BLADE-FINANCE Governance Node

Authority Governance for Financial-Sector AI Decision Systems Under the Treasury Financial Services AI Risk Management Framework

Burak Oktenli

Georgetown University, M.P.S. Applied Intelligence | ORCID: 0009-0001-8573-1667

Version 1.0 | May 2026 | Zenodo Research Paper | DOI: 10.5281/zenodo.20374692

Keywords: authority-governed autonomy, financial-sector AI governance, runtime assurance, Treasury FS AI RMF, NIST AI RMF, Executive Order 14179, economic security, Hierarchical Multi-Attribute Authority, HMAA, SHA-256 evidence chain, canonical-form serialization, deepfake authentication attack, AI-agent coordinated attack, retrospective swarm review, stigmergic ensemble, escalation-delta, triage metrics, Wilson interval, AUTHREX, BLADE, Kria K26, NVIDIA L4, YubiHSM 2, FIPS 186-5, ECDSA P-256

1. Zenodo Deposit Metadata

Field	Value
Title	BLADE-FINANCE Governance Node: Authority Governance for Financial-Sector AI Decision Systems Under the Treasury Financial Services AI Risk Management Framework
Author	Burak Oktenli | Independent Researcher, AUTHREX Systems; Georgetown University, M.P.S. Applied Intelligence | ORCID: 0009-0001-8573-1667
DOI / License / Version	10.5281/zenodo.20374692 | CC BY 4.0 | v1.0
Description	Simulation-validated, software-enforced authority-arbitration reference architecture for financial-sector AI decision systems, aligned to the U.S. Treasury Financial Services AI Risk Management Framework (FS AI RMF, released February 19, 2026) under the implementation framing of Executive Order 14179. An eight-stage AUTHREX pipeline with a four-tier HMAA authority model, a population-state coordination model across account, device, payee, and IP-cluster history, a SHA-256 canonical-form evidence chain, and a retrospective stigmergic swarm-review module that recovers coordinated rings the per-transaction path clears. Reference authority node: 36 hardware components, 33 electrical connections, 32 mechanical connections, approximately $9,228 BOM. Monte Carlo validation over 2,000 seeded trials per the headline scenario; nine embedded self-tests pass. TRL 3-4 (simulation) / TRL 2 (hardware); synthetic data only; no deployment in any financial institution.
Hardware	36 components / 33 electrical / 32 mechanical / approximately $9,228 BOM
Website	burakoktenli.com
Project Page	burakoktenli.com/blade-finance
Simulation	burakoktenli.com/blade-finance-simulation
Companion	AUTHREX → Treasury FS AI RMF Crosswalk (BLADE-FINANCE Working Paper No. 1)
Related	SATA: zenodo.18936251 / HMAA: zenodo.18861653 / CARA: zenodo.18917790 / FLAME: zenodo.19015618 / ADARA: zenodo.19043924 / MAIVA: zenodo.19015517 / BLADE-EDGE: zenodo.19177472 / BLADE-AV: zenodo.19232130 / BLADE-MARITIME: zenodo.19246785 / BLADE-INFRA: zenodo.19277887 / BLADE-SPACE: zenodo.20183269 / BLADE-CUAS: zenodo.20299604

Table 1: Zenodo deposit fields.

2. Contents of This Deposit

File	Description
blade-finance-governance-node.pdf	This research paper with embedded figures, governance equations, simulation results, and full reference list.
blade-finance-simulation.html	Interactive browser-native simulator (v2.2): six scenarios including free-play; deterministic mulberry32 PRNG; synchronous SHA-256 evidence chain over canonical-form (sorted-key, fixed-precision) serialization; schema validation; triage metrics with Wilson 95% intervals; Monte Carlo runner; external-dataset benchmark; golden-trace export; nine in-browser self-tests; and a retrospective stigmergic swarm-review module. Runs entirely client-side and offline.
blade-finance_authority_node_PARTS.csv	36-component bill of materials (approximately $9,228) with manufacturer URLs and quantities.
blade-finance_authority_node_ELECTRICAL_CONNECTIONS.json	33 electrical connections with protocol annotations, voltages, and currents.
blade-finance_authority_node_MECHANICAL_CONNECTIONS.json	32 mechanical connections with fastener and mount detail.
blade-finance_authority_node_CONFIG.json	Node definitions with pin arrays and provenance metadata.
blade-finance_authority_node_SCHEMATIC.svg	Vector wiring schematic, color-coded by subsystem.
blade-finance_authority_node_render.svg	Reference-node block diagram (dual-plane 1U appliance layout).

Table 2: Deposit file inventory.

3. Abstract

This paper presents the BLADE-FINANCE Governance Node, a simulation-validated, software-enforced authority-arbitration reference architecture for financial-sector AI decision systems operating under the policy framework established by the U.S. Treasury Financial Services AI Risk Management Framework (FS AI RMF), released February 19, 2026, and the implementation framing of Executive Order 14179. The eight-stage AUTHREX pipeline (VALIDATE, SATA, ADARA, MAIVA, HMAA, FLAME, ERAM, CARA) appends every decision to a SHA-256 evidence chain (the AUDIT stage). SATA scores per-input integrity and trust; ADARA scores adversarial, deepfake, and AI-agent attack indicators; a population-state model computes a coordination score across account, device, payee, and IP-cluster history; MAIVA composes a multi-signal consensus; HMAA implements a four-tier authority model (T3 / T2 / T1 / T0) that routes a transaction to autonomous clearance, supervised review, elevated confirmation, or manual hold; FLAME enforces a tier-dependent deliberation window; ERAM is an independent escalation-risk gate; and CARA coordinates recovery with a tamper-evident audit ledger. A retrospective stigmergic swarm-review module operates over a long look-back window of the ledger and recovers coordinated “low-and-slow” rings that the per-transaction path has already cleared, reporting an explicit escalation-delta. The companion simulator (v2.2) is deterministic (mulberry32 PRNG, real SHA-256 over

canonical-form serialization validated byte-exact against the FIPS 180-4 reference vector) and ships nine in-browser self-tests, a Monte Carlo runner with Wilson confidence intervals, and an external-dataset benchmark. Across a 2,000-trial Monte Carlo run on the realistic mixed stream, the pipeline produced zero deliberation-window breaches, a benign false-review rate of 0.013 (95% CI 0.8%-1.9%), and a triage F1 of 0.957; the reported recall is an *actionable-risk triage* measure, not an empirical fraud-detection rate. The reference authority node comprises 36 hardware components, 33 electrical connections, 32 mechanical connections, and approximately $9,228 BOM. The work is published as fundamental research under CC BY 4.0. It is a reference architecture at TRL 3-4 (simulation) and TRL 2 (hardware); it has not been deployed in any financial institution, and all data are synthetic.

4. Introduction

4.1 Motivation

Financial-sector adoption of artificial intelligence, in fraud screening, authentication, payments, and increasingly in autonomous and agentic transaction systems, has outpaced the assurance machinery that governs when an automated decision may proceed without a human. On February 19, 2026, the U.S. Department of the Treasury released the Financial Services AI Risk Management Framework (FS AI RMF) [1], a sector-specific operationalization of the NIST AI Risk Management Framework [5] developed in coordination with the Cyber Risk Institute, the Financial Services Sector Coordinating Council, and approximately 108 financial institutions. The framework establishes 230 control objectives organized under four NIST functions (Govern, Map, Measure, Manage) and four progressive maturity stages (Initial, Minimal, Evolving, Embedded). Treasury simultaneously released a companion AI Lexicon [2], established the Artificial Intelligence Transformation Office (AITO), and, in coordination with the Financial Stability Oversight Council, launched an AI Innovation Series whose first roundtable convened March 4, 2026 [3]. The Treasury Secretary's framing, that "leadership in AI adoption is a crucial component of economic security", anchors this work's policy posture under Executive Order 14179, *Removing Barriers to American Leadership in Artificial Intelligence* [4].

Commercial fraud and authentication products compete on detection metrics: probability of detection, false-positive rate, and model accuracy. None of these primitives provides what an FS-AI-RMF-aligned environment requires: a governance layer that arbitrates *how much authority an automated system should be granted* for a given decision, computes whether the resulting decision trace is auditable, and routes tier-appropriate decisions to credentialed reviewers. BLADE-FINANCE shares the AUTHREX governance pipeline with the other members of the BLADE family: BLADE-EDGE (defense), BLADE-AV (autonomous vehicles), BLADE-MARITIME, BLADE-INFRA (critical infrastructure), BLADE-SPACE (orbital autonomy), and BLADE-CUAS (counter-UAS). The finance variant adapts that shared core along four axes: input-integrity scoring over transaction features rather than physical sensors; a population-state coordination model over account, device, payee, and IP-cluster history; a four-tier authority model encoding autonomous clearance, supervised review, elevated confirmation, and manual hold; and a retrospective swarm-review capability for coordinated rings. The governance mathematics are shared across variants, which supports cross-domain architectural portability.

4.2 Scope and Contributions

- An eight-stage AUTHREX pipeline (VALIDATE, SATA, ADARA, MAIVA, HMAA, FLAME, ERAM, CARA) with a SHA-256 canonical-form evidence chain appended at every decision.

- A population-state coordination model that scores account velocity, payee and device concentration, IP burst, and amount-spike behavior over rolling windows, the primary defense against single-transaction false positives.
- A four-tier HMAA authority model (T3 / T2 / T1 / T0) that maps confidence and risk to autonomous clearance, supervised review, elevated confirmation, or manual hold.
- A retrospective stigmergic swarm-review module that recovers coordinated low-and-slow rings the per-transaction path cleared, reporting an explicit escalation-delta. The honest claim is ensemble agreement and ring detection, *not* a Byzantine-fault-tolerance guarantee.
- A deterministic, offline, byte-reproducible reference simulator (v2.2) with nine self-tests, Monte Carlo validation with Wilson intervals, an external-dataset benchmark, and golden-trace export.
- A reference authority node: a dual-plane 1U design (NVIDIA L4 inference plane, Xilinx Kria K26 governance plane) with a YubiHSM 2 hardware security module, 36 components at approximately $9,228 BOM.
- Alignment to the Treasury FS AI RMF, NIST AI 100-1, and EO 14179, with an explicit honest scope: a methodology and reference architecture, not an implementation report.

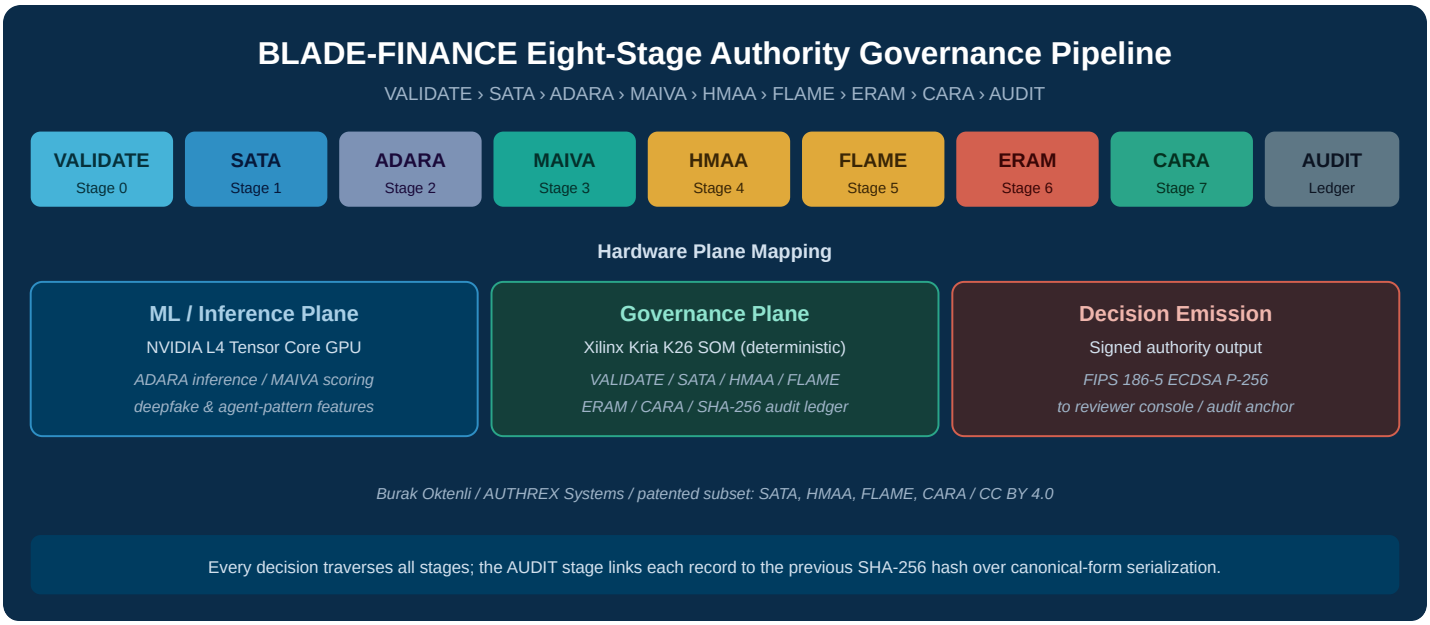

Figure 1: BLADE-FINANCE eight-stage authority governance pipeline with hardware plane mapping. The ML plane (NVIDIA L4) executes ADARA inference and MAIVA scoring under the authority of the governance plane (Kria K26 SOM), which holds the AUTHREX pipeline core and the audit ledger. The system emits only signed authority decisions to downstream reviewer consoles; it executes no transactions itself.

5. Threat Model

Threats are drawn from the financial-sector AI attack surface enumerated in the FS AI RMF, the control families of NIST SP 800-53 Rev. 5, and the AI-agent and synthetic-identity literature. Table 3 covers the governance-layer attack surface; upstream model-quality threats (training-data poisoning of vendor fraud models, raw feature integrity) remain the responsibility of the institution's detection stack.

Threat	Effect	Governance response
Deepfake authentication	Synthetic voice or video defeats step-up auth	ADARA deepfake-artifact scoring; SATA provenance and device-match trust; MAIVA requires corroborating signals before a non-T3 release
AI-agent coordinated attack	Botnet of agents issues many individually-plausible transactions	Population-state coordination score (payee, device, and IP concentration, velocity); HMAA defers to review; CARA recovery on misclassification
Low-and-slow mule ring	Sub-threshold legs to a shared mule payee clear the fast path	Retrospective swarm review over a long look-back window surfaces the ring; reported as escalation-delta
High-value transaction with risk signals	Large transfer with partial anomaly indicators	ERAM independent escalation gate; HIGH risk forces elevated confirmation even if HMAA would clear
Operator credential compromise	Stolen reviewer credentials escalate authority	HSM-bound credentials (YubiHSM 2); revocation; provisioning attestation
Audit ledger tamper	Online edit of historical decision records	SHA-256 prev-hash chain detects modification; full re-verification on export; periodic external anchor
Adversarial evasion on model features	Crafted inputs evade the ADARA classifier	Out of governance scope; population-state model provides a non-model-dependent corroboration path. Future work.

Table 3: Governance-layer threat model. The last row identifies a threat beyond current scope; model-robustness hardening is future work.

6. Governance Architecture

6.1 Pipeline

Table 4 enumerates the eight pipeline stages plus the audit append, with plane assignment, function, and provisional-patent reference where applicable.

Stage	Module	Function	Plane / Patent
0	VALIDATE	Schema and range validation of the transaction record; malformed inputs are rejected before scoring	Governance plane
1	SATA	Per-input integrity and trust score τ from device match, geo consistency, and provenance	Governance plane (Prov. 64/002,453)
2	ADARA	Adversarial, deepfake, and AI-agent indicator scoring; novel-payee flag	ML plane (patent pending)
3	MAIVA	Multi-signal consensus combining trust, adversarial, and coordination evidence	ML plane
4	HMAA	Four-tier authority arbitration (T3 / T2 / T1 / T0)	Governance plane (Prov. 63/999,105)
5	FLAME	Tier-dependent deliberation window; contracts under burst density	Governance plane (Prov. 64/005,607)
6	ERAM	Independent escalation-risk gate: severity × consequence × certainty	Governance plane
7	CARA	Coordinated recovery; state revert on misclassification with signed audit entry	Governance plane (Prov. 64/000,170)
,	AUDIT	SHA-256 prev-hash ledger append over canonical-form serialization	Governance plane

Table 4: Eight-stage authority-governed pipeline plus audit append. Patent numbers refer to U.S. provisional applications filed March 2026.

6.2 Trust, Adversarial, and Coordination Scoring

SATA computes a per-input trust score from weighted integrity features, bounded to [0,1]:

$$\text{sataTrust} = \text{clamp}(0.40 \cdot \text{deviceMatch} + 0.30 \cdot \text{geoConsistency} + 0.30 \cdot \text{provenance} , 0, 1)$$

Eq. (1): SATA per-input trust score.

A population-state model maintains rolling windows over account, device, payee, and IP-cluster history and composes a coordination score that rises when activity concentrates, the structural signature of a coordinated attack that no single transaction reveals:

$$\text{coordinationScore} = \text{clamp}(0.28 \cdot \text{payeeConc} + 0.26 \cdot \text{deviceConc} + 0.24 \cdot \text{ipBurst} + 0.12 \cdot \text{acctVelocity} + 0.10 \cdot \text{amountSpike} , 0, 1)$$

Eq. (2): Population-state coordination score over rolling windows.

ADARA combines independent attack indicators under a noisy-OR composition, and the adversarial score weights that composition against trust and coordination:

$$\text{pAttack} = 1 - (1 - \text{deepfakeArtifacts})(1 - \text{agentPattern})(1 - \text{coordinationScore})$$

Eq. (3): Noisy-OR attack probability across independent indicators.

$$\text{adv} = \text{clamp}(0.78 \cdot \text{pAttack} + 0.14 \cdot (1 - \text{sataTrust}) + 0.08 \cdot \text{coordinationScore} , 0, 1)$$

Eq. (4): ADARA adversarial score.

MAIVA composes a suspicion score that gates the HMAA classification; a novel-payee flag contributes a bounded increment so that provenance novelty cannot dominate the decision:

```
suspicion = clamp( 0.42·adv + 0.25·(1 − sataTrust) + 0.23·coordinationScore + 0.10·(novelPayee ? 0.4 : 0) , 0, 1 )
```

Eq. (5): MAIVA suspicion score gating HMAA.

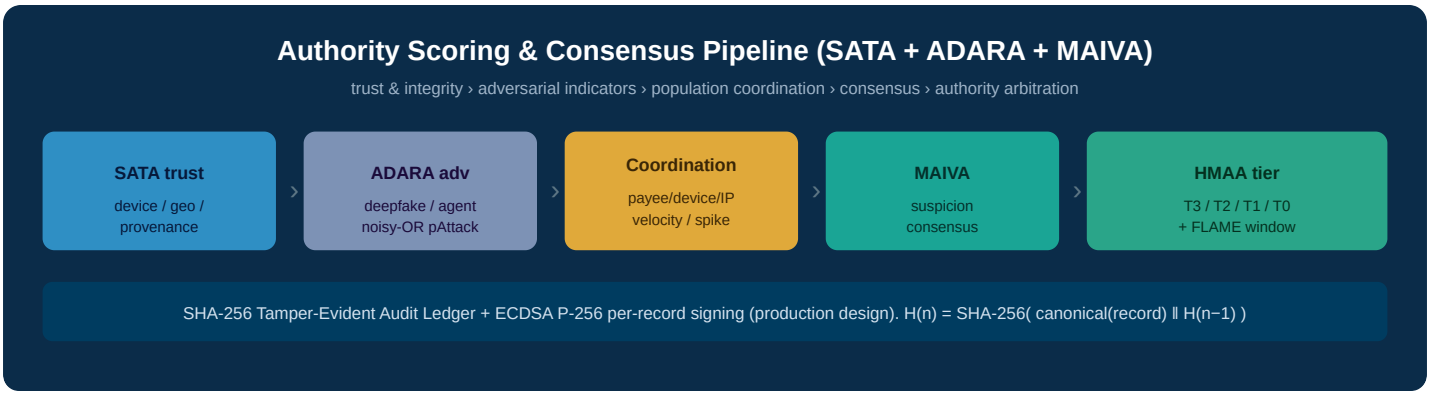

Figure 2: Authority scoring and consensus pipeline. Trust, adversarial, and coordination evidence compose a MAIVA suspicion score that gates HMAA tier arbitration; every decision is appended to the SHA-256 audit ledger.

6.3 HMAA Authority Tiers

HMAA maps the MAIVA suspicion score and the ERAM risk gate to a four-tier authority decision. Below the active tier threshold the decision defers to a credentialed reviewer within the FLAME window; on window expiration without action the system holds rather than auto-clearing.

Tier	Name	Reviewer class	Authority
T3	Autonomous clear	None (autonomous)	Transaction cleared and logged
T2	Supervised review	Operations / fraud analyst	Analyst CONFIRM within FLAME window
T1	Elevated confirmation	Senior / risk officer	Explicit elevated confirmation required
T0	Manual hold	Incident / suspected compromise	Full human control; transaction held

Table 5: Four-tier HMAA authority bindings. Tier transitions occur only on (a) reviewer action, (b) MAIVA classification change, or (c) policy-defined trigger.

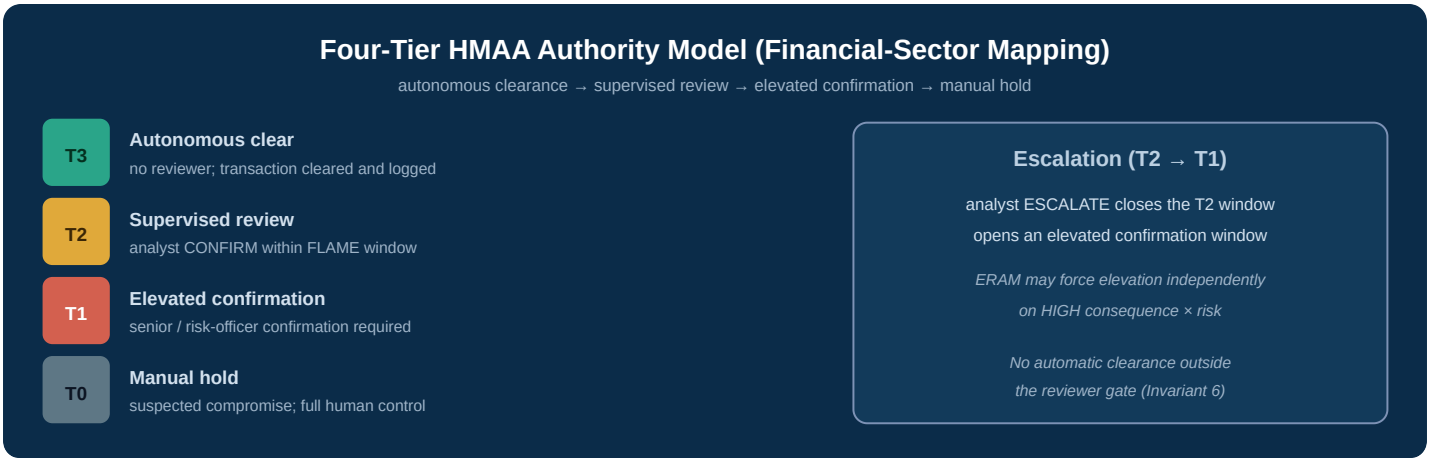

Figure 3: Four-tier HMAA authority model. On an analyst ESCALATE action at T2, the system closes the supervised-review window and opens an elevated-confirmation window. ERAM operates independently and can force elevation on HIGH consequence even when HMAA would clear.

6.4 FLAME and ERAM

FLAME enforces tier-dependent deliberation: a reviewer must act within the active window; the window contracts under burst density so that a flood of simultaneous decisions cannot exhaust deliberation time. ERAM (Engagement Risk Assessment Model) is independent of HMAA: it composes a risk score from classification severity, transaction consequence (a function of amount and exposure), and recent reviewer-certainty history, and forces an elevated-confirmation level when risk is HIGH, even if HMAA alone would clear. This separation guarantees that a high-consequence transaction with only partial anomaly signals still receives human attention.

6.5 CARA Recovery

Phase	Trigger	Action
Signal revalidation	Feature dropout or trust drift	Re-poll inputs; recompute SATA trust and coordination score
Partial restoration	Trust recovering on multiple signals	Resume monitoring; hold tier at the supervised-review ceiling
Full recovery	Sustained trust above threshold	Restore to autonomous clearance; emit signed recovery audit entry
Manual override	T0 incident intervention	Authenticated reset; audit entry logged with credential identity

Table 6: CARA recovery phases. State transitions are logged to the SHA-256 prev-hash chain.

6.6 Retrospective Swarm Review

The per-transaction path is, by construction, myopic: its short rolling windows cannot accumulate the evidence of a deliberately sparse, “low-and-slow” ring, many individually-benign legs sharing a mule payee or botnet IP, spread thinly across accounts so that no short window crosses the coordination threshold. These legs clear the fast path. The retrospective swarm-review module addresses this gap. On demand, it spawns an ensemble of lightweight agents over a long look-back window of the ledger (default 300 records, 24 agents). Each agent independently bootstrap-resamples the window and deposits evidence into a shared, stigmergic map keyed by payee, device, and IP-cluster; concentration that no single transaction revealed accumulates into a trail. A key is

reported as a confirmed ring only when at least a two-thirds quorum of agents independently flag it. The module then reports an *escalation-delta*: the transactions the fast path cleared that belong to a confirmed ring.

Two design boundaries keep this defensible. First, the claim is *ensemble agreement and ring detection*, not Byzantine fault tolerance: there is no quorum-intersection safety bound, and the bootstrap quorum is a stability filter over resamples of one rule, not a vote among independent models. Second, the capability is demonstrated on the failure mode at a realistic base rate (the low-and-slow scenario), measured as rings surfaced and the marginal recovery of actionable transactions, not as another near-unity score on a single-class scenario. The module is deterministic and runs against the existing SHA-256 ledger.

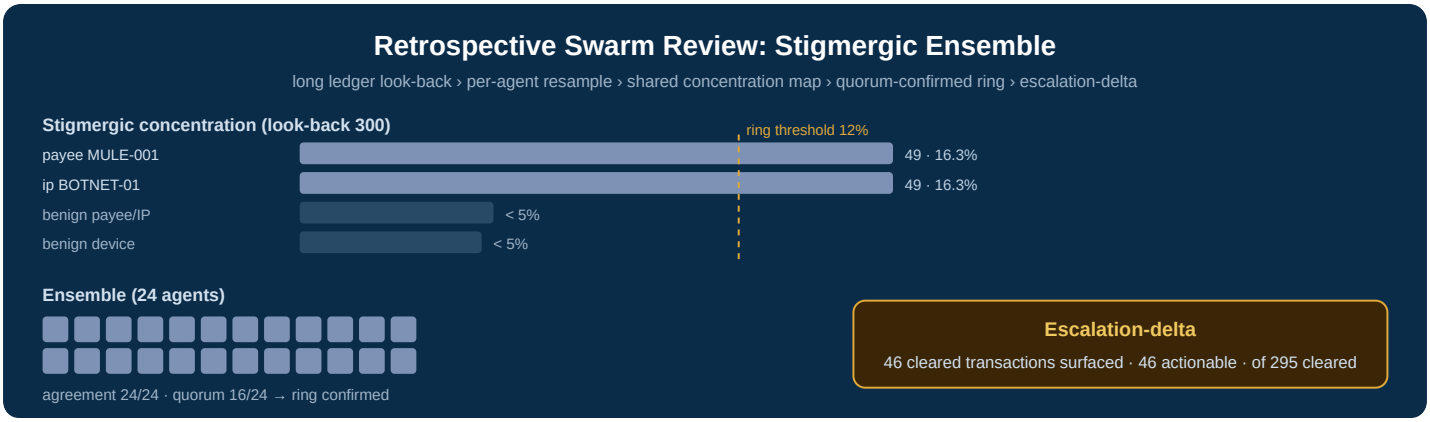

Figure 4: Retrospective swarm review on the low-and-slow scenario. Concentration accumulates on the shared mule payee and botnet IP (both above the ring threshold) while benign and spread keys stay below it; 24 of 24 agents agree, and 46 transactions the fast path cleared are surfaced, all 46 actionable. Result is deterministic.

6.7 Design Invariants

- **Inv 1:** No authority decision is emitted without traversing all pipeline stages.
- **Inv 2:** Below the active tier threshold, the decision defers to a credentialed reviewer within the FLAME window.
- **Inv 3:** ERAM can force elevated confirmation independently of HMAA on HIGH consequence.
- **Inv 4:** CARA recovery phases are mutually exclusive at any point in time.
- **Inv 5:** Every decision is recorded in the SHA-256 prev-hash audit chain; each record links to the prior hash over canonical-form serialization.
- **Inv 6:** On FLAME window expiration without reviewer action, the system holds; it never auto-clears outside the reviewer gate.
- **Inv 7:** The retrospective swarm review is read-only over the ledger and does not alter live decisions; it produces an analyst-facing escalation-delta.

7. Reference Authority Node

The reference platform is a dual-plane 1U rack-mount node (19-inch EIA-310) with redundant 600 W power. An NVIDIA L4 Tensor Core GPU serves the ML/inference plane (ADARA and MAIVA scoring); a Xilinx Kria K26 System-on-Module serves the deterministic governance plane (VALIDATE, SATA, HMAA, FLAME, ERAM, CARA, and the audit ledger). An Intel Xeon-D 1747NTE host with 64 GB ECC memory and dual 480 GB NVMe SSDs in RAID-1 provides the platform substrate. Cryptographic operations use a YubiHSM 2 in a FIPS 140-2 Level 3 tamper-evident potted enclosure together with an Infineon SLB 9670 TPM 2.0; networking is via dual 10GbE

SFP+ for transactions and isolated 1GbE management. The reference configuration is 36 components, 33 electrical connections, 32 mechanical connections, and approximately $9,228 BOM.

Subsystem	Component	Role
Governance plane	Xilinx Kria K26 SOM	Deterministic AUTHREX pipeline; audit ledger; ECDSA signing supervision
ML / inference plane	NVIDIA L4 Tensor Core GPU	ADARA inference; MAIVA scoring
Host	Intel Xeon-D 1747NTE + 64 GB ECC + 2×480 GB NVMe RAID-1	Platform host; storage
Crypto	YubiHSM 2 (FIPS 140-2 L3 potted) + Infineon SLB 9670 TPM 2.0	ECDSA P-256 signing (FIPS 186-5); credentials; key custody
Network	Dual 10GbE SFP+ + isolated 1GbE management	Transaction ingress; out-of-band management
Power / cooling	Redundant 600 W 80-Plus Platinum + 6× 40 mm N+1 hot-swap fans	Redundant power and thermal

Table 7: Key reference-node subsystems (36 components total).

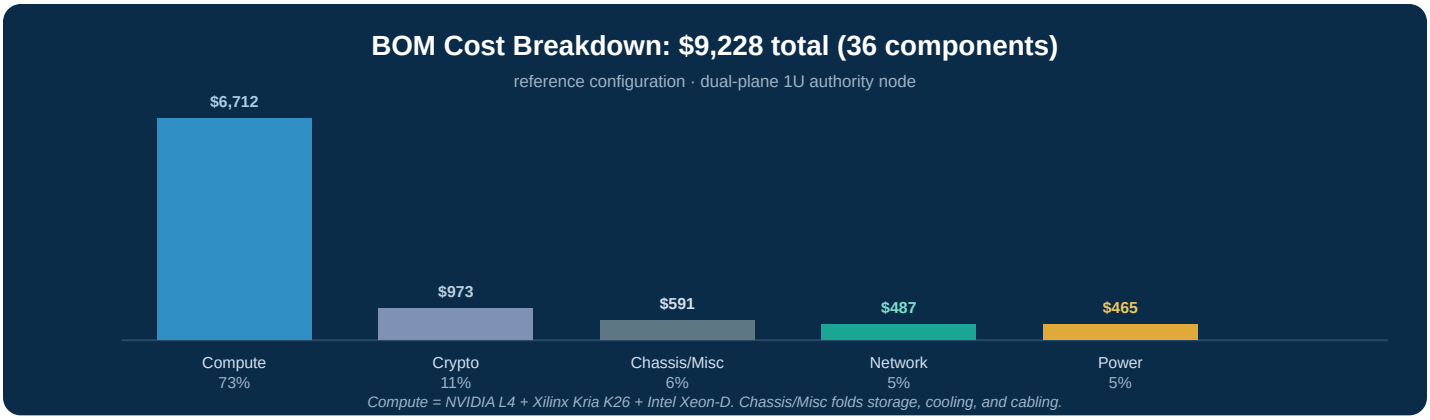

Figure 5: BOM cost breakdown by subsystem. Total approximately $9,228 at the reference configuration; compute (GPU, FPGA, host CPU) dominates at 73%.

8. Related Work

Commercial fraud, authentication, and transaction-monitoring products detect and alert; their sensor fusion is typically vendor-internal and proprietary, and they do not enforce a software-gated authority layer over whether an automated decision may proceed. BLADE-FINANCE operates at a complementary authority-arbitration layer designed to wrap, not replace, existing detection investments. Relative to the Simplex monitor-actuator paradigm, which switches between a verified safe controller and an unverified complex one, BLADE-FINANCE replaces binary switching with a continuous four-tier authority spectrum, adds a population-state coordination model, and adds deliberation windows that prevent cascading clearances even when authority conditions are met.

Feature	Simplex	Commercial fraud tools	This work
Continuous multi-tier authority	Binary switch	,	Yes; four-tier HMAA
Population coordination model	,	Vendor-proprietary	Yes; account/device/payee/IP
Cascade prevention	,	,	Yes; FLAME window
Retrospective ring recovery	,	Limited	Yes; swarm escalation-delta
Tamper-evident decision trace	,	,	Yes; SHA-256 + ECDSA
FS AI RMF alignment	,	,	Yes; function-domain-stage crosswalk
Open architecture spec	N/A	Proprietary	Yes; CC BY 4.0; ~$9,228

Table 8: Architectural comparison. Commercial tools detect and alert; BLADE-FINANCE adds continuous software-gated multi-tier authority with a tamper-evident decision trace.

9. Simulation Methodology and Results

9.1 Simulator Architecture

The companion reference simulator (blade-finance-simulation.html, v2.2) implements the full eight-stage pipeline in browser-native JavaScript. Reproducibility is provided by a seeded mulberry32 PRNG (no use of Math.random) and a synchronous SHA-256 hash chain computed over a canonical-form (sorted-key, fixed-precision) state serialization that verifies byte-identically across engines; the SHA-256 implementation was validated byte-exact against the FIPS 180-4 reference vector. Inputs are schema-validated before scoring, and malformed records are rejected rather than scored. The simulator separates four false-positive-relevant outcomes, cleared, review, blocked, and invalid, and reports triage precision, recall, and a benign false-review rate with Wilson 95% confidence intervals; a Monte Carlo runner repeats a scenario over N seeded trials; an external-dataset benchmark scores the pipeline against an externally-supplied label field that the detection path never reads; and a golden-trace export fixes an expected checksum and confusion matrix for regression. Nine in-browser self-tests cover the SHA-256 reference vector, canonical-JSON order-invariance, determinism, the population model, schema validation, ledger tamper detection, the external benchmark, and the retrospective swarm review. The ledger continuity check is maintained incrementally in O(1) per record, with full re-verification on export. The simulator validates the logical decision behavior of the governance algorithms; production cryptography (per-record ECDSA P-256, HSM custody) and hardware timing are specified in the design and are not exercised in the browser.

An important honesty note on the metric definition: because legitimate high-value transactions are deliberately counted as *operationally risky* (they should escalate), the recall figure measures *actionable-risk triage*, whether the pipeline routes transactions that warrant attention to review, and is *not* an empirical fraud-detection rate. The bypass condition (governance disabled) is reported alongside to show the governance lift.

9.2 Results

Table 9 reports a 2,000-trial Monte Carlo run on the realistic mixed stream (scenario MIX, seed 42, governance ON). All values are produced by the deposited simulator.

Metric	Value (governance ON)	Notes
Trials	N = 2,000 (seed 42)	scenario MIX
Triage split	cleared 86.8% / review 11.9% / blocked·invalid 1.4%	,
FLAME deliberation breaches	0	Inv 6 upheld
Precision (flagged actionable)	0.917	,
Recall (actionable-risk triage)	1.000 · 95% CI 98.4%-100.0%	triage, not fraud detection
Benign false-review rate	0.013 · 95% CI 0.8%-1.9%	Wilson interval
Triage F1	0.957	,
Recall (governance BYPASS)	0.000	governance lift baseline
Self-tests	9 / 9 pass	deterministic suite
Determinism (1,000 steps, seed 42)	head 7d5aaab4… · checksum 0x06acd6be	golden value

Table 9: Monte Carlo and reproducibility results for scenario MIX. Recall is an actionable-risk triage measure, not an empirical fraud-detection rate.

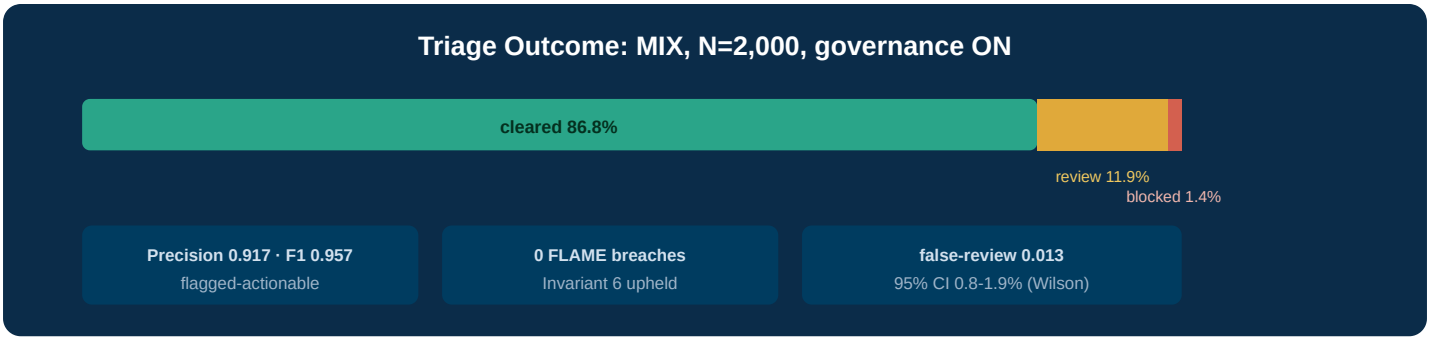

Figure 6: Triage outcome and headline metrics for the 2,000-trial mixed-stream run. The governance pipeline routed 11.9% of the stream to review and 1.4% to block/invalid with zero deliberation-window breaches.

9.3 Retrospective Swarm Review Results

On the low-and-slow scenario, the per-transaction path cleared the ring legs as designed: each leg presented benign per-transaction features and the short runtime window never accumulated the concentration. A single retrospective swarm review over the most recent 300 ledger records, with 24 agents and a 16/24 quorum, confirmed two rings, the shared mule payee and botnet IP, each at 24/24 agreement, and surfaced 46 cleared transactions as part of those rings, all 46 of which were actionable (escalation-delta = recovered = 46 of 295 cleared). The spread device dimension correctly did not cross the ring threshold, demonstrating that the module flags the genuinely-shared infrastructure rather than benign variation. The result is reproducible on the deposited seed.

9.4 Statistical Analysis

Triage intervals are reported as Wilson score 95% confidence intervals, which are well-behaved at proportions near 0 and 1 where the normal approximation fails. The reported recall interval (98.4%-100.0%) and benign false-review interval (0.8%-1.9%) are computed over the 2,000-trial run. The mulberry32 generator is adequate for the N reported here; high-N expansion beyond this range would warrant a wider-state generator, as used in companion BLADE deposits. Zero observed deliberation breaches is necessary but not sufficient evidence for a safety-relevant governance property; these bounds do not exclude rare modes at this sample size, and hardware-in-the-loop validation with operational data is required to tighten them.

10. Known Limitations and Future Work

Table 10 doubles as a threats-to-validity register: each row names a limitation, classifies it, and states the path that would retire it. The two that bound the contribution most are the absence of real transaction data and the simulation-only maturity; both are stated plainly rather than hedged.

Limitation	Category	Impact / path
Simulation-only validation	Scientific	No real transaction data; synthetic scenarios only. Path: hardware-in-the-loop with institutional data under appropriate agreement.
TRL 3-4 simulation / TRL 2 hardware	Engineering	No prototype built. Reference architecture; prototype is a post-deposit activity.
Production cryptography not exercised	Engineering	Simulator uses real SHA-256 over canonical form; per-record ECDSA P-256 and HSM custody are design-specified, not run in-browser.
Recall is triage, not fraud detection	Scientific	Metric counts actionable-risk routing; an empirical fraud-detection rate requires labeled operational data.
Swarm review is ensemble, not BFT	Scientific	Claims ring detection and escalation-delta; no quorum-intersection safety bound is claimed.
Scoring weights are synthetic	Scientific	Default weights are author-assigned; formal elicitation is future work.
Adversarial robustness of features	Security	Model-evasion hardening is out of governance scope and is future work.

Table 10: Limitations and mitigation paths.

10.1 Implementation Status

Component	Status	Evidence basis
HMAA four-tier authority	Implemented in simulator	Eqs. (1)-(5) across scenarios
Population-state coordination model	Implemented in simulator	Self-test; scenario C/E
FLAME deliberation window	Implemented in simulator	0 breaches over 2,000 trials
CARA recovery	Implemented in simulator	Scenario D behavior
SHA-256 audit chain	Implemented (real SHA-256, byte-exact)	Self-test vs FIPS 180-4 vector
Retrospective swarm review	Implemented in simulator	Self-test; escalation-delta on scenario E
Reference node (36 components)	Specified in design docs	BOM, electrical/mechanical JSON, schematic
Per-record ECDSA / HSM custody	Design-specified, not built	Production reference design
Operational data integration	Not tested	Planned HIL

Table 11: Implementation status as of v1.0 (paper) / v2.2 (reference simulation).

10.2 Notation

Symbol	Meaning	Domain
sataTrust	SATA per-input integrity and trust score	[0,1]
coordinationScore	Population-state coordination over rolling windows	[0,1]
pAttack	Noisy-OR attack probability across indicators	[0,1]
adv	ADARA adversarial score	[0,1]
suspicion	MAIVA suspicion score gating HMAA	[0,1]
escalation-delta	Cleared transactions surfaced as a confirmed ring	count

Table 12: Notation reference.

11. Regulatory Alignment

Treasury FS AI RMF [1] establishes 230 control objectives under four NIST functions and four maturity stages. BLADE-FINANCE control surfaces map to these functions and stages through the companion crosswalk (BLADE-FINANCE Working Paper No. 1); this paper does not claim a line-item mapping of the 230 individual control objectives, which requires the FS AI RMF Control Objective Reference Guide distributed through the Cyber Risk Institute. **NIST AI 100-1** [5] is the foundational framework the FS AI RMF operationalizes. **Executive Order 14179** [4] frames the policy environment. The evidence-chain design (canonical-form SHA-256 ledger, design-specified ECDSA P-256 per-record signing under FIPS 186-5, HSM key custody) supports auditability and records-integrity objectives. The work is positioned under the Treasury framing that leadership in AI adoption is a component of economic security; it is offered as fundamental research and a public contribution to the financial-sector AI-governance literature.

12. Data Availability

All artifacts are deposited under CC BY 4.0 at DOI 10.5281/zenodo.20374692, with no access restrictions. The deposit includes this research paper (PDF), the interactive simulator (HTML, client-side and offline), the reference-node specification files (CONFIG, ELECTRICAL, MECHANICAL in JSON), the bill of materials (CSV), and the vector wiring schematic (SVG) with a reference-node block diagram (SVG). This is fundamental research; it contains no controlled technical data and no information from any financial institution.

Reproducibility. The reported results are produced by the deposited simulator (v2.2), which is deterministic and runs offline in any modern browser with no network access. To reproduce: open blade-finance-simulation.html, run the in-browser Self-Test (expect 9 of 9 pass), then run the Monte Carlo control on the mixed-stream scenario at seed 42 with governance enabled. Stepping the deterministic stream to 1,000 records at seed 42 fixes the audit-ledger head hash to 7d5aaab4... and the running checksum to 0x06acd6be; any divergence indicates a modified engine or a non-conforming SHA-256 implementation. The triage confidence intervals are Wilson score intervals over the 2,000-trial run.

13. Cross-Domain Ethics Statement

The AUTHREX governance pipeline is architecturally shared with BLADE-EDGE (defense, directed energy), BLADE-AV (autonomous vehicles), BLADE-MARITIME, BLADE-INFRA (critical infrastructure), BLADE-SPACE (orbital autonomy), and BLADE-CUAS (counter-UAS). The financial-sector variant contains no transaction-execution authority of its own: it arbitrates whether an automated decision may proceed and emits a signed, tier-appropriate recommendation to credentialed reviewers and an audit anchor. It is a reference architecture and has not been deployed in any financial institution; the author is an independent researcher and has not been engaged by any institution to provide AI-governance services. All scenarios and data are synthetic.

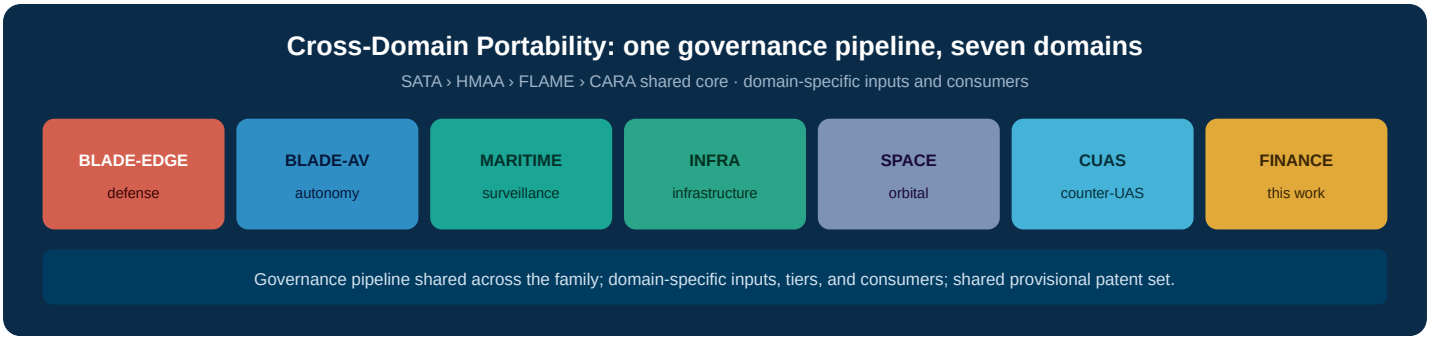

Figure 7: Cross-domain portability of the SATA / HMAA / FLAME / CARA core across the BLADE family. The same governance mathematics is applied across seven application domains within one architectural family.

14. Version History

Version	Date	Changes
v1.0	2026-05-22	Initial Zenodo deposit. Eight-stage AUTHREX pipeline (VALIDATE, SATA, ADARA, MAIVA, HMAA, FLAME, ERAM, CARA) with SHA-256 audit append; four-tier HMAA; population-state coordination model; retrospective stigmergic swarm review with escalation-delta. Reference simulator v2.2 (mulberry32 PRNG, real SHA-256 over canonical-form serialization, schema validation, triage metrics with Wilson intervals, Monte Carlo runner, external-dataset benchmark, golden-trace export, nine self-tests, incremental O(1) ledger continuity check). 36-component reference authority node (~$9,228 BOM). Companion: AUTHREX-FS AI RMF crosswalk (Working Paper No. 1). License CC BY 4.0.

Table 13: Deposit version history.

15. How to Cite

APA: Oktenli, B. (2026). *BLADE-FINANCE Governance Node* (v1.0). Zenodo. DOI 10.5281/zenodo.20374692.

```
@misc{oktenli2026bladefinance, author={Oktenli, Burak}, title={BLADE-FINANCE Governance Node: Authority Governance for Financial-Sector AI Decision Systems Under the Treasury Financial Services AI Risk Management Framework}, year={2026}, publisher={Zenodo}, note={DOI 10.5281/zenodo.20374692}, license={CC-BY-4.0}}
```

16. References

[1] U.S. Department of the Treasury (2026). *Financial Services AI Risk Management Framework (FS AI RMF).* Released February 19, 2026.

[2] U.S. Department of the Treasury (2026). *AI Lexicon.* Companion document to the FS AI RMF, February 19, 2026.

[3] U.S. Department of the Treasury, Office of Public Affairs (2026). *Treasury Launches the Artificial Intelligence (AI) Innovation Series.* First roundtable March 4, 2026.

[4] Executive Order 14179 (2025). *Removing Barriers to American Leadership in Artificial Intelligence.*

[5] National Institute of Standards and Technology (2023). *AI Risk Management Framework, NIST AI 100-1, Version 1.0.*

[6] Sha, L. (2001). Using simplicity to control complexity. *IEEE Software*, 18(4), 20-28.

[7] Hanley, J. A., & Lippman-Hand, A. (1983). If nothing goes wrong, is everything all right? Interpreting zero numerators. *JAMA*, 249(13), 1743-1745.

[8] Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *JASA*, 22(158), 209-212.

[9] National Institute of Standards and Technology. *FIPS PUB 186-5, Digital Signature Standard (DSS).*

[10] National Institute of Standards and Technology. *FIPS PUB 180-4, Secure Hash Standard (SHS).*

[11] Oktenli, B. (2026). *Architectural Crosswalk Between the AUTHREX Authority-Lifecycle Framework and the Treasury FS AI RMF.* BLADE-FINANCE Working Paper No. 1. Zenodo.

[12] Oktenli, B. (2026). *SATA: Sensor Authority and Trust Assessor.* Zenodo. doi:10.5281/zenodo.18936251

[13] Oktenli, B. (2026). *HMAA: Hierarchical Multi-Attribute Authority.* Zenodo. doi:10.5281/zenodo.18861653

[14] Oktenli, B. (2026). *FLAME: Faulted Logic Authority Mitigation Engine.* Zenodo. doi:10.5281/zenodo.19015618

[15] Oktenli, B. (2026). *CARA: Coordinated Authority Recovery Architecture.* Zenodo. doi:10.5281/zenodo.18917790

[16] Oktenli, B. (2026). *ADARA: Adversarial and Deception Attack Recognition Analyzer.* Zenodo. doi:10.5281/zenodo.19043924

[17] Oktenli, B. (2026). *MAIVA: Multi-Agent Intelligent Voting Architecture.* Zenodo. doi:10.5281/zenodo.19015517

[18] Oktenli, B. (2026). *BLADE-CUAS Governance Node.* Zenodo. doi:10.5281/zenodo.20299604