
GAISSF™ v1.0

Deployment Playbook

Step-by-Step Rollout, Production Transition and Operational Handover

Global AI Safety and Security Framework

GAISSF-IMP-011 | Public | Final Publication

Issued 29 June 2026 | Updated 05 October 2026

© 2026 ODA3 Pvt Ltd

Document Control

Attribute	Controlled value
Title	GAISSF™ v1.0 Deployment Playbook
Canonical framework	Global AI Safety and Security Framework
Document ID	GAISSF-IMP-011
Purpose	Provide a repeatable step-by-step deployment and rollout method
Primary audience	Program managers, security architects, AI engineers, platform teams, MLOps, DevSecOps, system owners, change managers, risk, compliance and assurance teams
Dependencies	GAISSF-NOR-001 through GAISSF-NOR-008; GAISSF-IMP-009 and GAISSF-IMP-010 where published and current
Precedence	GAISSF-NOR-001 governs framework interpretation; GAISSF-NOR-004 governs control records
Control baseline	59 controls across nine domains; Foundational = 52 D1-D8 controls; D9 additional/conditional where Physical AI or cyber-physical actuation is in scope
Status	Final Publication
Issued / Updated	Issued 29 June 2026; updated 05 October 2026
Contact	contact@oda3.org
Distribution	Public - docs.oda3.org / official ODA3 publication channels
Classification	Public

Legal and Reliance Notice

© 2026 ODA3 Pvt Ltd. Published market-facing by ODA3 Institute.

GAISSF™ is the framework mark used for the Global AI Safety and Security Framework. Third-party names and marks remain the property of their respective owners. Use of GAISSF™ materials is subject to the applicable GAISSF Ecosystem Licence (GEL) and published terms.

This playbook is informative implementation guidance. It does not amend GAISSF™ requirements, determine legal or regulatory compliance, replace system-specific safety or engineering obligations, or create certification eligibility. GAISSF certification is not currently offered; any future certification would require a separately issued scheme.

A deployment decision remains the responsibility of the organization and its designated accountable authorities. An independent assessor or reviewer may challenge the evidence and decision basis but does not assume the deployment authority.

Foreword

A controlled deployment converts an approved AI system design into an operating service without losing traceability, safety constraints, security controls or accountability. This playbook provides a practical sequence from rollout planning through pilot, production release, stabilization, handover and post-deployment assurance.

1. Purpose and Scope

This playbook applies to new AI systems, material model upgrades, significant architecture changes, new agentic capabilities, external tool integrations, major data-source changes and Physical AI deployments.

It covers deployment preparation, release strategy, change approval, pilot execution, production transition, rollback, operational acceptance and continuing assurance. Minor routine changes may use a proportionate workflow where organizational change policy permits.

Conformance-tier scope is governed by GAISSF-NOR-001. Foundational comprises the 52 canonical controls in D1-D8. D9 contains seven additional controls that become mandatory where Physical AI or cyber-physical actuation is in scope. Operational and Optimized retain the same applicable-control scope while increasing operating-effectiveness evidence, monitoring and assurance expectations.

2. Deployment Principles

Principle	Operational meaning
No production without scope	The deployed system, interfaces, data, users, suppliers and operating regions are explicitly defined.
No approval without evidence	Release decisions rely on current testing, risk and control evidence.

Principle	Operational meaning
Progressive exposure	Increase users, traffic, autonomy and consequence only after stage-gate criteria are met.
Reversible change	Every material deployment has a tested rollback, disablement or safe-state path.
Separation of duties	Build, approval, deployment and independent assessment or review are appropriately separated.
Observe before expanding	Telemetry and human review are used to validate actual behavior before scale-up.
Known limitations remain visible	Residual risks, constraints and unsupported use cases are communicated to operators and users.
Failed evidence is retained	Negative, inconclusive and failed results remain part of the decision record.

3. Rollout Model

Stage	Deployment state	Primary objective	Exit gate
0	Plan	Define scope, strategy, ownership and deployment risk	Approved rollout plan
1	Prepare	Complete controls, evidence, runbooks and environments	Readiness review passed
2	Validate	Execute pre-production technical and operational testing	Validation evidence accepted
3	Pilot	Expose a limited population under heightened oversight	Pilot success criteria met
4	Limited production	Expand traffic with constraints and rapid rollback	Stability criteria met
5	General production	Operate at approved scope and service level	Operational acceptance
6	Stabilize	Monitor post-release behavior and close deployment findings	Stabilization review passed
7	Handover	Transfer ownership to business-as-usual operations	Signed operational acceptance
8	Sustain	Maintain applicable-control effectiveness and reassess material change	Recurring assessment and improvement

4. Stage 0 - Plan the Deployment

4.1 Deployment plan minimum content

Planning element	Required content
Release identifier	Unique release, build, model and configuration identifiers
Scope	Systems, environments, users, geographies, interfaces and suppliers
Change description	What is new, modified, removed or inherited
Risk classification	Security, safety, privacy, legal, operational and reputational impact
Rollout strategy	Pilot, canary, blue-green, phased region, feature flag or controlled cutover
Owners	Release owner, system owner, control owners and approvers
Success criteria	Functional, security, safety, quality and operational thresholds
Stop criteria	Conditions requiring pause, rollback, containment or escalation
Rollback	Technical and business rollback sequence
Monitoring	Telemetry, dashboards, alert thresholds and review cadence
Communications	Operator, user, customer, regulator and stakeholder notices
Evidence	Required deployment and decision artifacts

4.2 Change impact assessment

1. Identify affected GAISSF™ controls and evidence packages.
2. Assess model, data, application, infrastructure, supplier and jurisdiction changes.
3. Determine whether prior testing remains valid.

4. Identify new or changed threat scenarios and failure modes.
5. Determine whether the Applicable-Control Baseline, assessment scope, any future-scheme certification scope, or public claims are affected.
6. Classify the change as routine, significant, major or emergency under organizational policy.

5. Stage 1 - Prepare the Release

Workstream	Preparation activities	Required output
Architecture	Confirm production topology, trust boundaries, access paths and dependencies	Approved architecture
Security	Harden environments, identities, secrets, network paths and logging	Security checklist
Data	Validate source, quality, lineage, privacy, retention and retrieval stores	Data readiness record
Model	Freeze model/version, provenance, configuration and evaluation baseline	Model release record
Application	Freeze prompt logic, orchestration, tools, policies and safeguards	Application release record
Operations	Prepare monitoring, support, incident, fallback and escalation runbooks	Operational runbooks
Evidence	Assemble release evidence and integrity records	Deployment evidence package
People	Train operators, support, approvers and affected users	Training record

5.1 Environment controls

- Separate development, test, staging and production environments.
- Restrict production access using least privilege and privileged-access monitoring.
- Use controlled configuration and secret-management processes.
- Prevent unapproved model, prompt, data or tool changes after release freeze.
- Synchronize release identifiers across code, model, configuration and evidence.

6. Stage 2 - Validate Before Production

Validation area	Minimum validation objective	Evidence
Functional	System performs approved use cases and rejects unsupported paths	Test results
Security	Threat scenarios and abuse cases are tested	Security evaluation
Robustness	Expected perturbations, malformed inputs and distribution changes are assessed	Robustness report
Privacy	Data flows, leakage, retention and user rights are verified	Privacy test record
Content safety	Harmful, prohibited and high-risk outputs are evaluated	Safety evaluation
Human oversight	Review, override and escalation mechanisms function	Oversight test
Resilience	Failover, recovery, degradation and rollback are tested	Resilience evidence
Observability	Required events, decisions and failures are logged	Telemetry verification
Physical safety	Safe state, override, actuation constraints and HIL/simulation are tested	Safety test package, using released/current reference assets or equivalent methods where applicable

Where a GAISSF Validation Test Suite (VTS) or other implementation asset is referenced, use it only where that asset is published, current and applicable to the deployment. Unreleased or internal tooling does not create a conformance requirement, and equivalent evidence-generating methods may be used unless a normative requirement states otherwise.

6.1 Pre-production gate

- All critical and high-risk test scenarios executed
- No unresolved release-blocking finding
- Residual risks documented and approved
- Evidence package complete and traceable

- Rollback tested
- Monitoring and alert thresholds verified
- Incident and communication runbooks exercised
- Release authority recorded

7. Stage 3 - Pilot Deployment

The pilot should limit exposure while producing representative evidence of actual operation. The pilot population, duration, transaction limits, autonomy, data and regions should be explicitly constrained.

Pilot control	Recommended practice
Population	Select a controlled user group with informed support channels
Traffic	Limit requests, transactions or physical operating hours
Autonomy	Restrict tool calls, approvals and consequential actions
Data	Use approved data classes and enhanced leakage monitoring
Human review	Increase review frequency and require manual approval for high-impact outputs
Telemetry	Enable detailed logging and daily control-health review
Support	Provide rapid incident triage and direct operator escalation
Decision	Use pre-defined success, pause and failure criteria

7.1 Pilot success criteria

Criterion	Example decision measure
Control stability	No repeated control failure above approved tolerance
Security	No unresolved exploit path or material abuse pattern
Safety	No uncontained high-severity harmful outcome
Quality	Performance remains within approved operating range
Operations	Support and escalation operate within service targets
Observability	Required events are captured and attributable
User impact	Complaints, confusion and misuse remain within approved thresholds
Rollback	Rollback remains available and tested

8. Stage 4 - Limited Production

Limited production expands real-world exposure while retaining enhanced controls. Expansion should be incremental and reversible.

Expansion dimension	Control method
Users	Percentage-based or role-based enablement
Geography	Region-by-region activation
Features	Feature flags and capability allowlists
Traffic	Rate caps and transaction ceilings
Autonomy	Tiered approval and tool-permission limits
Data	Progressive enablement of data classes
Physical operation	Restricted locations, speeds, loads or operating windows

8.1 Canary and blue-green patterns

Pattern	Use	Control concern
Canary	Expose a small percentage to the new release	Representative sampling and rapid rollback
Blue-green	Maintain old and new production stacks	Configuration parity and controlled cutover
Shadow	Run new system without affecting decisions	Data protection and non-action assurance
Feature flag	Enable selected capabilities independently	Flag governance and configuration evidence
Regional wave	Deploy by geography or business unit	Jurisdiction and dependency differences

9. Stage 5 - General Production Release

9.1 Go-live decision package

Decision input	Minimum content
Scope confirmation	Approved final production boundary
Control status	Applicable control implementation and open findings
Validation	Pre-production, pilot and limited-production results
Risk	Residual risks, exceptions and acceptance authority
Operations	Support, monitoring, incident, continuity and supplier readiness
Rollback	Current rollback point and safe-state confirmation
Communications	User disclosures and stakeholder notices
Approval	Named release authority, date and conditions

9.2 Go-live checklist

- Release identifiers verified
- Production configuration matches approved baseline
- Monitoring dashboards active
- On-call and escalation contacts confirmed
- User and operator documentation published
- Supplier dependencies available
- Public claims reviewed; any future-scheme certification claim or mark use separately authorized where applicable
- Rollback point preserved
- Approval recorded before activation

10. Rollback, Containment and Safe-State Procedures

Trigger	Immediate response	Decision owner
Critical security exploit	Disable affected function, isolate systems and initiate incident response	Security incident authority
Material harmful output	Suspend feature, constrain model or require human approval	System owner / safety authority
Privacy breach	Contain data flow, revoke access and activate breach process	Privacy and incident authority
Loss of monitoring	Pause expansion or revert to last observable release	Operations owner
Model degradation	Revert model/configuration or restrict use cases	Model owner
Supplier outage/change	Fail over, degrade safely or suspend dependent capability	Service owner
Physical safety threshold breach	Enter safe state and stop actuation	Safety authority

For Physical AI or cyber-physical actuation, rollback may require transition to a defined safe state rather than restoration of a prior software release. Safety authority, physical constraints and command-integrity controls should be explicit before activation.

10.1 Rollback record

- Trigger and detection time
- Affected release and scope
- Decision authority
- Actions taken
- Data and transaction handling
- Customer or regulator notification
- Evidence preserved
- Recovery validation
- Root-cause and corrective-action reference

11. Stage 6 - Stabilization

Time window	Primary activities
First 24 hours	Continuous telemetry review, incident watch, rollback readiness and release verification
Days 2-7	Daily control-health review, user feedback, supplier and performance monitoring
Weeks 2-4	Trend analysis, finding closure, evidence consolidation and operational tuning
End of stabilization	Formal review of success criteria, residual risk and handover readiness

11.1 Stabilization review questions

- Did actual use remain within approved scope?
- Did any control fail, degrade or require manual compensation?
- Were alerts actionable and reviewed?
- Did user behavior reveal new misuse or confusion?
- Did suppliers or data sources change?
- Did the release affect the Applicable-Control Baseline, assessment scope, future-scheme certification scope, or public claims?
- Are operational teams ready to accept ownership?

12. Stage 7 - Operational Handover

Handover item	Acceptance evidence
Ownership	Named system, service, model, data and control owners
Runbooks	Approved operating, incident, fallback and recovery procedures
Monitoring	Dashboards, alerts, thresholds and review responsibilities
Access	Production roles and privileged-access review
Evidence	Repository, index, retention and refresh ownership
Suppliers	Escalation, support, notification and continuity contacts
Training	Operator and support competence records
Backlog	Known defects, improvements and corrective actions
Risk	Residual-risk and exception register
Acceptance	Signed operational acceptance record

13. Stage 8 - Sustain and Improve

Trigger	Required reassessment
Model version change	Provenance, regression, security, safety and performance
Prompt/orchestration change	Abuse, authorization, output and workflow testing
New tool or agent	Permissions, transactions, delegation and containment
New data source	Lineage, privacy, quality, poisoning and licence
Infrastructure change	Hardening, identity, segmentation, resilience and logging
New jurisdiction	Legal, regulatory, data-transfer and disclosure analysis
Incident or near miss	Control effectiveness, root cause and corrective action
GAISSF™ update or correction notice	Version impact, Applicable-Control Baseline and migration assessment

14. Deployment Roles and Decision Rights

Role	Deployment responsibility
Executive sponsor	Resolve material scope, funding and risk issues
Release authority	Approve or reject production activation
System owner	Own business use, scope and residual risk
Security lead	Approve security readiness and containment
Safety lead	Approve safety case and physical-AI constraints where applicable

Role	Deployment responsibility
Model owner	Own model provenance, evaluation and version
Data owner	Approve data use, quality, privacy and retention
Operations lead	Accept monitoring, support and continuity
Independent assessor or reviewer	Challenge evidence, scope and gate decisions without assuming management authority
Change manager	Maintain release and rollback records

15. Deployment Evidence Package

Evidence group	Artifacts
Planning	Deployment plan, scope, strategy, impact assessment
Build	Release manifests, model/data/configuration identifiers
Validation	Functional, security, safety, privacy and resilience tests
Approval	Gate decisions, exceptions and residual-risk acceptance
Execution	Deployment logs, feature-flag records and environment evidence
Pilot	Pilot metrics, issues, reviews and decision record
Production	Go-live checklist, activation record and monitoring evidence
Stabilization	Control-health reviews, incidents and corrective actions
Handover	Operational acceptance and owner acknowledgements
Integrity	Hashes, signatures, timestamps and repository audit records

16. Deployment Metrics

Metric	Use	Caution
Deployment success rate	Track release reliability	Do not hide severity of failed releases
Rollback frequency	Detect instability or weak validation	Some rollback use demonstrates control effectiveness
Time to detect	Evaluate observability	Depends on event visibility
Time to contain	Evaluate incident readiness	Measure by severity class
Control-health failures	Identify operational degradation	Normalize by volume and scope
Pilot-to-production defects	Evaluate pilot representativeness	Classify materiality
Unapproved configuration drift	Detect release-control weakness	Requires authoritative baseline
Evidence completeness	Assess deployment traceability	Count does not equal quality

17. Emergency Deployment Procedure

Emergency deployment may shorten normal sequencing but shall not eliminate accountability, logging, rollback or safe-state capability, risk acceptance or retrospective review.

1. Document the emergency condition and affected scope.
2. Identify minimum mandatory validation and containment.
3. Obtain authorized emergency approval.
4. Preserve the pre-change state and rollback or safe-state option.
5. Enable enhanced monitoring.
6. Complete omitted documentation and testing after stabilization.
7. Conduct retrospective review and corrective action.

18. Common Deployment Failure Modes

Failure mode	Consequence	Prevention
Uncontrolled model update	Evidence and behavior no longer match approval	Pin versions and enforce release manifests
Pilot not representative	Production failures appear after scale-up	Select realistic users, data and workloads
Rollback untested	Containment fails during incident	Test rollback before release
Telemetry activated late	Early failures are invisible	Verify observability before pilot

Failure mode	Consequence	Prevention
Feature flags unmanaged	Unauthorized capabilities become active	Control flag ownership and logging
Supplier change unnoticed	Inherited controls silently change	Contractual notification and dependency monitoring
Handover assumed	No team owns post-release risks	Require signed operational acceptance
Success-only reporting	Negative evidence is lost	Retain failed and inconclusive results

19. Detailed Rollout Work Breakdown

Work package	Tasks	Owner	Exit evidence
Release planning	Scope, risk, strategy, schedule, dependencies	Release manager	Approved plan
Environment preparation	Access, secrets, network, logging, baseline	Platform owner	Environment checklist
Validation	Functional, security, safety, privacy, resilience	Test lead	Signed validation report
Pilot	Users, limits, oversight, support, telemetry	System owner	Pilot decision record
Scale-up	Traffic, feature, region and autonomy waves	Operations lead	Wave gate approvals
Handover	Runbooks, ownership, training and backlog	Service owner	Operational acceptance

20. Decision Logic for Pause, Rollback and Safe State

A pause is appropriate when evidence is incomplete, monitoring is impaired or a threshold breach has not created immediate material exposure. Rollback or safe-state activation is appropriate when a material security, safety, privacy or integrity condition exists and continued operation increases harm.

Condition	Default decision
Critical exploit or unsafe actuation	Contain and roll back or enter safe state
Material harmful output with uncertain scope	Pause affected capability and investigate
Monitoring unavailable for high-risk action	Restrict or suspend until observability returns
Minor performance degradation within risk tolerance	Continue under enhanced monitoring
Supplier model change without evaluation	Hold expansion and revalidate

21. Communications Plan

Audience	Message content	Trigger
Operators	Release, known limitations, monitoring and escalation	Before activation
Users	Capability, restrictions, disclosures and recourse	Before access
Executives	Risk, residual issues and go/no-go decision	Each material gate
Customers	Service impact and material change	Contractual trigger
Regulators	Required incident or change information	Legal trigger

22. Post-Deployment Validation

- Compare actual traffic and users with approved scope.
- Re-run critical tests against production configuration.
- Validate alert completeness and operator response.
- Review user feedback and unexpected use.
- Confirm evidence package matches the deployed release.
- Close or reclassify pilot and stabilization findings.

23. Physical AI Deployment Boundary

Where Physical AI or cyber-physical actuation is in scope, GAISSF™ Domain 9 applies in addition to the applicable D1-D8 baseline. PAI-SF™ v1.0 should be used alongside GAISSF™ D9 where relevant to physical/kinetic system risks.

CSX-PAISF-004 is a mapping only. It may show lineage or correspondence between GAISSF™ and PAI-SF™ concepts, but it does not establish equivalence, evidence sufficiency, automatic conformance or certification status.

24. Limitations

- This playbook does not prescribe a universal rollout duration.
- It does not replace sector-specific safety, quality or regulatory release processes.
- A successful pilot does not prove absence of low-frequency or emerging failures.
- Rollback may be constrained where irreversible transactions or physical actions occur.
- Deployment evidence remains subject to sampling, completeness and integrity limitations.
- Deployment readiness does not establish legal compliance, system safety, security, or eligibility for any future GAISSF certification scheme.

25. Notably Absent

- No authorization to deploy merely because a checklist is complete.
- No assumption that canary deployment is safe for every high-consequence system.
- No waiver of human oversight, risk acceptance or incident obligations.
- No promise that progressive rollout prevents all harmful outcomes.
- No substitution for independent physical-safety validation where such validation is required.
- No permission to erase failed, negative or inconclusive deployment evidence.
- No representation that an ODA3 certification service or certification mark is currently available.

Annex A - Deployment Plan Template

Field	Entry
Release identifier	
Scope	
Change description	
Risk classification	
Rollout strategy	
Owners	
Success criteria	
Stop criteria	
Rollback	
Monitoring	
Communications	
Evidence	
Approval	

Annex B - Stage-Gate Record

Gate	Criteria met	Open issues	Decision	Approver	Date
Plan					
Prepare					
Validate					
Pilot					
Limited production					
General production					
Stabilization					

Gate	Criteria met	Open issues	Decision	Approver	Date
Handover					

Annex C - Rollback Checklist

- Rollback authority confirmed
- Last known good release available
- Data and transaction consequences assessed
- Supplier dependencies addressed
- Monitoring active during rollback
- Users and stakeholders notified
- Evidence preserved
- Recovery validation completed
- Root-cause review initiated

Annex D - Operational Acceptance Template

Acceptance area	Status	Evidence	Owner sign-off
Monitoring			
Incident response			
Support			
Access and privileges			
Evidence maintenance			
Supplier management			
Known risks and exceptions			
Training			

Annex E - Source Traceability

Playbook subject	Controlled source
Normative requirements	GAISSF-NOR-001
Control records	GAISSF-NOR-004
Terminology	GAISSF-NOR-005
Release status	GAISSF-NOR-007
Historical decisions	GAISSF-NOR-008
Adoption program	GAISSF-IMP-009 - where published and current
Implementation architecture and workflows	GAISSF-IMP-010 - where published and current
Physical AI companion	PAI-SF™ v1.0, where applicable and current
GAISSF / PAI-SF mapping	CSX-PAISF-004 - mapping only; no equivalence, evidence-sufficiency or automatic-conformance effect

Annex F - Publication Record

Version	Date	Change	Status
1.0	29 June 2026	Initial deployment playbook for GAISSF™ v1.0	Final Publication
1.0	05 October 2026	Reconciled profile/D9, Applicable-Control Baseline, assessment/certification/VTS wording, R2 placement, PAI-SF™ boundary, dependencies and publication metadata; no GAISSF™ control changed.	Updated