade a finding's role in reaching one. The distinction matters in practice. An RR3 finding is not urgent on its own, but it may be the step that completes every route to an AR1 outcome.

CONSEQUENCE BAND: HARM		REACH RING: ROLE IN REACHING IT	
B1	Critical	RR1	Critical reach reaches B1 without another code flaw
B2	Major	RR2	Major reach
B3	Moderate	RR3	Chain-dependent contributes only with another flaw
B4	Limited	RR4	Bounded reach

A finding may appear in several routes, and a route may reach several outcomes, so the model does not aggregate findings into a score. It produces an **outcome profile**. For each protected outcome the profile gives its consequence band, the least demanding entry conditions, the number of distinct supported routes, and the **known-route closure set**: the smallest set of fixes or controls that interrupts every known route to that outcome. That last field is what turns a vulnerability list into a remediation plan.

The framework: five capabilities

FIGURE 2 The continuous loop

one continuous loop: each turn produces the next turn's input

The output of each capability becomes the input to the next, and the loop does not terminate: closure evidence from one turn becomes the baseline for the next.

TABLE 1 Capability definitions, outputs and owners

CAPABILITY	WHAT IT DOES	OUTPUT (EVIDENCE)	ACCOUNTABLE OWNER
R1 Observe	Maintains a current model of the estate: ingress points, identities and privilege tiers, data stores, and the routes between them. Refreshed on change, not only on schedule.	Route inventory with source-anchored references; asset and identity graph; drift log against the last baseline.	Application security engineering, jointly with platform owners.
R2 Validate	Tests whether candidate routes are actually reachable by executing them under authorization in an isolated environment. Separates candidate from confirmed.	Confirmed chains with reproducible proof-of-concept artefacts and recorded preconditions.	Independent validation function (internal or service-operated).
R3 Prioritise	Builds outcome profiles, identifies chokepoints, and computes the closure set for each protected outcome. Orders work by routes removed per unit of effort.	Outcome profiles; chokepoint register; closure sets; re-prioritised backlog.	CISO office with the risk function.
R4 Close	Remediates, then re-executes the exact confirmed path. Closure is recorded only on a failed re-run, not on a merged fix.	Closure record per chain: re-run output, date, operator, residual path if any.	Engineering delivery owner, verified independently.
R5 Attest	Maps evidence to control expectations and produces the statement the institution relies on, at a cadence the regulator's cycle requires.	Control evidence pack: metrics, trends, open residual routes, exceptions with owners.	Compliance and internal audit.

Maturity model

Maturity is defined by what the institution can demonstrate, not by the tooling it buys. Each level is a precondition for the next: an organisation that cannot map routes cannot validate them in any meaningful sense.

TABLE 2 Resilience maturity levels

LEVEL	NAME	DEFINING PROPERTY	CHARACTERISTIC FAILURE
L0 	Ad hoc	Scanning is periodic and tool-driven; no route model; findings triaged individually.	Assurance rests on a clean report; no ability to state exposure.
L1 	Instrumented	Consistent scanning across repositories and frameworks; findings tracked to closure.	Confuses coverage of code with coverage of flaw classes.
L2 	Mapped	Route inventory exists; findings placed in Access Rings; outcome owners named.	Maps are static; drift between the model and the estate goes undetected.
L3 	Validated	Candidate routes are executed and confirmed; POC artefacts exist; closure sets computed.	Validation is a periodic campaign rather than a control; coverage of the estate is partial.
L4 	Continuous	Loop runs on change; every closure is proven by re-execution; residual AR1/AR2 reach is a reported, owned metric.	Residual: assurance depends on the independence and quality of the validation function.

Most institutions that believe they are at L2 are at L1. They have findings and dashboards, but no model of what those findings confer. The gap between L1 and L2 is not closed by buying another tool. It closes when the institution decides to record positions and routes alongside issues.

Operating cadence

The loop runs at different rates at different layers, and cadence follows rate of change: a layer that changes daily cannot be assured quarterly.

TABLE 3 Cadence by layer

LAYER	CADENCE	TRIGGER	EVIDENCE PRODUCED
Application code	Per merge to a protected branch	Commit, dependency change, framework upgrade	Diff-scoped route review; new-route register entries
Identity & authorization	Continuous / weekly	Role, tenancy or policy change	AR3-AR5 position map; access-control coverage report
Integration & API perimeter	Weekly	New endpoint, partner onboarding, schema change	Ingress inventory delta; SSRF and handoff review
Runtime & configuration	Continuous	Deploy, config drift, feature flag change	Configuration-route interactions; enabling-step register
Data layer	Monthly + on change	Schema change, new store, retention change	AR1/AR2 outcome profiles; closure sets
Full validation	Quarterly baseline; event-driven	Major release, M&A, material incident, new business line	Confirmed-chain set; independent validation report

SECTION 6

Metrics that survive an audit

Metrics fail audits for two reasons: they measure activity rather than exposure, and nobody can reproduce them. The framework keeps a small set, defined closely enough that a third party can re-derive them.

TABLE 4 Framework metrics

METRIC	DEFINITION	WHY IT DEFENDS A POSITION
Verified closure rate	Chains closed by a failing re-run ÷ chains confirmed in the period.	Distinguishes remediation intent from demonstrated closure.
Median time-to-verified-close	Median days from confirmation to a failing re-run, by consequence band.	Answers the question regulators ask: how long was it exploitable?
Residual AR1/AR2 reachable outcomes	Count of distinct protected outcomes whose known routes still reach AR1 or AR2, and the number of routes per outcome.	States exposure in consequence terms rather than severity counts.
Closure-set coverage	Protected outcomes with a computed known-route closure set ÷ outcomes in scope.	Proves remediation is planned against routes, not tickets.
Chokepoint concentration	Fraction of routes to a given ring passing through a single finding or circle.	Identifies where one fix removes many paths, and where a single failure opens many.
Detection completeness by family	Benchmarked recall of the deployed detection stack against a labelled corpus, per vulnerability family.	Replaces "we scan everything" with a measured statement about what is visible.

THE ONE METRIC TO RETIRE

~~Total findings open~~, used as a headline assurance number. It rises when detection improves and falls when detection gets worse, so **it moves in the wrong direction as capability changes**. Treat it as a workload indicator and label it that way.

■ SECTION 7

Regulatory mapping

The framework is designed to produce evidence that maps onto existing expectations, not to create a new standard. The table below is indicative rather than exhaustive, and it is not legal advice. Confirm obligations with counsel and through the supervisory relationship in each jurisdiction.

TABLE 5 Indicative mapping of expectations to framework evidence

SOURCE	EXPECTATION (IN SUMMARY)	FRAMEWORK EVIDENCE
DORA Regulation (EU) 2022/2554	Continuous ICT risk management, protection and prevention, detection, and proportionate testing of digital operational resilience, including vulnerability handling and independent testing.	Route inventory and drift log (R1); confirmed-chain register with dates (R2); verified closure records (R4); residual AR1/AR2 reporting (R3, R5).
FFIEC IT Examination Handbook	Ongoing information security assessment, vulnerability and patch management, and independent testing proportionate to risk; evidence of management oversight.	Cadence-by-layer schedule (Table 3); verified closure rate and median time-to-close (Table 4); exception register with named owners.
MAS Technology Risk Management Guidelines	Identification and mitigation of technology risk, security assessments and penetration testing, and board-level oversight of residual risk.	Outcome profiles by consequence band; chokepoint concentration; independent validation report; board-level residual risk statement.

SECTION 8

Implementing in 90 days

Adoption is incremental. The first quarter should not chase full coverage. It should establish the loop on one consequential domain and produce evidence that survives challenge.

TABLE 6 First-quarter path

W1	2	3	4	5	6	7	8	9	10	11	12
PHASE	ACTIVITY					EXIT CRITERION					
 Weeks 1-2 Map	Select one domain with an AR1 or AR2 outcome (payments, customer data, identity root). Build the route inventory and name outcome owners.					A route map exists with source references; a named owner accepts it as accurate.					
 Weeks 3-6 Validate	Validate candidate routes by execution in an isolated environment. Separate confirmed chains from candidates rejected on evidence.					At least one confirmed chain per protected outcome, each with a reproducible artefact; rejected routes carry a stated reason.					
 Weeks 7-10 Prioritise & remediate	Compute closure sets and chokepoints; re-order the existing backlog against routes removed. Remediate the highest-leverage steps first.					Closure sets published for every in-scope outcome; backlog re-prioritised with an audit trail.					
 Weeks 11-12 Close & attest	Re-execute every confirmed chain; record closures and residual reach; present the evidence pack to the risk committee.					A closure record per chain, a residual reach statement, and a cadence committed for the next quarter.					

Five ways this becomes theatre

- × **Mapping without validation.** A route map full of unconfirmed candidates is a hypothesis. Label it as one.
- × **Validation without re-execution.** Closing on a merged fix brings back the gap between intent and evidence that this framework exists to remove.
- × **Closure sets that ignore shared mechanisms.** Twenty variants that run through one vulnerable operation may need a single fix, while three independent routes need three. Counting tickets overstates the work and understates the concentration.
- × **Cadence without change triggers.** A quarterly loop on a layer that changes daily is a sampling exercise reported as continuous assurance.

- × **Independence in name only.** Validation done by the team that must attest to the result is not independent, however rigorous the method.

■ SECTION 9

Limitations and honest scope

The evidence base is a working draft

The Rings of Risk evidence base covers 15 open-source applications. Its route records are source-supported candidates, and the applications were not executed in that study. The framework therefore separates candidate from confirmed at every step. That separation is built into the method, not appended as a caveat.

Validation coverage will be partial

Known-route closure is bounded by the routes you know. Every closure statement should say so: *"every known supported route to this outcome is interrupted."*

Benchmark results do not transfer automatically

RealVuln measures detection on a published corpus. It predicts performance on an institution's estate but does not certify it, so local validation remains necessary.

Aggressive validation carries operational risk

Keep it in isolated environments under written authorization. It does not belong in production monitoring.

This is not a certification

Adoption should be evidenced against the institution's actual obligations, not against this document.

■ SECTION 10

Self-assessment

There are ten questions. Each asks for evidence rather than intent, and each can be answered from records the institution already holds, or from records it does not. **Score one point for every yes backed by an artefact.**

0-2 pts L0-L1	3-5 pts L2 Mapped	6-8 pts L3 Validated	9-10 pts L4 Continuous
------------------	----------------------	-------------------------	---------------------------

TABLE 7 Capability self-assessment

#	QUESTION	ARTEFACT THAT WOULD SATISFY IT
☐ 1	Can you name the protected outcomes that matter most, with a consequence band for each?	Outcome register, approved by the risk function
☐ 2	Does a route inventory link each finding to the position it confers?	Route register with source references
☐ 3	Are findings assigned to Access Rings by a stated rule rather than by CVSS?	Ring classification with judgement calls marked
☐ 4	Is any candidate route validated by execution, not accepted from review alone?	Confirmed-chain register with dates
☐ 5	Does every confirmed chain carry a reproducible artefact?	Re-runnable proof-of-concept per chain
☐ 6	Do you hold closure sets for each outcome: the minimum fixes that interrupt every known route?	Published closure set per outcome
☐ 7	Is closure recorded on a failing re-run, not on a merged fix?	Re-run output attached to each closure
☐ 8	Is the loop triggered by change (merge, deploy, role change) rather than by calendar alone?	Change-trigger log and drift report
☐ 9	Is validation performed independently of the team that attests to the result?	Named independent validator; terms of reference
☐ 10	Can residual AR1/AR2 reachable outcomes be stated as a number today?	Residual reach statement in the last board pack

Most institutions score two or three. The operating model they inherited was never designed to record positions and routes, so the score reflects the model rather than the quality of their engineering. An honest first score is more useful than a confident assertion.

Definitions

Access Ring (AR)	Privilege tier of an attacker or protected asset: AR5 Outside, AR4 Standard user, AR3 Delegated authority, AR2 Control plane, AR1 Crown jewels. Numbered from the core.
Consequence band	Harm grading of an outcome: B1 Critical, B2 Major, B3 Moderate, B4 Limited. Independent of the ring at which the attacker stands.
Reach Ring (RR)	A finding's role in reaching an outcome: RR1 Critical reach, RR2 Major reach, RR3 Chain-dependent, RR4 Bounded reach.
Route	An ordered sequence of findings and steps leading from a stated start position to a protected outcome, under stated conditions.
Confirmed chain	A route executed and reproduced under authorization, with a minimal reproducible artefact and recorded preconditions.
Known-route closure set	The smallest feasible set of fixes or controls that interrupts every known supported route to a given outcome.
Chokepoint	A finding or position through which a disproportionate share of routes to a ring passes.

SOURCES

Rings of Risk: Consequence and Reachability in Application-Code Vulnerability Chains, working draft, 22 October 2026 (github.com/kolega-ai/rings-of-risk).

RealVuln Benchmark Comparative Report, Python & TypeScript/JavaScript stratum, v3.1.0, data generated 11 September 2026 (realvuln.com).

Both are published with open data, and both are subject to revision.

Disclaimer. This framework is provided for information and is not legal, regulatory or investment advice. Regulatory references are indicative summaries; obligations depend on jurisdiction, entity classification and supervisory expectation, and should be confirmed with counsel. Kolega Labs develops security products referenced in the companion documents and may derive commercial benefit from adoption of this framework; readers should weigh it accordingly and test its claims against their own evidence.

© 2026 Kolega Labs. Framework v1.0, issued 23 October 2026.