LAKANA SOS v71: A Local-First, Fail-Closed Safety Operating System for Civilian Protection Under Degraded Infrastructure

MarTaize K. Fails / LAKANA Sovereign Systems

April 2026

Title Page

Title: LAKANA SOS v71: A Local-First, Fail-Closed Safety Operating System for Civilian Protection Under Degraded Infrastructure

Subtitle: Architecture-level Monte Carlo evidence from the LAKANA SOS v71 Final Core Connected Empirical Run, with optional movement/accountability/winter post-stack excluded

Author: MarTaize K. Fails, Founder, LAKANA Sovereign Systems

Run label: LAKANA SOS v71 Final Core Connected Empirical Monte Carlo Run - optional movement/accountability/winter post-stack excluded

Primary archive:
sos_v71_full_connected_empirical_guide_temporal_20260428_234814.tar

Archive SHA-256:
b675716e16fb987c6f44a219f855d693a1c98e48871f42491093863694f4cfe7

Document type: Technical research manuscript for preprint preparation, funder diligence, and controlled expert review.

Abstract

Civilian safety systems increasingly depend on networked sensors, cloud inference, and centralized data flows in settings where infrastructure may be degraded, adversarial, congested, or unavailable. LAKANA SOS is proposed as a local-first, fail-closed safety operating system for civilian protection under such conditions. The architecture combines TSARO, a deterministic physics-first trust and risk-envelope layer; NICOLE Protocol, an evidence-integrity and auditability layer; and CivOS, a local survival substrate intended to support bounded execution, fail-closed behavior, and resilient cyber-physical operation. This manuscript reports the LAKANA SOS v71 final core connected empirical Monte Carlo run, labeled final for core evidence with the optional movement/accountability/winter post-stack excluded.

The run used smoke_test=false, 250,000 full trials, 500,000 adversarial trials, ten full batches, ten environmental empirical batches, a batch size of 25,000, 64 workers, an enabled empirical

stack, and environmental fail-closed behavior. The empirical connection review found an environmental trace loaded and sliced into per-batch arrays, empirical transport calibration connected, all six GUIDE proxy lanes loaded and validated, and TSARO temporal/timeline proxy inputs connected and validated. Evidence tiers are separated throughout: environmental trace behavior is treated as empirical measured/raw-derived, transport as empirical calibrated, GUIDE and TSARO temporal/timeline as derived proxy, survivability as simulation-internal, TSARO/NICOLE outputs as native diagnostics, and adversary models as heuristic/synthetic.

In the core survivability comparison, LAKANA+CivOS achieved a simulated safe-state survivability estimate of 0.975304, compared with 0.973176 for LAKANA without CivOS and 0.963996 for the industry-style centralized/cloud comparator. The LAKANA+CivOS minus industry delta was 0.011308, with a bootstrap BCa confidence interval of approximately 0.010908 to 0.011732. Transport success was 0.939064 for LAKANA versus 0.674332 for the industry comparator. TSARO escalation was threat-conditioned but imperfect, and NICOLE integrity behavior remained strong under the modeled assumptions. These findings support architecture-level simulation validation, not deployment-ready field validation. The optional movement/accountability/winter post-stack failed with a NumPy bitwise_and type error and is excluded from final claims.

Keywords

local-first safety; fail-closed systems; Monte Carlo simulation; civilian protection; cyber-physical safety; emergency infrastructure; deterministic bounded envelope; evidence sovereignty; TSARO; NICOLE Protocol; CivOS; LAKANA SOS

Plain-Language Summary

LAKANA SOS is a proposed safety operating system that is designed to keep basic protection logic local, bounded, and auditable when networks are unreliable or adversarial conditions are present. Instead of depending entirely on cloud inference, the system is designed to verify local physical evidence, restrict unsafe state expansion, preserve audit trails, and fail closed when confidence is not adequate.

This paper reports one large simulation run. The run does not prove that LAKANA SOS is ready for deployment, and it does not prove that real-world users will be protected in every emergency. It does show that, under the modeled conditions and with the connected empirical/proxy inputs available in the archive, LAKANA+CivOS preserved simulated safe states more often than an industry-style centralized/cloud comparator. The result is statistically supported in this run, but it remains a simulation result.

The main result is that LAKANA+CivOS reached 0.975304 simulated survivability, while the industry-style centralized/cloud comparator reached 0.963996. That is a difference of about 1.13 percentage points. CivOS also provided a smaller incremental improvement over LAKANA without CivOS. LAKANA's transport model showed stronger survival of communication paths than the comparator, and the TSARO and NICOLE diagnostics behaved coherently within the model.

The paper is careful about what the result means. The GUIDE and TSARO temporal/timeline inputs were connected, but they are proxy-derived rather than field-validated hardware replays. The transport model is empirical-calibrated, not a full live deployment benchmark. The adversary model is synthetic and should not be interpreted as validated attacker economics. One optional downstream post-stack failed and is explicitly excluded.

The safest interpretation is that LAKANA SOS v71 provides bounded, artifact-backed simulation evidence for a local-first, fail-closed architecture. The next step is not stronger language; it is field validation, hardware benchmarking, external audit, and additional sensitivity work.

Contribution Statement

This manuscript makes seven bounded contributions.

First, it presents LAKANA SOS as a local-first, fail-closed civilian safety operating-system architecture intended for degraded, adversarial, and infrastructure-fragile conditions. The architecture is not framed as an autonomous rescue system or deployment-certified emergency system. It is framed as a bounded safety substrate requiring further validation.

Second, it formalizes a deterministic bounded-envelope posture for civilian safety: when evidence is incomplete, compromised, stale, or contradictory, the system narrows permissible behavior rather than expanding into unsupported claims or unsafe action.

Third, it describes the integration of TSARO, NICOLE Protocol, and CivOS within the SOS architecture. TSARO supplies skeptical physical-state validation and escalation logic. NICOLE supplies evidence integrity, separation, document-control, and auditability functions. CivOS supplies a local survival substrate for fail-closed operation and resilient execution paths.

Fourth, it reports a final core connected empirical Monte Carlo run with 250,000 full trials and 500,000 adversarial trials. The run includes empirical environmental trace conditioning, empirical transport calibration, six complete GUIDE proxy lanes, and TSARO temporal/timeline proxy inputs.

Fifth, it reports a statistically supported survivability advantage for LAKANA+CivOS over an industry-style centralized/cloud comparator under modeled assumptions: 0.975304 versus 0.963996, with a delta of 0.011308 and a bootstrap BCa confidence interval of approximately 0.010908 to 0.011732.

Sixth, it separates core evidence from diagnostic, proxy, synthetic, and excluded evidence. This separation is central to the manuscript. The paper does not upgrade proxy-derived GUIDE lanes into field-validated hardware data, and it does not upgrade simulation-internal survivability into deployment validation.

Seventh, it provides a reproducibility and auditability framework for future review while protecting trade-secret implementation details. The intended path is staged: public-safe technical disclosure, restricted reproducibility artifacts for trusted reviewers, and private internal chain-of-custody for sensitive implementation details.

Main Manuscript

1. Introduction

Civilian safety increasingly depends on systems that fuse sensors, communication links, remote analytics, and automated decision support. These systems are often deployed in environments where networks are congested, infrastructure is brittle, and the consequences of unsafe inference are high. Stadiums, campuses, rural routes, industrial sites, disaster zones, farms, and other public-facing settings can all exhibit the same structural tension: safety decisions need to be made close to the physical event, while many modern data pipelines depend on remote cloud infrastructure and centralized data retention.

LAKANA SOS addresses this tension through a local-first safety operating-system architecture. Its central thesis is that safety-critical civilian protection should not depend exclusively on probabilistic cloud availability, behavioral extraction, or externally controlled data custody. Instead, the system should verify local physical evidence, constrain safety envelopes deterministically, preserve audit trails, and fail closed when evidence does not support safe expansion.

This paper reports the LAKANA SOS v71 final core connected empirical Monte Carlo run. The run is treated as final for core paper, white-paper, and funder-facing evidence, with one explicit carveout: the optional movement/accountability/winter post-stack failed downstream and is excluded from all final claims. The manuscript therefore evaluates the core SOS architecture, not the excluded optional post-stack.

The manuscript is deliberately conservative. It does not claim field validation, deployment certification, medical performance, law-enforcement suitability, or guaranteed real-world survival. It reports simulation evidence under modeled assumptions, names the artifacts that support each claim, and separates evidence tiers. This posture is important because safety infrastructure loses credibility when simulation results are described as real-world proof. The value of this run is not that it ends the validation process; it is that it provides a bounded, inspectable, statistically supported basis for the next stage of LAKANA SOS development.

2. Civil Safety Problem Statement

Many high-risk civilian environments operate under uncertainty. Networks may be intermittently unavailable. Sensor feeds may disagree. RF environments may be hostile or congested. Users may be vulnerable to coercion, surveillance, or evidence manipulation. A cloud-only system can still be useful in such environments, but if the system cannot preserve local evidence, operate during degraded transport, or enforce bounded behavior when confidence drops, it can become fragile in precisely the scenarios where safety support is most needed.

The civil safety problem addressed here is therefore not simply prediction accuracy. It is safe-state persistence under degraded infrastructure. The relevant question is not whether a remote model can produce a plausible output under normal conditions. The relevant question is whether an architecture can maintain bounded, auditable, local-first safety behavior when

network transport degrades, when evidence quality changes, when adversarial pressure exists, and when the system must avoid coercive or surveillance-like expansion.

LAKANA SOS is designed around this problem. Its architecture emphasizes fail-closed operation, deterministic bounded envelopes, local evidence custody, and explicit separation between evidence, inference, and release pathways. This design does not remove uncertainty. Rather, it forces uncertainty to be visible and prevents uncertain evidence from being silently converted into broad claims or uncontrolled action.

3. Related Work and Comparator Framing

LAKANA SOS is positioned at the intersection of trustworthy systems engineering, operational-technology security, emergency-communications resilience, cyber-physical safety, edge inference, and controlled simulation reporting. NIST SP 800-160 frames trustworthy secure systems as systems that must be engineered across the full life cycle rather than protected only through after-the-fact perimeter controls [1]. NIST SP 800-82 Rev. 3 extends this concern into operational technology, where programmable systems interact with the physical environment and must preserve safety, reliability, and performance under cyber and infrastructure constraints [2]. NIST AI RMF 1.0 supplies an additional risk-management reference for AI-enabled systems by emphasizing trustworthy, rights-preserving, use-case-aware risk governance rather than unbounded model performance claims [3].

Emergency-management doctrine supplies a second reference frame. FEMA's Community Lifelines construct treats communications, energy, transportation, health and medical systems, safety and security, water systems, hazardous materials, and food/shelter functions as interdependent services whose disruption can cascade across a community [4]. FEMA's BRIC program is likewise organized around pre-disaster mitigation, community capability, partnerships, infrastructure resilience, and innovation [5]. CISA's emergency-communications mission and SAFECOM guidance emphasize interoperable communications that can support emergency response across jurisdictions and conditions, including degraded or unavailable conventional infrastructure [6,7]. Those sources motivate an evaluation posture focused on degraded communications, bounded local operation, and auditable state preservation rather than cloud-only average-case performance.

The safety and control literature adds a third reference frame. Systems-theoretic safety work emphasizes that software-intensive, sociotechnical systems can fail through unsafe control structures and interactions, not only through isolated component failure [8]. Simplex-style architectures and runtime assurance concepts show a long-standing engineering pattern in which a higher-performance controller or function is supervised by a safety-preserving backup or envelope that can retain control when uncertainty or unsafe behavior appears [9,10]. High-assurance cyber-physical systems work, including DARPA HACMS, demonstrates the relevance of formal methods and proof-oriented engineering for safety- and security-critical embedded systems [11]. LAKANA SOS does not claim the same level of formal proof or hardware validation in this manuscript. The relevance is architectural: a local-first, fail-closed safety substrate should be evaluated through explicit boundaries, fallback behavior, and evidence integrity, not only through endpoint accuracy.

The statistical and simulation methodology follows established tools rather than claiming novelty for the tools themselves. Bootstrap confidence intervals are used for cross-architecture contrasts [12]. Confidence-sequence and empirical-Bernstein ideas inform the anytime-precision diagnostic [13,14]. Variance-based sensitivity analysis and Sobol-style estimators support bounded parameter-family interpretation [15,16]. Nearest positive semidefinite projection supports numerical hygiene for correlation structures [17]. Copula models provide a standard way to express correlated multi-channel failure assumptions [18,19]. The manuscript's contribution is the disciplined composition of these methods into an evidence-tiered, claim-bounded civilian safety architecture evaluation.

The comparator used in the v71 Monte Carlo is an industry-style centralized/cloud architecture, not a claim about every existing system or any named competitor. It is a reference architecture used to test whether LAKANA's local-first, fail-closed posture produces a measurable difference under modeled assumptions. The manuscript may state that LAKANA+CivOS outperformed this modeled centralized/cloud comparator in the v71 simulation. It must not state that LAKANA outperforms all industry systems, all emergency platforms, all cloud architectures, or every possible centralized design. Broader comparator generalization would require a broader suite of comparator families and external review of comparator assumptions.

The value of the present comparator is that it provides a public-safe, auditable baseline for architecture-level analysis. It allows the paper to examine whether fail-closed local operation, resilient transport, TSARO escalation logic, NICOLE integrity behavior, and CivOS local survival behavior produce a detectable survivability difference. That is a bounded and defensible research claim.

4. LAKANA SOS System Overview

LAKANA SOS is a Safety Operating System for civilian protection. It is designed to operate as a local-first cyber-physical safety layer that constrains unsafe state expansion and preserves evidence integrity under degraded, adversarial, or infrastructure-fragile conditions. The system is not positioned as a replacement for existing emergency services, weather systems, medical systems, law-enforcement systems, or human judgment. It is positioned as a protective safety substrate that can run alongside existing tools while enforcing bounded evidence-aware behavior.

The core SOS stack contains three central architectural elements: TSARO, NICOLE Protocol, and CivOS.

TSARO is the deterministic physics-first trust/risk envelope layer. In SOS, TSARO evaluates physical-state evidence, applies skeptical validation, and determines whether escalation or protective posture should occur. It is not described as an unconstrained autonomous agent. It is a bounded logic layer that responds to evidence within explicit thresholds and integrity conditions.

NICOLE Protocol is the evidence integrity and auditability layer. It supports separation, document-control logic, key-erasure behavior, evidence integrity, release-path discipline, and controlled auditability. In this paper, NICOLE is discussed as a SOS integrity mechanism. The

paper does not independently validate the full LAKANA privacy/sovereignty doctrine; that claim lane is omitted from the SOS claim matrix.

CivOS is the civic/cyber-physical kernel substrate. In the v71 run, CivOS contributes local survival and fail-closed behavior. The observed incremental survivability contribution is measurable but smaller than the total LAKANA-versus-comparator effect.

GUIDE and TSARO temporal/timeline inputs also appear in the run. They are connected and validated as proxy-derived inputs, not as field-validated hardware replay.

5. Design Principles

The LAKANA SOS design principles are conservative. The system prioritizes local evidence, bounded action, auditability, and fail-closed behavior over expansive prediction. Five principles govern the architecture.

First, physics-first evidence takes priority over remote interpretation. Sensor-derived physical evidence is not assumed to be perfect, but it is treated as the primary input to local safety state estimation. Stale, contradictory, or low-integrity evidence should narrow system behavior.

Second, the system must be local-first. Local-first does not mean that external systems are never used. It means that the basic safety posture must not collapse merely because external transport is degraded. Cloud systems, partners, and external services can augment SOS, but they should not be required for the minimal protective posture.

Third, SOS must fail closed. When evidence is insufficient, the system should not invent confidence or expand into broad assertions. It should reduce permissions, remain silent when necessary, preserve evidence, or enter a safer local mode.

Fourth, evidence integrity must be treated as a first-class output. The NICOLE layer is designed to support auditability, evidence integrity, separation, and controlled release. This matters because safety claims without evidence custody can become unverifiable.

Fifth, the architecture must remain non-coercive. SOS is not designed as surveillance infrastructure or behavioral manipulation infrastructure. The manuscript therefore avoids claims that require raw behavioral extraction, centralized biometric control, or law-enforcement-style deployment assumptions.

6. Deterministic Bounded-Envelope Safety Model

The bounded-envelope model is the core conceptual mechanism of SOS. Instead of allowing uncertain evidence to expand system authority, SOS treats safety behavior as a constrained envelope. If evidence integrity is high and local physical conditions support escalation, the system may move into a stronger protective posture. If evidence is compromised, missing, or contradictory, the permitted envelope shrinks.

This model should not be described as eliminating uncertainty. It is better understood as a governance structure for uncertainty. In a cloud-tethered probabilistic system, uncertainty can be hidden behind a confident output. In a fail-closed bounded-envelope system, uncertainty becomes a constraint on action.

In the v71 run, the bounded-envelope posture is tested indirectly through survivability, transport resilience, TSARO escalation, NICOLE integrity, and failure-mode diagnostics. These are simulation-internal and native diagnostic evidence tiers. They support architecture-level claims but do not constitute field certification.

A reviewer may reasonably ask whether deterministic envelopes are too rigid. The appropriate response is not to claim universality. The appropriate response is that deterministic bounded envelopes are suited for cases where unsupported expansion is more dangerous than conservative silence or safe-state persistence. The architecture should be evaluated against broader comparator families and field pilots before deployment claims are made.

7. TSARO Safety Logic

TSARO is the safety logic layer responsible for skeptical physical-state validation and bounded escalation. In the v71 run, TSARO diagnostics show that escalation is threat-conditioned but imperfect. The core activation diagnostics report TSARO escalation overall at approximately 0.252532 and TSARO escalation given threat at approximately 0.587850. The native TSARO summary reports P(escalate | threat) at approximately 0.801202 and false escalation given no threat at approximately 0.080300.

These values must be interpreted carefully. They indicate that TSARO behavior is not arbitrary and is meaningfully conditioned by threat state in the model. They do not indicate perfect classification, and they do not support a claim of zero false positives. The difference between the core diagnostic escalation rate and the native P(escalate | threat) should be described as arising from different diagnostic definitions and layers, not as a contradiction.

The strongest allowed claim is that TSARO produced meaningful, threat-conditioned escalation behavior under modeled assumptions. The strongest forbidden claim is that TSARO is field-proven, guarantees correct escalation, or eliminates false escalation.

8. NICOLE Evidence and Auditability Model

NICOLE Protocol provides SOS with evidence integrity, separation, and auditability behavior. In the v71 run, NICOLE diagnostics are strong under modeled assumptions. The NICOLE integrity summary reports evidence integrity at approximately 0.999884, separation violation rate at 0.0, replay compromise at approximately 0.000040, and stream compromise at approximately 0.000076. The native NICOLE summary reports integrity OK at approximately 0.999884 and DCL written given release or override path at 1.0.

These results are meaningful, but the paper must not overstate them. They support modeled integrity behavior, not independent privacy audit certification. NICOLE's integrity metrics can be discussed as part of SOS auditability and evidence custody. The manuscript should not create a separate privacy/sovereignty validation claim lane. That lane is intentionally omitted in this SOS paper.

The reviewer risk is that integrity language can sound like certification. The manuscript should therefore state that NICOLE performed strongly under the v71 model and that independent security/privacy audit remains future validation work.

9. CivOS Local Survival Substrate

CivOS is the local survival substrate within the SOS architecture. Its role is to support bounded execution, local survival, and fail-closed operation under degraded conditions. In the v71 survivability results, LAKANA+CivOS achieved 0.975304, while LAKANA without CivOS achieved 0.973176. The implied incremental survivability delta is approximately 0.002128.

This is a measurable but modest effect. It should be described as incremental rather than dominant. The larger architecture-level difference is between LAKANA+CivOS and the industry-style centralized/cloud comparator. CivOS contributes value, but the manuscript should not imply that all results are driven by CivOS alone.

The CivOS energy summary reports battery_after_core_mean at approximately 0.548200 and Lazarus success rate at approximately 0.999988. These values are architecture diagnostics and should be used carefully. They support discussion of local survival behavior within the model, not deployment claims about battery performance in real devices.

10. Transport and Fail-Closed Communication Model

Transport resilience is a central result in the v71 run. The transport survival artifact reports LAKANA transport success at 0.939064. The industry comparator transport success reported in survivability artifacts is approximately 0.674332. Channel success values include BLE at 0.777308, UWB at 0.827036, WiFiNan at 0.681064, LoRa at 0.581312, and Satellite at 0.390472. The transport model uses a copula with rho 0.55 and degrees of freedom 5.0, with degradation beta 0.85.

This evidence is empirical-calibrated. It supports the claim that empirical transport calibration was connected and that LAKANA's modeled multi-transport design preserved communication success more often than the centralized/cloud comparator. It does not prove field transport performance or guarantee communication under all real-world RF conditions.

The fail-closed aspect is important. Transport failure should not cause the system to hallucinate certainty. The architecture should preserve local evidence and safe-state behavior even when external delivery is degraded. The v71 transport results are therefore relevant to the core SOS thesis.

11. GUIDE Proxy Connection Layer

GUIDE inputs were connected and complete. The results/tsaro_guide_inventory.json artifact reports guide_inputs_status=complete, loaded_count=6, provided_failed_count=0, and guide_partial=false. The six loaded lanes are guide_geo_history, guide_mesh_sector, guide_responder_sector, guide_last_safe_sector, guide_rf_quiet_map, and guide_threat_sector.

These facts support the claim that GUIDE proxy lanes were loaded, parsed, and validated in the Monte Carlo. They do not support the claim that field-validated GUIDE hardware data was used. The file paths themselves identify the inputs as derived guide temporal proxy files. The manuscript therefore uses the phrase "GUIDE proxy-derived lanes" throughout.

The reviewer risk is evidence-tier inflation. The paper should preempt that risk by explicitly stating that GUIDE was connected as proxy-derived simulation input and that field GUIDE hardware validation remains future work.

12. TSARO Temporal / Timeline Inputs

TSARO temporal and timeline inputs were also connected. The results/tsaro_temporal_inventory.json artifact reports tsaro_temporal_inputs_status=complete and tsaro_temporal_mode=validated. It contains 12,000 delay frame entries, with temporal behavior used as a proxy-derived validated input.

The strongest allowed claim is that TSARO temporal/timeline proxy inputs were connected and consumed in the core Monte Carlo. The strongest forbidden claim is that the run used direct TSARO hardware temporal history or a fully temporal field replay. This distinction matters because temporal proxy inputs can improve architecture realism but do not replace hardware validation.

13. Threat and Failure Model

The v71 run includes threat, transport, integrity, RF aggression, adversary, and failure-mode components. Threat-related diagnostics include TSARO escalation rates, threat-stratified precision outputs, adversary Markov summaries, attacker cost models, and worst-case perturbation summaries.

The adversarial components must be handled conservatively. The results/adversary_markov_summary.json artifact explicitly marks the adversary Markov model as synthetic and uncertain. The results/adversary_worstcase.json artifact states that the heuristic worst-case search is not a guaranteed worst-case bound. Therefore, the paper may describe these as stress-test diagnostics but must not call them empirically validated attacker economics or formal adversarial guarantees.

Failure modes are more directly useful for model diagnosis. The tables/failure_modes.csv artifact reports 243,826 trials with failure code 0 and 6,174 trials with failure code 1. The results/failure_code1_diagnostics.json artifact reports failure code 1 rate at 0.024696 and identifies a dominant FDE/suppression/context family. This supports the claim that modeled failures are concentrated and diagnostically interpretable.

14. Empirical Data and Simulation Inputs

The run uses multiple evidence tiers. The environmental trace is empirical measured/raw-derived. The results/env_trace_used.json artifact reports that env_trace.csv was loaded once and sliced into per-batch .npz files. The schema maps rf_score_col to rf_score_proxy and wx_silence_col to wx_silence_proxy. The ten environmental empirical batch files each contain rf_score and wx_silence arrays of 25,000 records. Across the combined 250,000 environmental records, RF score mean is approximately 0.033279, RF p95 is approximately 0.084300, RF p99 is approximately 0.146200, and wx_silence mean is approximately 0.245612.

Transport is empirical calibrated through empirical_transport_calibration.csv and empirical_transport_schema.json, identified in the parameter manifest. GUIDE and TSARO temporal/timeline are derived proxies. Survivability is simulation-internal. TSARO and NICOLE summaries are native diagnostics. Adversary Markov and attacker cost outputs are heuristic/synthetic.

The evidence-tier separation is not cosmetic. It is the difference between a credible manuscript and an overclaiming manuscript. The paper's strongest results come from simulation-internal survivability combined with empirical/proxy connection evidence and statistical uncertainty. Its weakest or most caveated evidence comes from synthetic adversary diagnostics and the excluded optional post-stack.

15. Monte Carlo Experimental Design

The v71 run was configured with smoke_test=false, 250,000 full trials, 500,000 adversarial trials, ten full batches, ten environmental empirical batches, batch size 25,000, and 64 workers. The empirical stack was enabled. Environmental fail-closed behavior was enabled. The run completed in approximately 600.49 seconds according to results/runtime_manifest.json.

The batch structure is coherent: ten tmp/batch_*.npz files and ten tmp/env_emp_batch_*.npz files are present. The run archive contains required manifests, result files, tables, figures, batch caches, and amended gate reports. The amended gate report returns PASS_CANONICAL_CANDIDATE, noting that the first gate report falsely failed due to retained parameter path keys while the generated inventory artifacts verified connection.

The parameter manifest retains full_connected_run=false. This flag is treated as a documentation caveat rather than a connection failure because the explicit generated inventory artifacts confirm the environmental trace, empirical transport calibration, GUIDE lanes, and TSARO temporal/timeline connection. Future runs should remove ambiguity by retaining a true full-connected flag when the inventories pass.

16. Statistical Methods

The primary statistical comparison is architecture-level survivability. The run estimates survivability for LAKANA+CivOS, LAKANA without CivOS, and an industry-style centralized/cloud comparator. Confidence intervals are reported through multiple methods, including Wilson and bootstrap percentile/BCa intervals. The primary delta of interest is LAKANA+CivOS minus industry.

Bootstrap BCa intervals are used for the main difference estimate [12]. The bootstrap design uses 10,000 resamples and reports a primary BCa interval for LAKANA+CivOS minus industry of approximately 0.010908 to 0.011732. This interval excludes zero and supports a statistically meaningful positive difference under the model.

Statistical significance must be separated from operational significance. A 1.13 percentage point survivability increase can matter in safety contexts, but its operational value depends on deployment conditions, event rates, cost, usability, and field validation. The manuscript

therefore treats the effect as statistically supported and plausibly operationally relevant, not as deployment proof.

The anytime precision procedure did not stop under the strict eps=0.001 rule. At n=5,000,000, the final halfwidth for architecture estimates was approximately 0.001202 and for deltas approximately 0.002403 under the fail-closed criterion requiring both Hoeffding and empirical-Bernstein confidence sequences within tolerance [13,14]. This limits strict precision-certification language but does not invalidate the main 250,000-trial core result.

17. Results

17.1 Archive and Run Configuration

The archive checksum matched the expected SHA-256 value. The uploaded file is gzip-compressed despite its .tar filename, and it extracted as the expected run directory. The archive contains 78 tar members and 68 files. The run includes the required manifests, result files, tables, figures, batch caches, and amended gate report.

The full run configuration is coherent: 250,000 full trials, 500,000 adversarial trials, batch size 25,000, ten full batches, ten environmental empirical batches, 64 workers, parallel execution, and smoke_test=false. The empirical stack and environmental fail-closed settings are enabled.

The core run is treated as the final core connected empirical Monte Carlo evidence package, with the optional post-stack explicitly excluded.

17.2 Empirical Connection Findings

The environmental trace was loaded once and sliced into per-batch environmental empirical .npz files. Ten environmental empirical batches exist, each with 25,000 rows. The arrays include rf_score and wx_silence. Combined environmental RF score mean is approximately 0.033279, p95 approximately 0.084300, and p99 approximately 0.146200. The combined wx_silence mean is approximately 0.245612.

Empirical transport calibration is connected through the parameter manifest and reflected in transport channel success estimates. GUIDE shows complete six-lane connection with no provided failures. TSARO temporal/timeline inputs show complete validated status.

17.3 Survivability Results

The main survivability estimates are:

Architecture	Survivability estimate	Evidence tier
LAKANA+CivOS	0.975304	Simulation-internal
LAKANA without CivOS	0.973176	Simulation-internal
Industry central cloud comparator	0.963996	Simulation-internal

The LAKANA+CivOS minus industry delta is 0.011308. The bootstrap BCa interval is approximately 0.010908 to 0.011732. This interval is entirely above zero, supporting a statistically meaningful advantage under modeled assumptions.

in the simulation, LAKANA+CivOS preserved safe-state survivability about 1.13 percentage points more often than the industry-style centralized/cloud comparator. This does not prove real-world performance; it supports architecture-level simulation evidence.

17.4 CivOS Incremental Contribution

LAKANA+CivOS survivability was 0.975304, while LAKANA without CivOS was 0.973176. The difference is approximately 0.002128. CivOS therefore provides a measurable but smaller incremental contribution compared with the total LAKANA-versus-industry difference.

The CivOS energy summary reports Lazarus success rate at approximately 0.999988 and battery_after_core_mean at approximately 0.548200. These are internal diagnostics and should not be described as deployed device performance.

17.5 Transport Resilience

LAKANA transport success was 0.939064. Industry comparator transport success was approximately 0.674332. Empirical channel success values include BLE at 0.777308, UWB at 0.827036, WiFiNan at 0.681064, LoRa at 0.581312, and Satellite at 0.390472.

Technical interpretation: the modeled LAKANA transport layer preserved communication success substantially more often than the centralized/cloud comparator under the calibrated transport model.

LAKANA's modeled multi-transport posture held up better than the industry-style comparator when communications degraded.

17.6 TSARO Escalation Behavior

TSARO escalation overall was approximately 0.252532. Core escalation given threat was approximately 0.587850. The native TSARO summary reports P(escalate | threat) at approximately 0.801202 and false escalation given no threat at approximately 0.080300. The TSARO score mean was approximately 2.848566, with threshold 4.0, p95 approximately 5.294939, and p99 approximately 6.239061.

Technical interpretation: TSARO behavior is threat-conditioned, but not perfect. The nonzero false escalation rate is important and should remain visible.

TSARO responded much more often when modeled threats were present, but it still escalated sometimes when no threat was present.

17.7 NICOLE Integrity Behavior

NICOLE evidence integrity rate was approximately 0.999884. Separation violation rate was 0.0. Replay compromise rate was approximately 0.000040, and stream compromise rate approximately 0.000076. Native NICOLE metrics include DCL written given release or override path at 1.0 and key erasure triggered given destructive attempt at 1.0.

Technical interpretation: NICOLE performed strongly within the modeled SOS integrity pathways. This supports a strong integrity/auditability claim under simulation assumptions.

in the simulation, NICOLE preserved evidence integrity very reliably and did not produce modeled separation violations.

17.8 Failure Mode Concentration

The failure table reports 243,826 code-0 cases and 6,174 code-1 cases. Failure code 1 rate is 0.024696. Subclass rates given code 1 show FDE suppression at approximately 57.34%, context interference at approximately 14.56%, confirmation suppressed at approximately 11.01%, far below threshold at approximately 6.38%, near threshold at approximately 5.56%, mid-threshold at approximately 4.16%, bounded suppression at approximately 0.94%, and hard veto at approximately 0.049%.

Technical interpretation: failures are concentrated in a diagnosable family related to suppression and context behavior, rather than spread randomly across unrelated failure codes.

when the system failed in this run, the main weakness was not total architecture collapse; it was concentrated in a specific detection/suppression region that can be targeted for future work.

17.9 Adversarial / Degraded Infrastructure Diagnostics

RF score mean was approximately 0.033279, with p95 approximately 0.084300 and p99 approximately 0.146200. Attacker effort mean was approximately 1.379694, attacker detection probability mean approximately 0.500325, attacker cost mean approximately 2.525366, cost p95 approximately 9.413804, and cost p99 approximately 23.828262.

The adversary Markov summary is explicitly synthetic and uncertain. The worst-case artifact states that the heuristic search is not a guaranteed worst-case bound. These diagnostics are useful for stress-test framing but not for validated adversary economics.

17.10 Replication and Precision Diagnostics

The replication diagnostic used eight replications of 20,000 trials each. Mean survivability estimates were approximately 0.966150 for LAKANA+CivOS, 0.951781 for LAKANA without CivOS, and 0.945125 for the industry comparator. These diagnostic means differ from headline estimates because they come from a separate replication diagnostic design, but their ordering remains consistent.

The anytime precision procedure did not stop. At 5,000,000 final samples, architecture estimate halfwidth under the fail-closed criterion was approximately 0.001202, and delta halfwidth approximately 0.002403, above the strict eps=0.001 requirement for all monitored metrics. This should be stated plainly.

17.11 Optional Post-Stack Exclusion

Source artifact: results/sos_v66_poststack_error.json.

The optional movement/accountability/winter post-stack failed with the error: ufunc 'bitwise_and' not supported for the input types. The artifact note states that the post-stack failed while core artifacts remain valid. This paper excludes the failed optional post-stack from final claims.

The failure is a limitation, not a core invalidation. It affects any claim about the optional movement/accountability/winter post-stack. It does not affect the core survivability, empirical connection, transport, GUIDE, TSARO temporal, TSARO native, NICOLE native, bootstrap, failure-mode, replication, or batch-cache evidence described above.

18. Civil Emergency Management Interpretation

From a civil emergency-management perspective, the manuscript's most important result is not a single number. It is the combination of local-first architecture, transport resilience, evidence integrity, and explicit claim boundaries. Emergency and public-safety systems are evaluated not only by average-case performance but also by their behavior under degraded infrastructure and uncertainty.

The v71 run supports the idea that a local-first, fail-closed safety operating system can preserve simulated safe-state behavior more effectively than a centralized/cloud comparator under modeled degraded conditions. It also shows that the system's failure modes are concentrated and diagnosable, that NICOLE integrity behavior remains strong in the model, and that TSARO escalation is threat-conditioned but imperfect.

For civil emergency management, this is a useful stage of evidence. It is not sufficient for procurement, deployment, or operational certification. It is sufficient to justify further validation: field transport benchmarking, hardware-in-the-loop tests, GUIDE hardware validation, TSARO temporal replay, independent security/privacy audit, and controlled pilots with explicit human oversight.

The paper should therefore position SOS as a candidate safety infrastructure architecture with bounded simulation support, not as a deployed emergency-response replacement.

19. Discussion

The v71 results support three broad conclusions. First, the core SOS architecture produced a statistically supported survivability advantage over the modeled industry-style centralized/cloud comparator. Second, the run shows coherent connection of empirical/proxy inputs, including environmental trace, empirical transport calibration, GUIDE proxy lanes, and TSARO temporal/timeline proxy inputs. Third, diagnostic behavior for TSARO, NICOLE, transport, and failure modes is interpretable rather than opaque.

The magnitude of the headline survivability delta should be interpreted with discipline. A 1.13 percentage point improvement is not trivial in safety contexts, but it is also not a universal guarantee. Its value depends on event rates, deployment setting, false escalation costs, user trust, operational procedures, and field performance. The paper therefore frames it as an architecture-level simulation effect that warrants additional validation.

CivOS contributes an incremental improvement but should not be overstated as the sole driver. TSARO produces meaningful escalation but has nonzero false escalation. NICOLE integrity is strong under the model but requires external audit before certification-style claims. GUIDE and temporal inputs strengthen realism but remain proxy-derived. The adversary model adds stress context but not validated attacker economics.

The optional post-stack failure is important because it demonstrates why claim boundary discipline matters. A weaker manuscript would bury the failure or treat it as irrelevant. This manuscript states the failure directly and excludes the affected module. That strengthens, rather than weakens, the credibility of the core claims.

20. Evidence-Tier and Claim Boundary Review

The paper's admissible claims depend on evidence tiers. Environmental trace data are empirical measured/raw-derived. Transport calibration is empirical calibrated. GUIDE and TSARO temporal/timeline are derived proxies. Survivability is simulation-internal. TSARO and NICOLE summaries are native diagnostics. Adversary models are heuristic/synthetic. The optional post-stack is downstream excluded.

This separation prevents common overclaiming errors. For example, GUIDE connection can be claimed, but field GUIDE hardware validation cannot. NICOLE integrity can be claimed under modeled assumptions, but independent privacy certification cannot. Transport resilience can be claimed under empirical-calibrated simulation, but not guaranteed RF performance. Survivability advantage can be claimed relative to the modeled comparator, but not universal superiority.

A paper that respects these boundaries will be more defensible to a skeptical reviewer, even if the reviewer requests additional validation.

21. Limitations

This study has substantial limitations.

First, it is simulation validation, not deployment validation. The run supports architecture-level claims under modeled assumptions. It does not establish field performance, procurement readiness, emergency-service integration, or operational certification.

Second, the comparator is a modeled industry-style centralized/cloud architecture. It does not represent every possible industry system, every cloud-edge hybrid architecture, or any named commercial product. Broader comparator families are needed before broader market claims can be made.

Third, GUIDE lanes are proxy-derived. The run shows that six GUIDE proxy lanes were connected and validated in the Monte Carlo. It does not show field-validated GUIDE hardware performance.

Fourth, TSARO temporal/timeline inputs are proxy-derived. The run shows complete validated temporal/timeline inputs. It does not show direct hardware replay or deployment logs.

Fifth, empirical transport calibration has limits. It improves realism, but it is not a substitute for field RF benchmarking under diverse conditions and hardware configurations.

Sixth, the environmental trace uses RF/weather proxy behavior. The run consumed environmental inputs, but it is not a full environmental field validation across all relevant civilian settings.

Seventh, the anytime precision procedure did not meet the strict eps=0.001 stopping rule. The main result remains statistically supported, but strict anytime certification should not be claimed.

Eighth, rare-event claims require caution. A 250,000-trial full run and 500,000-trial adversarial run are substantial, but they do not certify extremely rare failure probabilities below the resolution of the design without additional rare-event methods.

Ninth, the adversary models are heuristic/synthetic. They support stress-test interpretation but not validated attacker economics or formal worst-case guarantees.

Tenth, the optional movement/accountability/winter post-stack failed and is excluded. Any claim about that module must wait for a patched and verified rerun or separate validation.

22. Reproducibility and Auditability

The primary run archive is sos_v71_full_connected_empirical_guide_temporal_20260428_234814.tar, with SHA-256 hash b675716e16fb987c6f44a219f855d693a1c98e48871f42491093863694f4cfe7. The archive contains manifests, result artifacts, table outputs, figure outputs, batch caches, empirical environmental batch caches, and gate reports.

A public-safe reproducibility package should disclose: the run label; checksum; high-level code manifest; parameter manifest fields needed to interpret the results; environment manifest; runtime manifest; env trace usage summary; GUIDE inventory summary; TSARO temporal inventory summary; survivability summary; bootstrap outputs; transport survival outputs; TSARO/NICOLE native summaries; failure modes; replication table; anytime precision trace; and the optional post-stack exclusion note.

A restricted reviewer package under NDA may include fuller parameter manifests, code manifests, batch-cache metadata, representative .npz schema descriptions, full table outputs, and reproduction instructions. Sensitive implementation details, private thresholds, anti-coercion triggers, proprietary kernels, and protected architecture details should remain restricted or internal depending on legal/IP strategy.

The public/private split should not be used to hide weaknesses. It should be used to preserve trade-secret discipline while maintaining enough transparency for expert review.

Evidence and Claim Boundary Matrix

Claim	Primary artifacts / metrics	Evidence tier	Public support level
LAKANA+CivOS improves simulated survivability versus the modeled centralized/cloud comparator	survivability_summary_all_arch.json; 0.975304 versus 0.963996; delta 0.011308; BCa CI approximately 0.010908-0.011732	Simulation-internal + bootstrap	Strong within modeled assumptions
CivOS contributes incremental value inside the LAKANA architecture	0.975304 versus 0.973176; delta approximately 0.002128	Simulation-internal	Moderate; measurable but not dominant
Environmental trace conditioning was connected	env_trace_used.json; ten environmental batch files; RF and weather-silence arrays	Empirical measured/raw-derived	Strong connection evidence; not field validation
Transport calibration was connected	transport_survival.json; LAKANA transport success 0.939064 versus industry comparator 0.674332	Empirical calibrated	Strong modeled transport-resilience evidence
GUIDE proxy lanes were connected	tsaro_guide_inventory.json; six of six lanes loaded; zero provided failures	Derived proxy	Strong inventory evidence; not hardware validation
TSARO temporal/timeline inputs were connected	tsaro_temporal_inventory.json; complete; validated; 12,000 delay frames	Derived proxy	Strong connection evidence; not direct hardware replay
TSARO escalation behavior was threat-conditioned	Activation/native summaries; P(escalate | threat) approximately 0.801202; false escalation approximately 0.080300	Native diagnostic	Native diagnostic
NICOLE integrity behavior was strong under modeled assumptions	Evidence integrity approximately 0.999884; separation violation rate 0.0	Native diagnostic	Strong simulation-internal integrity evidence

Claim	Primary artifacts / metrics	Evidence tier	Public support level
Failure modes were diagnostically concentrated	Code-1 failure rate approximately 0.024696; FDE/suppression/context concentration	Diagnostic	Moderate to strong; useful for targeted hardening
Adversarial diagnostics provide stress-test context	Markov and worst-case artifacts explicitly synthetic/uncertain	Heuristic/synthetic	Weak to moderate; not validated attacker economics
Optional movement/accountability/winter post-stack	sos_v66_poststack_error.json records failure	Optional/ downstream excluded	Not part of final claims

Appendices

Appendix A - Artifact Inventory

The archive contains 78 members and 68 files. Required files include manifests/parameter_manifest.json, manifests/environment_manifest.json, manifests/code_manifest.json, results/runtime_manifest.json, results/env_trace_used.json, results/tsaro_guide_inventory.json, results/tsaro_temporal_inventory.json, results/survivability_summary_all_arch.json, results/bootstrap_and_tost.json, results/activation_diagnostics.json, results/transport_survival.json, results/tsaro_native_summary.json, results/nicole_native_summary.json, results/sos_v66_poststack_error.json, tables/failure_modes.csv, tables/replication_survivability.csv, tables/anytime_precision_trace.csv, ten full batch files, and ten environmental empirical batch files.

Appendix B - Evidence-Tier Table

Evidence source	Tier	Use in this manuscript
Environmental trace / RF and weather-silence arrays	Empirical measured/raw-derived	Environmental conditioning only
Transport calibration	Empirical calibrated	Transport resilience simulation only
GUIDE lanes	Derived proxy	Connected proxy context only
TSARO temporal/timeline	Derived proxy	Temporal/timeline conditioning only
Survivability outputs	Simulation-internal	Architecture comparison

Evidence source	Tier	Use in this manuscript
		only
TSARO/NICOLE summaries	Native diagnostic	Mechanism interpretation only
Adversary Markov/cost/worst-case outputs	Heuristic/synthetic	Stress-test context only
Movement/accountability/ winter post-stack	Optional/downstream excluded	Limitation; no final claims

Appendix C - Run Configuration Table

Parameter	Value
smoke_test	false
n_trials_full	250,000
n_trials_adv	500,000
batch_size	25,000
full batch files	10
env empirical batch files	10
workers	64
empirical stack enabled	true
environmental fail-closed	true
runtime seconds	~600.49
amended gate verdict	PASS_CANONICAL_CANDIDATE
final manuscript label	FINAL_CANONICAL_CORE_RUN, optional post-stack excluded

Appendix D - Survivability and Bootstrap Details

Primary survivability values come from results/survivability_summary_all_arch.json. Bootstrap BCa results come from results/bootstrap_and_tost.json. The manuscript uses the BCa interval for the LAKANA+CivOS minus industry delta as the primary uncertainty statement.

Appendix E - Transport Calibration Summary

Transport success was 0.939064 for LAKANA. Industry transport success was approximately 0.674332. Channel success values were BLE 0.777308, UWB 0.827036, WiFiNan 0.681064, LoRa 0.581312, and Satellite 0.390472.

Appendix F - GUIDE Proxy Inventory

All six GUIDE proxy lanes were loaded and validated: guide_geo_history, guide_mesh_sector, guide_responder_sector, guide_last_safe_sector, guide_rf_quiet_map, and guide_threat_sector. The manuscript does not describe these as field-validated hardware inputs.

Appendix G - TSARO Temporal / Timeline Inventory

TSARO temporal inputs were complete and validated, with 12,000 delay-frame entries. The manuscript describes these as proxy-derived validated inputs, not direct hardware replay.

Appendix H - Failure Mode Details

Failure code 1 rate was approximately 0.024696. Among code-1 cases, FDE suppression accounted for approximately 57.34%, context interference approximately 14.56%, and confirmation suppressed approximately 11.01%. This concentration suggests targeted hardening paths.

Appendix I - Optional Post-Stack Exclusion

The optional movement/accountability/winter post-stack failed with a NumPy bitwise_and type error. It is excluded from the final claims. The core Monte Carlo artifacts remain valid for the bounded claims stated in this manuscript.

Appendix J - Public / Restricted / Private Disclosure Split

Public-safe disclosure should include run label, checksum, headline metrics, evidence tiers, survivability results, transport results, GUIDE/temporal status, limitations, and optional post-stack exclusion. Restricted reviewer disclosure may include fuller manifests, code manifest details, batch-cache schemas, and reproduction scripts under NDA. Private chain-of-custody should retain implementation-sensitive kernels, protected thresholds, anti-coercion logic, and proprietary architecture details.

Appendix K - Future Validation Protocols

Future work should include hardware-in-the-loop TSARO tests, field transport benchmarking, GUIDE hardware validation, independent NICOLE security/privacy audit, broader comparator-family analysis, rare-event stress methods, proxy sensitivity analysis, and controlled pilot studies with human oversight.

Conclusion

The LAKANA SOS v71 final core connected empirical Monte Carlo run provides bounded, artifact-backed simulation evidence for a local-first, fail-closed civilian safety operating-system architecture. The core results show statistically supported higher simulated survivability for LAKANA+CivOS compared with a modeled industry-style centralized/cloud comparator, coherent empirical/proxy connection, strong modeled transport resilience, meaningful but imperfect TSARO escalation behavior, strong NICOLE integrity behavior, and diagnostically

interpretable failure modes. The result is strong enough to support a technical manuscript, white paper, and funder-facing evidence package, provided the claim boundary remains strict. The paper should proceed as an expert-facing technical manuscript, not as deployment proof. Future work should move from simulation evidence to hardware-in-the-loop testing, external audit, broader comparator families, field transport benchmarks, and controlled pilots.

Declarations

Scope statement. This manuscript reports simulation evidence and connected empirical/proxy input evidence. It does not claim field validation, deployment certification, emergency-response certification, medical performance, law-enforcement suitability, independent privacy/security certification, or guaranteed real-world survival.

Competing interests. MarTaize K. Fails is the founder of LAKANA Sovereign Systems and the inventor/developer of the evaluated architecture.

Data and code availability. The public manuscript reports aggregated outputs, manifest-level summaries, selected diagnostics, and bounded reproducibility details. Full implementation artifacts, exact thresholds, protected internal schedules, and reconstructive engine details are withheld for intellectual-property, safety, and anti-abuse reasons. Controlled-access review may be pursued separately under appropriate confidentiality terms.

Ethics and human subjects. The reported work is simulation-only and does not report human-subject deployment, clinical intervention, or field emergency-response testing.

References

[1] Ross, R., Winstead, M., and McEvilley, M. *Engineering Trustworthy Secure Systems*. NIST Special Publication 800-160 Volume 1 Revision 1, National Institute of Standards and Technology, 2022. DOI: 10.6028/NIST.SP.800-160v1r1.

[2] Stouffer, K., Pease, M., Tang, C.-Y., Zimmerman, T., Pillitteri, V., Lightman, S., Hahn, A., Saravia, S., Sherule, A., and Thompson, M. *Guide to Operational Technology (OT) Security*. NIST Special Publication 800-82 Revision 3, National Institute of Standards and Technology, 2023. DOI: 10.6028/NIST.SP.800-82r3.

[3] Tabassi, E. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1, National Institute of Standards and Technology, 2023. DOI: 10.6028/NIST.AI.100-1.

[4] Federal Emergency Management Agency. *Community Lifelines Implementation Toolkit 2.1*. FEMA, 2023.

[5] Federal Emergency Management Agency. *Building Resilient Infrastructure and Communities (BRIC)*. FEMA Hazard Mitigation Assistance program materials, current public program page accessed 2026.

[6] Cybersecurity and Infrastructure Security Agency. *Emergency Communications*. CISA public emergency communications program materials.

[7] Cybersecurity and Infrastructure Security Agency. *SAFECOM Guidance on Emergency Communications Grants*. CISA SAFECOM guidance materials.

[8] Leveson, N. G. *Engineering a Safer World: Systems Thinking Applied to Safety*. MIT Press, 2011. DOI: 10.7551/mitpress/8179.001.0001.

[9] Seto, D., Krogh, B. H., Sha, L., and Chutinan, A. Dynamic Control System Upgrade Using the Simplex Architecture. *IEEE Control Systems*, 18(4):72-80, 1998. DOI: 10.1109/37.710880.

[10] Rivera, J. G., Danylyszyn, A. A., Weinstock, C., Sha, L. R., and Gagliardi, M. J. *An Architectural Description of the Simplex Architecture*. CMU/SEI-96-TR-006, Software Engineering Institute, Carnegie Mellon University, 1996.

[11] Defense Advanced Research Projects Agency. *High-Assurance Cyber Military Systems (HACMS)*. DARPA program and case-study materials.

[12] Efron, B., and Tibshirani, R. J. *An Introduction to the Bootstrap*. Chapman and Hall/CRC, 1993.

[13] Hoeffding, W. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13-30, 1963. DOI: 10.1080/01621459.1963.10500830.

[14] Maurer, A., and Pontil, M. Empirical Bernstein bounds and sample variance penalization. arXiv:0907.3740, 2009.

[15] Jansen, M. J. W. Analysis of variance for model output. *Computer Physics Communications*, 117(1):35-43, 1999. DOI: 10.1016/S0010-4655(98)00154-4.

[16] Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. Variance based sensitivity analysis of model output: design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259-270, 2010. DOI: 10.1016/j.cpc.2009.09.018.

[17] Higham, N. J. Computing a nearest symmetric positive semidefinite matrix. *Linear Algebra and its Applications*, 103:103-118, 1988. DOI: 10.1016/0024-3795(88)90223-6.

[18] Demarta, S., and McNeil, A. J. The t copula and related copulas. *International Statistical Review*, 73(1):111-129, 2005. DOI: 10.1111/j.1751-5823.2005.tb00254.x.

[19] Sklar, A. Fonctions de repartition a n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Universite de Paris*, 8:229-231, 1959.

[20] Wilson, E. B. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22(158):209-212, 1927. DOI: 10.1080/01621459.1927.10502953.