

ODA3 INSTITUTE

IMPLEMENTATION GUIDANCE

GAISSF™ v1.0 Troubleshooting Guide

Issue diagnosis, corrective action, recovery and verified resolution

GAISSF-IMP-016 | Public

© ODA3 Pvt Ltd

Document Control

Document ID	GAISST-IMP-016
Title	GAISST™ v1.0 Troubleshooting Guide
Doc Type	Implementation guidance — informative
Purpose	A structured method for diagnosing, containing, correcting and verifying common GAISST™ implementation and operational problems
Primary Audience	Control owners, AI engineers, MLOps, DevSecOps, architects, security operations, incident response, risk, compliance, safety teams, assessors and program leads
Dependencies	GAISST-NOR-001 to GAISST-NOR-008; GAISST-IMP-009 to GAISST-IMP-015
Precedence	GAISST-NOR-001, GAISST-NOR-004, approved incident procedures and legal obligations prevail
Control Baseline	59 controls across nine domains
Status	Final Publication
Publish Date / Review Cycle	Issued 29 June 2026; updated 05 October 2026; review on material change to GAISST-NOR-001 or GAISST-NOR-004
Contact	contact@oda3.org
Classification	Public

GAISST™ v1.0 Troubleshooting Guide

Issue diagnosis, corrective action, recovery and verified resolution

Reliance notice: This guide provides practical issue-resolution methods. It does not replace emergency procedures, legal obligations, specialist safety engineering, incident-response plans or the normative GAISST™ control requirements. Where human safety, unlawful processing, active compromise or irreversible action is possible, containment and specialist escalation take precedence over routine troubleshooting. This guide is not legal, regulatory, safety, engineering, certification or audit advice. Certification is not currently offered.

1. How to Use This Guide

Use the guide after confirming that the issue falls within the approved troubleshooting authority of the team. For material incidents, physical-safety events, suspected compromise, unlawful processing or major public-claim misstatement, activate the relevant incident, legal or safety procedure first.

Each issue pattern identifies symptoms, probable causes, immediate actions, verification and escalation criteria. Resolution is not complete until the corrective action is tested, evidence is preserved, related controls are reviewed and recurrence risk is assessed.

1.1 Reader guide

Reader	Recommended use
Executive or board reader	Review purpose, decision rights, legal and regulatory integration, limitations and Notably Absent sections.
Program or control owner	Use workflows, templates, gates, evidence fields and escalation criteria as implementation aids.
Engineering or operations team	Translate guidance into system-specific procedures, configurations, runbooks and testable acceptance criteria.
Legal, privacy or compliance team	Validate jurisdiction-specific obligations, deadlines, retention, disclosure, privilege and regulator-facing requirements.
Assessor or auditor	Use the document as informative context only; test claims against the authoritative normative sources and applicable assessment criteria.

2. Troubleshooting Principles

Principle	Application
Contain before optimizing	Stop or restrict harmful behavior before attempting performance tuning.
Preserve evidence	Do not destroy logs, model versions, prompts, outputs, state or physical telemetry while correcting the issue.
Reproduce safely	Use isolated test environments and representative data where possible.
Separate symptom from cause	A visible output failure may originate in data, policy, routing, identity, supplier or monitoring.
Verify the approved baseline	Confirm model, prompt, policy, tool, data and infrastructure versions before diagnosis.
Change one variable at a time	Avoid introducing multiple untracked fixes that obscure cause and evidence.
Retest related controls	A failure in one component may invalidate other control evidence.
Document uncertainty	Record when root cause is unknown or only partially supported.

3. Severity and Response Model

Severity	Example condition	Response expectation
S1 — Critical	Active compromise, physical danger, major unlawful disclosure, uncontrolled autonomous action	Immediate containment, incident command and executive/specialist escalation
S2 — High	Material control failure, repeated harmful output, major monitoring blind spot, failed rollback	Urgent restriction, investigation and risk-owner decision
S3 — Moderate	Degraded control, incomplete evidence, recurring noncritical alert, partial mapping defect	Time-bound corrective action and control-owner review
S4 — Low	Documentation defect, minor usability issue, isolated non-material mismatch	Normal improvement workflow

These severity levels are an operational example for troubleshooting within a GAISSE™ programme. Where incidents are classified or handled under UAIF™ or AI-IRF™, those frameworks' severity scales apply.

3.1 S1 response summary

Action	Expectation
Contain	Isolate, disable or enter safe state as appropriate.
Notify	Activate incident, safety, legal and executive procedures.
Preserve	Capture volatile and persistent evidence before further change where feasible.
Escalate	Use designated specialist and regulatory authorities.
Validate	Retest affected controls before restoration.

4. Standard Diagnostic Workflow

1. Confirm the issue and affected scope.
2. Classify severity and determine whether incident or safety procedures apply.
3. Preserve evidence and establish the last known good state.
4. Contain or restrict affected capability.
5. Compare the current state with the approved baseline.
6. Identify candidate causes across people, process, technology, data and suppliers.
7. Reproduce the problem safely where appropriate.
8. Implement the least disruptive corrective action.
9. Validate recovery and retest affected controls.
10. Document root cause, limitations, evidence and recurrence prevention.

5. Root-Cause Analysis Methods

Method	Best use	Caution
Five Whys	Simple recurring process or ownership failures	May oversimplify complex technical causes
Fault tree	Safety, infrastructure and multi-cause failures	Requires disciplined event definitions
Timeline reconstruction	Incidents, unauthorized changes and agent actions	Depends on complete and synchronized logs
Change comparison	Failures appearing after release or configuration update	May miss latent conditions
Control-path analysis	Bypasses, policy failures and missing enforcement	Requires accurate architecture
Data lineage analysis	Poisoning, leakage, drift and retrieval defects	Supplier data may be opaque

Method	Best use	Caution
Hypothesis testing	Complex model behavior and intermittent faults	Avoid changing multiple variables simultaneously

6. Governance and Scope Problems

6.1 Control has no clear owner

Element	Guidance
Symptoms	Tasks are delayed; evidence is missing; multiple teams assume another team is responsible.
Likely causes	RACI not approved; control split across functions; organizational change; owner lacks authority.
Immediate actions	Assign interim accountable owner; map operating roles; escalate authority gaps; update control register.
Verification	Owner accepts accountability; recurring activities and evidence are assigned; escalation path works.
Escalate when	No function accepts ownership or material risk remains unmanaged.

6.2 Scope is inconsistent across documents

Element	Guidance
Symptoms	Inventory, Applicable-Control Baseline, architecture and public claims describe different systems or boundaries.
Likely causes	Uncontrolled scope changes; multiple source lists; unclear supplier boundary; outdated diagrams.
Immediate actions	Freeze current scope; reconcile system inventory, data flows, suppliers and claims; issue controlled scope record.
Verification	All artifacts use the same boundary and identifiers; excluded elements have approved rationale.
Escalate when	Public claims may be misleading.

6.3 Profile applied inconsistently with GAISSE-NOR-001

Element	Guidance
Symptoms	Teams apply different control sets for the same profile — for example, counts taken from superseded drafts — or apply a profile without an approved Applicable-Control Baseline.
Likely causes	Reliance on superseded drafts; informal interpretation; incomplete Applicable-Control Baseline.
Immediate actions	Stop profile claims; apply the profile exactly as defined in GAISSE-NOR-001; maintain an Applicable-Control Baseline covering the 52 D1-D8 Foundational controls and determine D9 applicability wherever Physical AI or cyber-physical actuation is within scope.

Element	Guidance
Verification	The Applicable-Control Baseline matches GAISSE-NOR-001; all 52 D1-D8 controls are accounted for, and each D9 control has a documented applicability decision where Physical AI or cyber-physical actuation is being assessed.
Escalate when	Customer claims or assurance reports relied on the incorrect scope.

6.4 Exception has expired or become permanent

Element	Guidance
Symptoms	Deviation remains open beyond expiry; compensating controls are untested; no remediation progress.
Likely causes	Weak governance; no automated reminders; risk owner unavailable; exception treated as normal.
Immediate actions	Escalate; reassess risk; restrict affected use if necessary; renew only with new evidence and authority.
Verification	New decision, expiry, monitoring and remediation are recorded.
Escalate when	Risk exceeds delegated authority or control failure is material.

7. Inventory, Configuration and Change Problems

7.1 Shadow AI or unknown system discovered

Element	Guidance
Symptoms	Unregistered model/API use, browser extensions, local tools or embedded AI functionality appears.
Likely causes	Procurement bypass; weak discovery; unmanaged SaaS; developer experimentation.
Immediate actions	Identify users, data and actions; restrict access if needed; add to inventory; assess scope and risk.
Verification	System is approved, contained or removed; data exposure and evidence are reviewed.
Escalate when	Sensitive data, consequential action or external disclosure occurred.

7.2 Deployed model version does not match approval

Element	Guidance
Symptoms	Behavior differs from test results; provider changed model; registry and runtime disagree.
Likely causes	Automatic provider update; unpinned version; release bypass; stale documentation.
Immediate actions	Freeze deployment; identify actual version; compare behavior; revert or re-evaluate.
Verification	Runtime, registry, evidence and release record align.
Escalate when	Unknown provider change affects high-risk operation or cannot be reversed.

7.3 Configuration drift detected

Element	Guidance
Symptoms	Policies, prompts, tools, thresholds or infrastructure differ from approved baseline.
Likely causes	Manual edits; emergency changes; weak configuration management; environment mismatch.
Immediate actions	Capture current state; block further change; compare baseline; restore or approve through change control.
Verification	Drift is eliminated or formally accepted; regression tests pass.
Escalate when	Drift bypassed security, safety or evidence controls.

7.4 Rollback package is unavailable or unusable

Element	Guidance
Symptoms	Last known good version cannot be restored; dependencies or data schema changed.
Likely causes	Incomplete release package; retention failure; untested rollback; supplier dependency.
Immediate actions	Restrict new activity; reconstruct known-good state; create emergency recovery plan; test in isolation.
Verification	Rollback or safe degradation works and evidence is preserved.
Escalate when	Service cannot be safely contained or irreversible actions continue.

8. Data, Retrieval and Privacy Problems

8.1 Sensitive data appears in model output

Element	Guidance
Symptoms	PII, confidential data, secrets or restricted content is reproduced.
Likely causes	Training memorization; retrieval authorization failure; prompt leakage; logging exposure.
Immediate actions	Stop affected output path; preserve evidence; identify source; revoke access; notify privacy/incident teams.
Verification	Leakage path is closed; affected data and users are assessed; regression tests pass.
Escalate when	Breach notification, legal reporting or material harm may apply.

8.2 Retrieval returns unauthorized documents

Element	Guidance
Symptoms	Users receive content outside role, tenant or purpose.

Element	Guidance
Likely causes	Authorization applied after retrieval; filter bypass; index contamination; shared cache.
Immediate actions	Disable retrieval or affected index; correct pre-retrieval authorization; clear contaminated caches.
Verification	Negative-access tests pass across tenants and roles.
Escalate when	Cross-tenant or regulated-data exposure occurred.

8.3 Knowledge base is stale or inconsistent

Element	Guidance
Symptoms	Answers cite outdated policy; different users receive conflicting information.
Likely causes	Failed indexing; missing freshness metadata; duplicate sources; delayed synchronization.
Immediate actions	Identify authoritative source; re-index; expire old content; add freshness and precedence rules.
Verification	Outputs use current approved sources and show correct citations.
Escalate when	Outdated information caused consequential action or compliance risk.

8.4 Data poisoning suspected

Element	Guidance
Symptoms	Unexpected behavior follows data-source change; anomalies cluster around new records.
Likely causes	Untrusted source; compromised pipeline; malicious document; weak provenance.
Immediate actions	Quarantine source; stop affected training/retrieval; preserve lineage; compare before/after behavior.
Verification	Source integrity is restored; model/index is rebuilt or validated; tests pass.
Escalate when	Poisoning affects production decisions, safety or multiple systems.

8.5 Data drift alert is persistent

Element	Guidance
Symptoms	Input or embedding distributions remain outside baseline; false alarms recur.
Likely causes	Real population change; monitoring defect; seasonal behavior; pipeline change.
Immediate actions	Validate signal quality; segment by population; compare business context; adjust only with approval.
Verification	Threshold and response remain justified; model performance and risk are reassessed.
Escalate when	Drift affects fairness, safety, legal obligations or critical decisions.

9. Model Behavior and Output Problems

9.1 Harmful or prohibited output increases

Element	Guidance
Symptoms	Safety filters trigger more often; users report abuse; refusal behavior changes.
Likely causes	Model/provider change; policy drift; adversarial campaign; new use context.
Immediate actions	Restrict affected capability; compare model and policy versions; increase review; run targeted evaluation.
Verification	Harm rate returns within approved limits and regression evidence is recorded.
Escalate when	Severe harm, repeated bypass or vulnerable populations are affected.

9.2 Hallucination or unsupported answers remain high

Element	Guidance
Symptoms	Confident incorrect answers; missing citations; conflicting evidence.
Likely causes	Poor grounding; retrieval failure; ambiguous task; excessive temperature; unsupported use case.
Immediate actions	Constrain task; require citations; improve retrieval; lower randomness; add abstention and human review.
Verification	Representative tests meet approved accuracy and abstention criteria.
Escalate when	Errors create legal, medical, financial, safety or material business consequences.

9.3 Refusal rate is too high

Element	Guidance
Symptoms	Legitimate requests are blocked; users route around controls.
Likely causes	Overbroad policy; model change; poorly calibrated classifier; context loss.
Immediate actions	Sample false refusals; refine policy boundaries; retrain or calibrate; maintain blocked harmful cases.
Verification	Legitimate completion improves without increasing unsafe output.
Escalate when	Users create uncontrolled workarounds or critical service is impaired.

9.4 Model extraction or abnormal querying suspected

Element	Guidance
Symptoms	High-volume systematic queries; repeated boundary probing; unusual similarity requests.
Likely causes	Credential abuse; automated scraping; weak rate limits; exposed endpoint.
Immediate actions	Rate-limit, revoke credentials, isolate account, preserve traffic, review output leakage.
Verification	Attack pattern stops; access and monitoring controls are improved.

Element	Guidance
Escalate when	Proprietary model, sensitive data or broad customer exposure is at risk.

9.5 Model performance differs by group

Element	Guidance
Symptoms	Error, refusal or harmful-output rates vary materially across populations.
Likely causes	Unrepresentative data; label bias; deployment context; threshold choice.
Immediate actions	Segment evaluation; confirm data and outcome definitions; involve legal/ethics owners; redesign or restrict use.
Verification	Differences are understood, mitigated and monitored with approved criteria.
Escalate when	Protected groups, rights or high-impact decisions are affected.

10. Prompt Injection and Adversarial Problems

10.1 Direct prompt injection bypasses policy

Element	Guidance
Symptoms	User instructions override system rules; prohibited tools or outputs become available.
Likely causes	Prompt-only control; weak instruction hierarchy; missing policy enforcement.
Immediate actions	Disable high-risk tools; enforce external policy gate; add input classification and adversarial tests.
Verification	Injection scenarios fail safely and policy decisions are logged.
Escalate when	Consequential actions or sensitive data were exposed.

10.2 Indirect injection from documents or web content

Element	Guidance
Symptoms	Retrieved content changes agent behavior or requests tool actions.
Likely causes	Untrusted instructions treated as authority; no content isolation; overpowered agent.
Immediate actions	Quarantine content; separate data from instructions; constrain tools; require approval.
Verification	Malicious document tests cannot alter policy or trigger unauthorized action.
Escalate when	External content caused transactions, data disclosure or persistent compromise.

10.3 Jailbreak test results are inconsistent

Element	Guidance
Symptoms	Same test passes and fails across runs; teams dispute severity.
Likely causes	Non-determinism; model version drift; incomplete test harness; inconsistent settings.

Element	Guidance
Immediate actions	Fix versions and parameters; repeat statistically; record prompts, seeds and environment.
Verification	Results are reproducible enough for decision and limitations are documented.
Escalate when	High-risk bypass is observed even intermittently.

10.4 Multi-modal input bypasses safeguards

Element	Guidance
Symptoms	Text controls work but images, audio or files trigger unsafe behavior.
Likely causes	Modality-specific parser or classifier gap; unscanned attachments.
Immediate actions	Restrict modality; scan and normalize input; test cross-modal attacks; align policies.
Verification	Equivalent policy outcomes apply across supported modalities.
Escalate when	Unsafe action or data extraction occurred.

11. Agentic AI Problems

11.1 Agent enters a loop or fails to terminate

Element	Guidance
Symptoms	Repeated tool calls, rising cost, duplicate actions or stalled workflow.
Likely causes	No step limit; ambiguous goal; tool failure; recursive delegation.
Immediate actions	Activate circuit breaker; revoke credentials; preserve state; impose step/time limits.
Verification	Agent terminates predictably and duplicate actions are prevented.
Escalate when	Financial, operational or physical consequences continue.

11.2 Agent performs unauthorized tool action

Element	Guidance
Symptoms	Action exceeds user role, destination, value or approved objective.
Likely causes	Shared credential; missing action-level policy; confused deputy; injection.
Immediate actions	Stop agent; revoke token; reverse action where possible; inspect approval and policy logs.
Verification	Per-agent identity and action authorization block recurrence.
Escalate when	Transaction is irreversible, regulated or materially harmful.

11.3 Agent memory leaks between users

Element	Guidance
Symptoms	One user sees another user's context or prior task data.

Element	Guidance
Likely causes	Shared memory namespace; tenant filter failure; cache contamination.
Immediate actions	Disable memory; isolate tenants; purge affected state; assess disclosure.
Verification	Cross-user negative tests pass and retention/deletion rules work.
Escalate when	Personal, confidential or regulated information was exposed.

11.4 Sub-agent delegation is untraceable

Element	Guidance
Symptoms	Actions occur without clear parent, sponsor or permission chain.
Likely causes	Implicit delegation; shared identities; missing correlation IDs.
Immediate actions	Suspend delegation; require parent-child identity and inherited permission limits.
Verification	Every action is attributable through the delegation chain.
Escalate when	High-impact actions cannot be reconstructed.

11.5 Human approval is bypassed

Element	Guidance
Symptoms	Agent executes before approval or approval is recorded after the fact.
Likely causes	Race condition; asynchronous queue; UI flaw; emergency override misuse.
Immediate actions	Disable affected transaction; enforce server-side approval token; test timing and replay.
Verification	No action executes without valid pre-action approval.
Escalate when	Consequential or irreversible action occurred.

12. Deployment and Platform Problems

12.1 Pilot succeeds but production fails

Element	Guidance
Symptoms	Errors, harms or latency appear only after scale-up.
Likely causes	Pilot not representative; traffic, users, regions or data differ; hidden dependency.
Immediate actions	Pause expansion; segment failures; compare pilot and production conditions; roll back if needed.
Verification	Representative load and scenario tests pass before re-expansion.
Escalate when	Material harm, outage or control failure persists.

12.2 Canary metrics look healthy but users report problems

Element	Guidance
Symptoms	Aggregate dashboards pass while specific groups or workflows fail.
Likely causes	Metrics too coarse; sampling bias; alert thresholds miss local impact.
Immediate actions	Segment metrics; inspect complaints and traces; add outcome and user-impact measures.
Verification	Dashboard detects the previously hidden failure.
Escalate when	Vulnerable users or critical workflows are affected.

12.3 Environment parity is broken

Element	Guidance
Symptoms	Staging tests pass but production behavior differs.
Likely causes	Different model, policy, data, secrets, network or runtime configuration.
Immediate actions	Compare manifests; align environments; record justified differences; repeat tests.
Verification	Production-relevant configuration is represented in validation.
Escalate when	Security or safety controls exist only in non-production.

12.4 Feature flag enables unapproved capability

Element	Guidance
Symptoms	Hidden feature becomes active without release approval.
Likely causes	Weak flag governance; default-on configuration; stale experiment.
Immediate actions	Disable flag; audit changes and users; restrict permissions; require approval workflow.
Verification	Flag inventory, ownership and logging prevent recurrence.
Escalate when	Unapproved capability processed sensitive data or took actions.

12.5 Supplier outage or silent service change

Element	Guidance
Symptoms	Latency, behavior, refusal or output format changes unexpectedly.
Likely causes	Provider outage; model update; API change; region shift.
Immediate actions	Activate fallback or degradation; compare provider notices; validate output and risk.
Verification	Service returns to approved baseline or migration is completed.
Escalate when	No safe fallback exists for critical operation.

13. Monitoring and Evidence Problems

13.1 Monitoring blind spot discovered

Element	Guidance
Symptoms	Actions cannot be reconstructed; required events are missing.
Likely causes	Logging disabled; unsupported provider; pipeline failure; privacy over-filtering.
Immediate actions	Restrict high-risk operation; restore telemetry; identify affected period; document uncertainty.
Verification	End-to-end event coverage and alerting are tested.
Escalate when	The missing period includes incident, high-risk action or assessment evidence.

13.2 Alert volume overwhelms operators

Element	Guidance
Symptoms	Backlog grows; acknowledgements are delayed; critical alerts are missed.
Likely causes	Poor thresholds; duplicate sources; no risk prioritization; noisy model behavior.
Immediate actions	Prioritize by harm; suppress duplicates; tune with evidence; add ownership and service levels.
Verification	Actionable alert rate and response time improve without masking true positives.
Escalate when	Critical events are unreviewed or staffing is inadequate.

13.3 Evidence is stale or incomplete

Element	Guidance
Symptoms	Assessment cannot confirm current operation; artifacts refer to old versions.
Likely causes	Manual collection; unclear owner; change not linked to evidence refresh.
Immediate actions	Identify affected controls; refresh evidence; automate capture; update expiry rules.
Verification	Evidence matches current scope, version and period.
Escalate when	Customer or public claims relied on invalid evidence.

13.4 Evidence conflicts across sources

Element	Guidance
Symptoms	Dashboard says control passed while logs or test results show failure.
Likely causes	Different periods; metric aggregation; source tampering; definition mismatch.
Immediate actions	Freeze conclusion; reconcile timestamps, scope and definitions; preserve all versions.
Verification	A single supported conclusion and limitation record are approved.
Escalate when	Conflict indicates manipulation or material control failure.

13.5 Failed test result was removed or altered

Element	Guidance
Symptoms	Negative evidence is missing; report history differs from repository.
Likely causes	Pressure to pass; weak access control; manual overwrite.
Immediate actions	Preserve repository; restrict access; investigate integrity; restore audit trail.
Verification	All results and changes are attributable and immutable enough for assurance.
Escalate when	Intentional evidence manipulation or public-claim misstatement is suspected.

14. Incident, Recovery and Continuity Problems

14.1 Incident severity is disputed

Element	Guidance
Symptoms	Teams disagree whether to escalate; response is delayed.
Likely causes	No AI-specific severity criteria; unclear affected persons or consequence.
Immediate actions	Use worst credible impact; convene security, safety, legal and system owners; document decision.
Verification	Severity and authority are agreed and response times are met.
Escalate when	Delay may increase harm or reporting deadlines are uncertain.

14.2 Root cause remains unknown

Element	Guidance
Symptoms	System is restored but no supported explanation exists.
Likely causes	Incomplete logs; supplier opacity; non-determinism; multiple concurrent changes.
Immediate actions	Document unknown cause; increase monitoring; restrict risk; run structured hypotheses.
Verification	Residual uncertainty is accepted by proper authority and recurrence controls exist.
Escalate when	High-severity risk cannot be bounded.

14.3 Recovery restores service but not controls

Element	Guidance
Symptoms	System is available but monitoring, policy or evidence remains degraded.
Likely causes	Recovery focused only on uptime; incomplete runbook.
Immediate actions	Keep service restricted; validate each critical control before full return.
Verification	Control-health and evidence checks pass, not only functional tests.
Escalate when	High-risk operation resumes without assurance.

14.4 Incident evidence cannot be synchronized

Element	Guidance
Symptoms	Prompts, tool calls, logs and physical telemetry have inconsistent timestamps.
Likely causes	Clock drift; missing correlation IDs; separate supplier systems.
Immediate actions	Preserve original records; establish offset; use sequence and transaction identifiers.
Verification	Timeline is sufficiently reliable and limitations are recorded.
Escalate when	Legal, safety or fraud investigation requires precise reconstruction.

15. Physical AI Problems

15.1 Emergency stop does not work as expected

Element	Guidance
Symptoms	Stop command is delayed, ignored or dependent on the AI controller.
Likely causes	Shared failure path; maintenance defect; inaccessible interface.
Immediate actions	Enter alternate safe state; remove system from service; activate safety incident process.
Verification	Independent stop is repaired and tested under realistic conditions.
Escalate when	Any person may be exposed to harm.

15.2 System leaves approved operational envelope

Element	Guidance
Symptoms	Speed, location, force, load or environmental limits are exceeded.
Likely causes	Planner error; sensor fault; stale map; safety supervisor bypass.
Immediate actions	Stop or restrict system; preserve telemetry; inspect supervisor and sensor data.
Verification	Envelope enforcement blocks recurrence and safe recovery is tested.
Escalate when	Actual or near-miss physical harm occurred.

15.3 Sensor disagreement causes unstable behavior

Element	Guidance
Symptoms	Rapid switching, erratic motion or inconsistent perception.
Likely causes	Calibration drift; obstruction; time synchronization; sensor compromise.
Immediate actions	Reduce capability or stop; isolate faulty sensor; use validated fallback or voting.
Verification	Fault detection and degraded mode perform as designed.
Escalate when	Safe operation cannot be demonstrated.

15.4 Safe state cannot be reached

Element	Guidance
Symptoms	System remains active or creates new hazards during failure.
Likely causes	Poor safe-state design; energy or mechanical constraint; software-only assumption.
Immediate actions	Activate emergency procedure; isolate energy; involve specialist safety authority.
Verification	Safe-state transition is redesigned and physically tested.
Escalate when	Immediate danger exists or fail-safe assumptions are invalid.

15.5 Physical incident evidence is incomplete

Element	Guidance
Symptoms	Cannot reconstruct commands, sensor state, overrides or actuator response.
Likely causes	Insufficient telemetry; storage loss; unsynchronized systems.
Immediate actions	Preserve hardware and available logs; photograph state; obtain supplier records.
Verification	Future events record synchronized safety-relevant data.
Escalate when	Investigation, reporting or liability may be affected.

16. Assessment and Public-Claim Problems

16.1 Assessment finding is challenged

Element	Guidance
Symptoms	Control owner disputes evidence or interpretation.
Likely causes	Scope mismatch; incomplete sampling; ambiguous criterion; missing context.
Immediate actions	Separate factual and interpretive issues; review exact requirement, scope and evidence.
Verification	Decision and rationale are documented through approved review or appeal route.
Escalate when	Independence, fairness or public claims are affected.

16.2 Public claim misstates scope or status

Element	Guidance
Symptoms	Marketing implies GAISST TM certification (which is not currently offered), enterprise-wide coverage or a higher profile than evidence supports.
Likely causes	Poor claims review; vague wording; outdated assessment.
Immediate actions	Correct or withdraw the claim; involve legal and communications; preserve evidence.
Verification	All public statements match version, profile, scope and status.
Escalate when	Customers, regulators or investors may have relied on the claim.

16.3 Crosswalk is treated as proof of conformance

Element	Guidance
Symptoms	Teams close GAISSF controls based only on ISO/NIST/OWASP mapping.
Likely causes	Confusion between relationship and implementation; evidence not reassessed.
Immediate actions	Reopen controls; test scope and operating effectiveness; classify mapping limitations.
Verification	Each GAISSF control has direct implementation and evidence.
Escalate when	Assessment readiness or compliance claims rely on unsupported closure.

16.4 Assessment evidence period is insufficient

Element	Guidance
Symptoms	Current records do not cover the required operating period.
Likely causes	Recent implementation; evidence loss; late scope change.
Immediate actions	Defer claim or assessment; extend evidence period; use limited scope only if valid.
Verification	Required period is represented and stable operation is demonstrated.
Escalate when	Continued claims would be misleading.

17. Decision Trees

Question	Yes	No
Could the issue cause immediate human harm?	Stop/enter safe state and activate safety response	Continue severity assessment
Is active compromise or unauthorized action occurring?	Contain, revoke access and activate incident response	Continue diagnosis
Is sensitive or regulated data exposed?	Restrict processing and involve privacy/legal	Continue technical troubleshooting
Can the issue be reproduced safely?	Use isolated environment and controlled variables	Rely on evidence, simulation and bounded hypotheses
Is the last known good state available?	Rollback or compare directly	Construct recovery baseline and restrict operation
Does evidence show the fix worked?	Retest related controls and close formally	Do not close; revise corrective action

18. Corrective Action and Closure

Closure criterion	Required evidence
Containment effective	Harmful or uncontrolled behavior has stopped
Root cause supported	Evidence supports the cause, or uncertainty is explicitly accepted
Corrective action implemented	Change record, configuration, code, process or training evidence

Closure criterion	Required evidence
Regression passed	Original failure and related scenarios no longer reproduce
Control retested	Affected GAISSF controls return to approved state
Scope reviewed	Related systems, suppliers and populations assessed
Evidence preserved	Original and corrected-state records retained
Owner approval	Control and risk owners accept closure
Prevention established	Monitoring, design or workflow reduces recurrence

19. Troubleshooting Metrics

Metric	Purpose	Caution
Mean time to contain	Assess ability to limit harm	Measure by severity and system type
Mean time to diagnose	Assess observability and expertise	Unknown root cause may be honest
Repeat issue rate	Detect weak corrective action	Normalize by system and volume
Rollback success rate	Assess recovery readiness	A rollback may still lose data or transactions
Evidence completeness	Assess reconstructability	Artifact count is not quality
Alert-to-incident conversion	Assess alert usefulness	Low rate may reflect under-escalation
Corrective-action overdue rate	Assess governance discipline	Prioritize by risk, not count alone

20. Common Troubleshooting Anti-Patterns

Anti-pattern	Why it fails	Required correction
Restart and forget	Removes symptom without preserving cause	Preserve evidence and complete root-cause review
Tune thresholds until alerts disappear	Can hide real failures	Validate signal and risk before adjustment
Patch only the visible output	Underlying data or control path remains broken	Review end-to-end control chain
Close when service returns	Controls and evidence may remain degraded	Require control-health verification
Blame the model provider immediately	Local integration may be the cause	Compare supplier and local evidence
Delete failed tests	Destroys assurance integrity	Retain and explain all results
Fix directly in production	Creates uncontrolled change and weak evidence	Use emergency change only with authority and rollback
Assume one incident is isolated	Systemic weakness may persist	Assess related systems and controls

21. Legal and Regulatory Integration

GAISSF provides an operational security, safety, governance and evidence framework. It does not determine which laws apply, establish statutory compliance, create a legal safe harbour, define a legal standard of care, replace competent legal advice, or guarantee that evidence collected for GAISSF purposes will satisfy a regulator, court or other authority. Organizations should maintain a jurisdiction-specific legal and regulatory obligations register and map applicable obligations to their GAISSF scope, controls, evidence, reporting processes and decision authorities.

- Identify applicable jurisdictions, regulators, sector rules, contractual duties and internal legal authorities before relying on this guide.
- Maintain a controlled obligations register, reporting-clock register, retention schedule, legal-hold process and conflicts-of-obligation procedure.
- Treat statutory, regulatory and contractual analysis as a parallel workstream to GAISSF implementation and assurance.
- Record the legal basis, responsible owner, scope, decision, evidence, deadline and review date for each material obligation.
- Do not represent GAISSF conformance, maturity or certification as legal approval, permission to operate, immunity from enforcement or proof that harm cannot occur.

21.1 Preservation-aware pre-change checkpoint

- Determine whether litigation, enforcement, investigation or a preservation duty is pending or reasonably anticipated.
- Check for active legal holds, regulator instructions, consent orders or change restrictions.
- Identify whether reboot, scaling, redeployment, cache clearing, log rotation or instance termination could destroy relevant evidence.
- Acquire forensic or evidentiary copies before destructive remediation where required.
- Record authorization for emergency changes and unavoidable evidence loss.

21.2 Technical containment versus legal preservation

Troubleshooting actions may create preservation risk where relevant evidence is destroyed after a duty to preserve has arisen. Technical containment and legal preservation should therefore be coordinated, but neither automatically overrides immediate life-safety or critical-service actions.

21.3 RCA governance and privilege caution

For severe events, organizations should determine with qualified counsel whether legal direction, separate factual and legal workstreams, distribution restrictions or other measures are appropriate. Counsel involvement does not by itself guarantee attorney-client privilege or work-product protection, and protection may vary or be waived.

21.4 Known-issues and failed-test repository

- Retain the original test script, harness configuration, raw result, environment log and failed output.
- Record verified fixes, applicable versions, side effects, recurrence indicators and evidence.
- Do not convert an unverified workaround into an approved baseline.

22. Limitations

- The issue patterns are not exhaustive and cannot replace specialist investigation.
- Some failures cannot be reproduced because of non-determinism, supplier opacity or missing telemetry.
- A successful fix does not prove that all related risks are eliminated.
- Physical-AI troubleshooting may require qualified safety engineers and controlled test facilities.
- Legal and regulatory reporting obligations may apply before root cause is known.

23. Notably Absent

- No universal severity threshold for every AI system.
- No permission to delay emergency containment while completing diagnosis.
- No guarantee that rollback restores all data, transactions or physical state.
- No assumption that absence of alerts means absence of failure.
- No approval to alter evidence, hide failed tests or suppress material findings.
- No substitution for legal advice, incident command or specialist safety authority.
- No assertion that every loss of data is legally actionable spoliation.
- No guarantee of privilege from counsel involvement.
- No permission to alter systems contrary to regulator, court or consent-order restrictions.

Annex A — Troubleshooting Record Template

Field	Entry
Issue ID	
Date/time	
System/scope	
Reporter	
Severity	
Symptoms	
Containment	
Evidence preserved	
Baseline compared	
Hypotheses	
Root cause	
Corrective action	
Verification	
Related controls	

Field	Entry
Residual risk	
Approver	
Closure date	

Annex B — Root-Cause Worksheet

Category	Questions	Findings
People	Were roles, competence, authority or handoffs inadequate?	
Process	Did workflow, approval, change or escalation fail?	
Technology	Did model, code, platform, policy or integration fail?	
Data	Were provenance, quality, access, drift or retention defective?	
Supplier	Did an inherited control, service or notification fail?	
Environment	Did scale, geography, physical context or threat change?	

Annex C — Verification Checklist

- ☐ Original issue no longer reproduces under representative conditions
- ☐ Containment can be removed safely
- ☐ Affected control evidence is current
- ☐ Related systems and suppliers were assessed
- ☐ Monitoring detects recurrence
- ☐ Rollback or safe state remains available
- ☐ Residual risk is approved
- ☐ Users and stakeholders were informed where required
- ☐ Corrective action is documented and closed

Annex D — Escalation Matrix Template

Condition	Primary authority	Secondary authority	Target response	Required evidence
Physical danger	Safety authority	Executive/incident command		

Condition	Primary authority	Secondary authority	Target response	Required evidence
Active compromise	Security incident lead	System owner/legal		
Data exposure	Privacy/legal	Security/system owner		
Agent unauthorized action	System owner	Security/risk		
Public-claim misstatement	Legal/communications	Executive/communications		
Control evidence failure	Control owner	Internal assurance/risk		

Annex E — Source Traceability

Troubleshooting subject	Controlled source
Normative controls	GAISST-NOR-001 and GAISST-NOR-004
Terminology	GAISST-NOR-005
Release and correction status	GAISST-NOR-007 and GAISST-NOR-008
Adoption governance	GAISST-IMP-009
Implementation workflows	GAISST-IMP-010
Deployment and rollback	GAISST-IMP-011
Migration and evidence reuse	GAISST-IMP-012
Architecture and trust boundaries	GAISST-IMP-013
Operations and monitoring	GAISST-IMP-014
Common interpretation	GAISST-IMP-015

Related guidance	Document
Incident, safety, legal and regulatory procedures	Organisation-approved procedures
Operations	GAISST-IMP-014
Evidence preservation	GAISST-IMP-017

Annex F — Publication Record

Date	Change
29 June 2026	Initial troubleshooting guide for GAISST TM v1.0.

Date	Change
05 October 2026	Reframed Issue 6.3 so the Applicable-Control Baseline follows GAISSTM-NOR-001 explicitly: 52 D1-D8 Foundational controls, with D9 applicability determined where Physical AI or cyber-physical actuation is in scope; reframed certification issues as assessment and public-claim issues (certification is not currently offered); replaced 'SoA' with 'Applicable-Control Baseline'; noted UAIFTM/AI-IRFTM severity scales; moved legal, preservation, privilege, known-issues and S1 guidance into the body. No GAISSTM control was added, removed, renamed or amended.

GAISSTM is a framework of ODA3 Institute, referenced under the GAISSTM Ecosystem Licence (GEL) v1.0. | GAISSTM Ecosystem developed by ODA3 Institute, licensed under GEL v1.0. | © 2026 ODA3 Pvt Ltd