CANONICAL LIVING OVERVIEW GUIDE

Software and AI Supply-Chain Security

A software and AI supply-chain guide covering source, dependencies, build systems, CI/CD, provenance, signing, SBOMs, VEX, model and dataset lineage, release, distribution, updates, suppliers, validation, incident response, and governed automation.

PERMANENT ID
MYTHOS-FG-SEC-0014

PUBLICATION CLASS
FIELD GUIDE

VERSION / STATUS
1.0.0-candidate / CANDIDATE

PERMANENT PUBLICATION IDENTITY

Software and AI Supply-Chain Security

Document ID
MYTHOS-FG-SEC-0014

Version
1.0.0-candidate

Status
Candidate Focused Field Guide

Review date
2026-10-26

Publication position
This is a living field guide. It is designed for frequent revision while retaining a stable permanent identity. Candidate status means the content is ready for structured use and review but is not represented as an external certification or authority publication.

LEVEL	CLASS	CADENCE	EVIDENCE
L1-L4	FIELD GUIDE	QUARTERLY	SOURCE-BOUND
Foundational through frontier	Practical and educational	Interim updates as needed	Limitations remain visible

Reader orientation

Dimension	Engineering position
Primary domain	MYTHOS-SEC - Cybersecurity, Security Engineering & Cryptography
Secondary domain	MYTHOS-SWE; MYTHOS-AI; MYTHOS-CLD
Audience	Software, platform, DevSecOps, application security, AI engineering, procurement, product, and supply-chain security teams.
Prerequisites	Software development, CI/CD, packages, containers, cloud, cryptography, identity, and risk fundamentals.
Publisher / owner	Mythos Systems / Aaron Stovall
Public classification	Public technical guidance

Expected outcomes

Map the complete software and AI artifact chain from source and data through deployment and update.

Protect build, dependency, signing, release, model, and distribution authority.

Generate and verify provenance, signatures, SBOMs, model metadata, and policy evidence.

Reduce supplier, dependency, CI/CD, model, dataset, plugin, and update risk.

Respond to compromised components or build systems with traceable scope and recovery.

How to use this guide

Read the guide as a progressive system rather than a disconnected encyclopedia. Foundational material establishes the mental model; practitioner sections convert it into repeatable implementation and operations; advanced sections expose architecture, scale, security, and failure behavior; frontier sections describe technologies whose evidence or maturity is still evolving.

Suggested reading path
New practitioners should follow the chapters in order and complete the laboratories. Experienced engineers may begin with the architecture, failure, validation, or frontier sections, then use the related-document map to deepen a specific topic.

Learning and assurance model

Level	Reader outcome
L1 - Foundational	Terminology, first principles, component roles, packet or state flow, and simple working examples.
L2 - Practitioner	Implementation, configuration, automation, testing, troubleshooting, and operational evidence.
L3 - Advanced	Architecture, scale, concurrency, resilience, threat models, performance, and complex migrations.
L4 - Frontier	Emerging systems, research directions, technical uncertainty, readiness gates, and adoption signals.

Contents

CHAPTER 01

Supply-chain architecture and threat model

Model every authority, artifact, transformation, and dependency in the delivery path.

Chapter operating position
Model every authority, artifact, transformation, and dependency in the delivery path.

1.1 Artifact and lifecycle inventory

Artifact and lifecycle inventory must be implemented as an observable engineering mechanism, not a product label or maturity claim. Identify source, dependencies, tools, containers, packages, firmware, models, datasets, prompts, plugins, configurations, and release assets. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for artifact and lifecycle inventory.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating artifact and lifecycle inventory as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for artifact and lifecycle inventory.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

1.2 Actors and authority

Actors and authority must be implemented as an observable engineering mechanism, not a product label or maturity claim. Map developers, maintainers, reviewers, bots, builders, registries, signers, vendors, mirrors, deployers, and consumers. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for actors and authority.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Implementation sketch

```
release:
  source_digest: sha256:...
  builder_identity: https://ci.example/workflows/release
  build_type: hermetic
  artifact_digest: sha256:...
  provenance: signed
  sbom:
    formats: [spdx, cyclonedx]
    verified_against_artifact: true
  policy:
    require_reviewed_source: true
    require_pinned_dependencies: true
    require_signature: true
    reject_unknown_builder: true
  rollback:
    previous_verified_release: sha256:...
```

Failure patterns

- Treating actors and authority as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for actors and authority.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

1.3 Threats and compromise paths

Threats and compromise paths must be implemented as an observable engineering mechanism, not a product label or maturity claim. Analyze source tampering, dependency confusion, account takeover, poisoned builds, malicious maintainers, model poisoning, and update abuse. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for threats and compromise paths.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating threats and compromise paths as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for threats and compromise paths.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

1.4 Trust boundaries and evidence

Trust boundaries and evidence must be implemented as an observable engineering mechanism, not a product label or maturity claim. Define repositories, runners, networks, secrets, registries, signing, transparency, deployment, and runtime boundaries. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

Define authority, ownership, scope, dependencies, limits, and acceptance criteria for trust boundaries and evidence.

Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.

Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.

Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.

Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.

Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

Treating trust boundaries and evidence as a product feature instead of an end-to-end engineering system.

Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for trust boundaries and evidence.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

Chapter 01 laboratory and review

Practical laboratory
Create a complete supply-chain graph for one software and AI-enabled product, including hidden tools and third-party services.

Laboratory evidence package

- Architecture or topology showing boundaries, identities, dependencies, and authoritative inputs.
- Versioned configuration or code with explanatory comments and reproducible setup instructions.
- Nominal-path validation plus at least two injected failure or boundary conditions.
- Structured logs, traces, metrics, packet captures, test output, or other appropriate evidence.
- A concise decision record covering findings, limitations, residual risks, and next actions.

Frontier extension

Extend the model to autonomous coding agents, model marketplaces, hardware accelerators, and post-quantum signing.

Review gate
Advance only when another engineer can reproduce the result, explain the architecture, identify the failure boundary, and distinguish measured evidence from assumptions or projections.

CHAPTER 02

Source, dependency, and contributor security

Protect the inputs and identities that enter the development process.

Chapter operating position
Protect the inputs and identities that enter the development process.

2.1 Repository and branch protection

Repository and branch protection must be implemented as an observable engineering mechanism, not a product label or maturity claim. Use strong identity, protected branches, review, signed commits where appropriate, ownership, status checks, and immutable history. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for repository and branch protection.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating repository and branch protection as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for repository and branch protection.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

2.2 Dependency selection and pinning

Dependency selection and pinning must be implemented as an observable engineering mechanism, not a product label or maturity claim. Control registries, names, versions, lockfiles, hashes, mirrors, typosquatting, provenance, and maintenance risk. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for dependency selection and pinning.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Implementation sketch

```
release:
  source_digest: sha256:...
  builder_identity: https://ci.example/workflows/release
  build_type: hermetic
  artifact_digest: sha256:...
  provenance: signed
  sbom:
    formats: [spdx, cyclonedx]
    verified_against_artifact: true
  policy:
    require_reviewed_source: true
    require_pinned_dependencies: true
    require_signature: true
    reject_unknown_builder: true
  rollback:
    previous_verified_release: sha256:...
```

Failure patterns

- Treating dependency selection and pinning as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for dependency selection and pinning.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

2.3 Contributor and maintainer trust

Contributor and maintainer trust must be implemented as an observable engineering mechanism, not a product label or maturity claim. Manage onboarding, roles, review authority, tokens, automation, dormant accounts, transfers, and emergency revocation. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for contributor and maintainer trust.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating contributor and maintainer trust as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for contributor and maintainer trust.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

2.4 Secrets and development environments

Secrets and development environments must be implemented as an observable engineering mechanism, not a product label or maturity claim. Prevent secret leakage across local systems, IDEs, agents, logs, CI, tests, tickets, and generated artifacts. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

Define authority, ownership, scope, dependencies, limits, and acceptance criteria for secrets and development environments.

Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.

Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.

Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.

Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.

Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

Treating secrets and development environments as a product feature instead of an end-to-end engineering system.

Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for secrets and development environments.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

Chapter 02 laboratory and review

Practical laboratory
Audit a repository for branch, dependency, identity, token, secret, and automation risk; implement controls and verify bypass attempts.

Laboratory evidence package

- Architecture or topology showing boundaries, identities, dependencies, and authoritative inputs.
- Versioned configuration or code with explanatory comments and reproducible setup instructions.
- Nominal-path validation plus at least two injected failure or boundary conditions.
- Structured logs, traces, metrics, packet captures, test output, or other appropriate evidence.
- A concise decision record covering findings, limitations, residual risks, and next actions.

Frontier extension

Explore reproducible developer environments, ephemeral workstations, signed agent actions, and capability-based contribution.

Review gate
Advance only when another engineer can reproduce the result, explain the architecture, identify the failure boundary, and distinguish measured evidence from assumptions or projections.

CHAPTER 03

Build-system and CI/CD security

Treat the build plane as high-privilege production infrastructure.

Chapter operating position
Treat the build plane as high-privilege production infrastructure.

3.1 Runner and builder isolation

Runner and builder isolation must be implemented as an observable engineering mechanism, not a product label or maturity claim. Use ephemeral execution, least privilege, network restrictions, clean state, trusted images, resource controls, and secret isolation. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for runner and builder isolation.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating runner and builder isolation as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for runner and builder isolation.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

3.2 Pipeline identity and authorization

Pipeline identity and authorization must be implemented as an observable engineering mechanism, not a product label or maturity claim. Bind jobs to source, branch, event, workflow, environment, reviewer, and deployment target. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for pipeline identity and authorization.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Implementation sketch

```
release:
  source_digest: sha256:...
  builder_identity: https://ci.example/workflows/release
  build_type: hermetic
  artifact_digest: sha256:...
  provenance: signed
  sbom:
    formats: [spdx, cyclonedx]
    verified_against_artifact: true
  policy:
    require_reviewed_source: true
    require_pinned_dependencies: true
    require_signature: true
    reject_unknown_builder: true
  rollback:
    previous_verified_release: sha256:...
```

Failure patterns

- Treating pipeline identity and authorization as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for pipeline identity and authorization.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

3.3 Hermeticity and reproducibility

Hermeticity and reproducibility must be implemented as an observable engineering mechanism, not a product label or maturity claim. Control inputs, network access, clocks, randomness, compilers, dependencies, environment, and output comparison. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for hermeticity and reproducibility.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating hermeticity and reproducibility as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for hermeticity and reproducibility.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

3.4 Pipeline testing and failure handling

Pipeline testing and failure handling must be implemented as an observable engineering mechanism, not a product label or maturity claim. Test policy, credentials, malicious pull requests, untrusted forks, poisoned caches, compromised actions, and rollback. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

Define authority, ownership, scope, dependencies, limits, and acceptance criteria for pipeline testing and failure handling.

Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.

Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.

Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.

Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.

Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

Treating pipeline testing and failure handling as a product feature instead of an end-to-end engineering system.

Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for pipeline testing and failure handling.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

Chapter 03 laboratory and review

Practical laboratory
Build a pipeline threat model and inject a malicious dependency, untrusted pull request, leaked token, and poisoned cache in a safe lab.

Laboratory evidence package

- Architecture or topology showing boundaries, identities, dependencies, and authoritative inputs.
- Versioned configuration or code with explanatory comments and reproducible setup instructions.
- Nominal-path validation plus at least two injected failure or boundary conditions.
- Structured logs, traces, metrics, packet captures, test output, or other appropriate evidence.
- A concise decision record covering findings, limitations, residual risks, and next actions.

Frontier extension

Explore confidential builds, hardware-attested runners, deterministic agentic coding, and verifiable distributed build farms.

Review gate
Advance only when another engineer can reproduce the result, explain the architecture, identify the failure boundary, and distinguish measured evidence from assumptions or projections.

CHAPTER 04

Provenance, signing, transparency, and SLSA

Create verifiable statements about where artifacts came from and how they were built.

Chapter operating position
Create verifiable statements about where artifacts came from and how they were built.

4.1 Build provenance

Build provenance must be implemented as an observable engineering mechanism, not a product label or maturity claim. Record source, digest, builder identity, invocation, parameters, dependencies, environment, outputs, and timestamps. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for build provenance.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating build provenance as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for build provenance.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

4.2 Artifact and metadata signing

Artifact and metadata signing must be implemented as an observable engineering mechanism, not a product label or maturity claim. Use identity-bound signing, key custody, short-lived certificates, offline or hardware keys, and verification policy. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for artifact and metadata signing.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Implementation sketch

```
release:
  source_digest: sha256:...
  builder_identity: https://ci.example/workflows/release
  build_type: hermetic
  artifact_digest: sha256:...
  provenance: signed
  sbom:
    formats: [spdx, cyclonedx]
    verified_against_artifact: true
  policy:
    require_reviewed_source: true
    require_pinned_dependencies: true
    require_signature: true
    reject_unknown_builder: true
  rollback:
    previous_verified_release: sha256:...
```

Failure patterns

- Treating artifact and metadata signing as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for artifact and metadata signing.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

4.3 Transparency and accountability

Transparency and accountability must be implemented as an observable engineering mechanism, not a product label or maturity claim. Publish append-only records, inclusion proof, monitor misuse, handle privacy, and support revocation or correction. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for transparency and accountability.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating transparency and accountability as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for transparency and accountability.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

4.4 SLSA adoption and verification

SLSA adoption and verification must be implemented as an observable engineering mechanism, not a product label or maturity claim. Select appropriate build and source properties, document gaps, verify consumers, and avoid badge-only compliance. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

Define authority, ownership, scope, dependencies, limits, and acceptance criteria for slsa adoption and verification.

Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.

Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.

Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.

Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.

Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

Treating slsa adoption and verification as a product feature instead of an end-to-end engineering system.

Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for slsa adoption and verification.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

Chapter 04 laboratory and review

Practical laboratory
Generate provenance and signatures for an artifact, verify them in a clean environment, then simulate source, builder, and identity mismatch.

Laboratory evidence package

- Architecture or topology showing boundaries, identities, dependencies, and authoritative inputs.
- Versioned configuration or code with explanatory comments and reproducible setup instructions.
- Nominal-path validation plus at least two injected failure or boundary conditions.
- Structured logs, traces, metrics, packet captures, test output, or other appropriate evidence.
- A concise decision record covering findings, limitations, residual risks, and next actions.

Frontier extension

Evaluate threshold signing, post-quantum signatures, confidential provenance generation, and machine-verifiable release policy.

Review gate
Advance only when another engineer can reproduce the result, explain the architecture, identify the failure boundary, and distinguish measured evidence from assumptions or projections.

CHAPTER 05

SBOM, VEX, inventory, and dependency intelligence

Represent what is inside products and what security conclusions can legitimately be drawn.

Chapter operating position

Represent what is inside products and what security conclusions can legitimately be drawn.

5.1 SBOM scope and formats

SBOM scope and formats must be implemented as an observable engineering mechanism, not a product label or maturity claim. Describe components, versions, hashes, relationships, licenses, suppliers, build, services, files, models, and incomplete coverage. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for sbom scope and formats.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating sbom scope and formats as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for sbom scope and formats.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

5.2 Generation and validation

Generation and validation must be implemented as an observable engineering mechanism, not a product label or maturity claim. Combine source, build, package, image, binary, runtime, and supplier evidence while detecting omissions and drift. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for generation and validation.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Implementation sketch

```
release:
  source_digest: sha256:...
  builder_identity: https://ci.example/workflows/release
  build_type: hermetic
  artifact_digest: sha256:...
  provenance: signed
  sbom:
    formats: [spdx, cyclonedx]
    verified_against_artifact: true
  policy:
    require_reviewed_source: true
    require_pinned_dependencies: true
    require_signature: true
    reject_unknown_builder: true
  rollback:
    previous_verified_release: sha256:...
```

Failure patterns

- Treating generation and validation as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for generation and validation.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

5.3 VEX and exploitability statements

VEX and exploitability statements must be implemented as an observable engineering mechanism, not a product label or maturity claim. Record affected, not affected, fixed, or under investigation with justification, scope, authority, and expiry. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for vex and exploitability statements.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating vex and exploitability statements as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for vex and exploitability statements.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

5.4 Operational integration

Operational integration must be implemented as an observable engineering mechanism, not a product label or maturity claim. Link inventory to vulnerabilities, policy, procurement, deployment, incident, cryptography, and end-of-life decisions. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

Define authority, ownership, scope, dependencies, limits, and acceptance criteria for operational integration.

Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.

Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.

Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.

Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.

Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

Treating operational integration as a product feature instead of an end-to-end engineering system.

Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for operational integration.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

Chapter 05 laboratory and review

Practical laboratory
Generate SPDX- or CycloneDX-style inventories from source and build evidence, find a missing transitive dependency, and create a justified VEX record.

Laboratory evidence package

- Architecture or topology showing boundaries, identities, dependencies, and authoritative inputs.
- Versioned configuration or code with explanatory comments and reproducible setup instructions.
- Nominal-path validation plus at least two injected failure or boundary conditions.
- Structured logs, traces, metrics, packet captures, test output, or other appropriate evidence.
- A concise decision record covering findings, limitations, residual risks, and next actions.

Frontier extension

Explore real-time runtime bills of materials, cryptographic bills of materials, model BOMs, and signed dependency graphs.

Review gate
Advance only when another engineer can reproduce the result, explain the architecture, identify the failure boundary, and distinguish measured evidence from assumptions or projections.

CHAPTER 06

AI model, dataset, and agent supply chains

Protect artifacts and processes unique to machine learning and agentic systems.

Chapter operating position
Protect artifacts and processes unique to machine learning and agentic systems.

6.1 Data lineage and governance

Data lineage and governance must be implemented as an observable engineering mechanism, not a product label or maturity claim. Record sources, licenses, consent, transformations, filtering, labeling, poisoning controls, versions, access, and deletion. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for data lineage and governance.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating data lineage and governance as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for data lineage and governance.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

6.2 Model and training pipeline integrity

Model and training pipeline integrity must be implemented as an observable engineering mechanism, not a product label or maturity claim. Protect code, parameters, checkpoints, weights, optimizers, evaluation, hardware, distributed jobs, and experiment records. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for model and training pipeline integrity.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Implementation sketch

```
release:
  source_digest: sha256:...
  builder_identity: https://ci.example/workflows/release
  build_type: hermetic
  artifact_digest: sha256:...
  provenance: signed
  sbom:
    formats: [spdx, cyclonedx]
    verified_against_artifact: true
  policy:
    require_reviewed_source: true
    require_pinned_dependencies: true
    require_signature: true
    reject_unknown_builder: true
  rollback:
    previous_verified_release: sha256:...
```

Failure patterns

- Treating model and training pipeline integrity as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for model and training pipeline integrity.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

6.3 Model distribution and loading

Model distribution and loading must be implemented as an observable engineering mechanism, not a product label or maturity claim. Verify model identity, format, code execution, custom operators, quantization, adapters, tokenizers, and runtime compatibility. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for model distribution and loading.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating model distribution and loading as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for model distribution and loading.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

6.4 Agent tools, skills, and plugins

Agent tools, skills, and plugins must be implemented as an observable engineering mechanism, not a product label or maturity claim. Govern capability packages, MCP servers, prompts, policies, credentials, updates, sandboxing, and delegated authority. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

Define authority, ownership, scope, dependencies, limits, and acceptance criteria for agent tools, skills, and plugins.

Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.

Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.

Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.

Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.

Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

Treating agent tools, skills, and plugins as a product feature instead of an end-to-end engineering system.

Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for agent tools, skills, and plugins.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

Chapter 06 laboratory and review

Practical laboratory
Build a model release record with dataset lineage, training provenance, evaluation evidence, signature, runtime dependencies, and rollback.

Laboratory evidence package

- Architecture or topology showing boundaries, identities, dependencies, and authoritative inputs.
- Versioned configuration or code with explanatory comments and reproducible setup instructions.
- Nominal-path validation plus at least two injected failure or boundary conditions.
- Structured logs, traces, metrics, packet captures, test output, or other appropriate evidence.
- A concise decision record covering findings, limitations, residual risks, and next actions.

Frontier extension

Explore confidential model training, verifiable inference, model watermarking, and agent capability transparency.

Review gate
Advance only when another engineer can reproduce the result, explain the architecture, identify the failure boundary, and distinguish measured evidence from assumptions or projections.

CHAPTER 07

Distribution, deployment, updates, and suppliers

Preserve integrity after the build leaves the producer.

Chapter operating position
Preserve integrity after the build leaves the producer.

7.1 Registries, repositories, and mirrors

Registries, repositories, and mirrors must be implemented as an observable engineering mechanism, not a product label or maturity claim. Secure authentication, authorization, immutability, signing, retention, replication, compromise detection, and recovery. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for registries, repositories, and mirrors.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating registries, repositories, and mirrors as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for registries, repositories, and mirrors.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

7.2 Deployment verification

Deployment verification must be implemented as an observable engineering mechanism, not a product label or maturity claim. Enforce signatures, provenance, policy, environment, configuration, secrets, identity, canaries, and rollback before activation. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for deployment verification.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Implementation sketch

```
release:
  source_digest: sha256:...
  builder_identity: https://ci.example/workflows/release
  build_type: hermetic
  artifact_digest: sha256:...
  provenance: signed
  sbom:
    formats: [spdx, cyclonedx]
    verified_against_artifact: true
  policy:
    require_reviewed_source: true
    require_pinned_dependencies: true
    require_signature: true
    reject_unknown_builder: true
  rollback:
    previous_verified_release: sha256:...
```

Failure patterns

- Treating deployment verification as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for deployment verification.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

7.3 Update and rollback security

Update and rollback security must be implemented as an observable engineering mechanism, not a product label or maturity claim. Protect manifests, channels, staged rollout, anti-rollback, emergency fixes, offline systems, and failed updates. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for update and rollback security.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating update and rollback security as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for update and rollback security.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

7.4 Supplier and acquisition assurance

Supplier and acquisition assurance must be implemented as an observable engineering mechanism, not a product label or maturity claim. Set requirements for development, evidence, vulnerabilities, support, notification, access, sub-suppliers, and exit. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

Define authority, ownership, scope, dependencies, limits, and acceptance criteria for supplier and acquisition assurance.

Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.

Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.

Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.

Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.

Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

Treating supplier and acquisition assurance as a product feature instead of an end-to-end engineering system.

Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for supplier and acquisition assurance.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

Chapter 07 laboratory and review

Practical laboratory
Design a secure release and update system, then simulate registry compromise, revoked signing identity, rollback attack, and supplier emergency.

Laboratory evidence package

- Architecture or topology showing boundaries, identities, dependencies, and authoritative inputs.
- Versioned configuration or code with explanatory comments and reproducible setup instructions.
- Nominal-path validation plus at least two injected failure or boundary conditions.
- Structured logs, traces, metrics, packet captures, test output, or other appropriate evidence.
- A concise decision record covering findings, limitations, residual risks, and next actions.

Frontier extension

Explore peer-to-peer verified distribution, sovereign mirrors, transparency federation, and quantum-safe update signing.

Review gate
Advance only when another engineer can reproduce the result, explain the architecture, identify the failure boundary, and distinguish measured evidence from assumptions or projections.

CHAPTER 08

Assurance, incidents, and continuous improvement

Detect compromise, scope affected artifacts, recover trust, and improve the system.

Chapter operating position
Detect compromise, scope affected artifacts, recover trust, and improve the system.

8.1 Continuous verification

Continuous verification must be implemented as an observable engineering mechanism, not a product label or maturity claim. Monitor source, dependencies, builder identity, provenance, signatures, SBOM drift, policy, registries, deployment, and runtime. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for continuous verification.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating continuous verification as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for continuous verification.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

8.2 Supply-chain incident response

Supply-chain incident response must be implemented as an observable engineering mechanism, not a product label or maturity claim. Revoke identities, stop releases, quarantine artifacts, determine scope, notify consumers, rebuild, replace trust, and verify recovery. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for supply-chain incident response.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Implementation sketch

```
release:
  source_digest: sha256:...
  builder_identity: https://ci.example/workflows/release
  build_type: hermetic
  artifact_digest: sha256:...
  provenance: signed
  sbom:
    formats: [spdx, cyclonedx]
    verified_against_artifact: true
  policy:
    require_reviewed_source: true
    require_pinned_dependencies: true
    require_signature: true
    reject_unknown_builder: true
  rollback:
    previous_verified_release: sha256:...
```

Failure patterns

- Treating supply-chain incident response as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.
- Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.
- Changing several variables before preserving evidence and stating a falsifiable hypothesis.
- Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.
- Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for supply-chain incident response.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

8.3 Metrics and maturity

Metrics and maturity must be implemented as an observable engineering mechanism, not a product label or maturity claim. Measure verified builds, unsigned artifacts, dependency freshness, critical exposures, provenance gaps, mean recovery, and supplier performance. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

- Define authority, ownership, scope, dependencies, limits, and acceptance criteria for metrics and maturity.
- Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.
- Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.
- Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.
- Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.
- Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

- Treating metrics and maturity as a product feature instead of an end-to-end engineering system.
- Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for metrics and maturity.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

8.4 Automation and AI assistance

Automation and AI assistance must be implemented as an observable engineering mechanism, not a product label or maturity claim. Use inventory, policy, review, triage, and remediation assistance with deterministic evidence and human release authority. The design should name authoritative inputs, identities, state transitions, dependencies, cryptographic or policy boundaries, limits, owners, telemetry, failure behavior, and acceptance criteria.

This guide traces the subject through the source, contributor, dependency, build, provenance, signature, inventory, distribution, deployment, and runtime chain. A configured integration, successful API response, established cryptographic session, green dashboard, or closed ticket is not sufficient proof. Intended state must be reconciled with applied state, negative tests, degraded behavior, raw evidence, recovery, and the resulting security or mission outcome.

A production implementation starts with a minimal known-good baseline and explicit invariants. Introduce one dependency or capability at a time, preserve before-state evidence, test authorized and prohibited behavior, inject realistic failures, measure convergence or recovery, and retain a tested rollback. Automation must stop safely when authority, freshness, identity, integrity, cryptographic state, or observed outcomes are ambiguous.

Troubleshooting locates the first divergence from the expected state machine or end-to-end transaction. Inspect every authority boundary and account for time, caching, eventual consistency, stale identity, incomplete collection, parser drift, asymmetric paths, encapsulation, supply-chain state, side channels, partial failure, and missing evidence. Closure requires causal explanation, reproducibility, validated restoration, residual limitations, and prevention or monitoring actions.

Engineering practices

Define authority, ownership, scope, dependencies, limits, and acceptance criteria for automation and ai assistance.

Model expected state, trust boundaries, data flow, identity, and failure behavior before implementation.

Validate intended configuration, active state, applied enforcement, raw evidence, and end-to-end outcome independently.

Test normal, prohibited, degraded, failed, scale, recovery, and rollback conditions.

Preserve policy, source, version, identity, telemetry, artifact, packet, log, time, and decision evidence.

Automate through typed contracts, least privilege, deterministic gates, canaries, bounded authority, and reversible changes.

Failure patterns

Treating automation and ai assistance as a product feature instead of an end-to-end engineering system.

Relying on intended policy, rule code, or controller state without proving applied and operational behavior.

Ignoring data quality, identity, time, cryptographic lifecycle, supplier state, or partial failure.

Changing several variables before preserving evidence and stating a falsifiable hypothesis.

Allowing automation or AI to act on stale, conflicting, unverified, or unauthorized inputs.

Declaring success after one positive result without negative tests, rollback, and recurrence monitoring.

Validation and evidence

Method	Expected evidence
Examine	Requirements, architecture, identities, data, algorithms, policy, configuration, ownership, and dependencies for automation and ai assistance.
Observe	Active state, applied policy, raw events, telemetry, artifacts, cryptographic results, resources, and endpoint behavior.
Test	Expected behavior, prohibited behavior, malformed inputs, dependency loss, partial failure, scale, recovery, and rollback.
Measure	Availability, latency, detection quality, evidence completeness, cryptographic cost, risk, resource use, and mission outcome.
Reconcile	Intended state against observed security, identity, application, cloud, cryptographic, data, and evidence records.

Chapter 08 laboratory and review

Practical laboratory
Run a compromised-builder exercise and produce an affected-artifact inventory, trust-reset plan, consumer communication, clean rebuild, and proof.

Laboratory evidence package

- Architecture or topology showing boundaries, identities, dependencies, and authoritative inputs.
- Versioned configuration or code with explanatory comments and reproducible setup instructions.
- Nominal-path validation plus at least two injected failure or boundary conditions.
- Structured logs, traces, metrics, packet captures, test output, or other appropriate evidence.
- A concise decision record covering findings, limitations, residual risks, and next actions.

Frontier extension

Develop an evidence-governed supply-chain supervisor that blocks unverifiable artifacts and coordinates specialized analysis agents.

Review gate
Advance only when another engineer can reproduce the result, explain the architecture, identify the failure boundary, and distinguish measured evidence from assumptions or projections.

Integrated learning roadmap

The guide is intended to be revisited. First establish fluency, then build a small implementation, operate it under failure, automate repeatable work, and finally evaluate emerging techniques against a stable baseline. Progress is demonstrated through evidence rather than reading completion alone.

Stage	Demonstrated capability
Foundation	Explain the system from first principles and trace its critical path without relying on product terminology.
Build	Implement the smallest representative system with explicit contracts, identities, telemetry, and tests.
Operate	Diagnose injected failures, recover safely, and preserve evidence of the causal chain.
Automate	Convert a proven manual workflow into bounded, idempotent, observable automation.
Architect	Design for scale, resilience, security, interoperability, and lifecycle change.
Explore	Evaluate frontier techniques using stated hypotheses, limitations, and adoption gates.

Source authority register

Source policy

Official standards, specifications, and project documentation remain with their issuing organizations. Mythos Systems provides original engineering interpretation, synthesis, implementation guidance, and relationship mapping. Source verification dates do not guarantee that an authority has not changed since review.

01. NIST - SP 800-218 - Secure Software Development Framework Version 1.1
Role: Final secure-software-development practices and tasks.
Verified: 2026-07-26
Official source: https://csrc.nist.gov/pubs/sp/800/218/final

02. NIST - Draft SP 800-218 Rev. 1 - Secure Software Development Framework Version 1.2
Role: Initial public draft of the next SSDF revision; used with explicit draft status.
Verified: 2026-07-26
Official source: https://csrc.nist.gov/pubs/sp/800/218/r1/ipd

03. NIST - SP 800-218A - Secure Software Development Practices for Generative AI and Dual-Use Foundation Models
Role: SSDF community profile for AI model and AI system development.
Verified: 2026-07-26
Official source: https://csrc.nist.gov/pubs/sp/800/218/a/final

04. NIST - SP 800-204D - Strategies for the Integration of Software Supply Chain Security in DevSecOps CI/CD Pipelines
Role: Software-supply-chain security integration for cloud-native DevSecOps pipelines.
Verified: 2026-07-26
Official source: https://csrc.nist.gov/pubs/sp/800/204/d/final

05. SLSA - Supply-chain Levels for Software Artifacts Specification
Role: Build provenance and software supply-chain integrity framework.
Verified: 2026-07-26
Official source: https://slsa.dev/spec/

06. OpenSSF - OpenSSF Scorecard
Role: Automated checks for open-source project security practices.
Verified: 2026-07-26
Official source: https://scorecard.dev/

07. Sigstore - Sigstore Documentation
Role: Keyless and identity-bound artifact signing, transparency, and verification.
Verified: 2026-07-26
Official source: https://docs.sigstore.dev/

08. Linux Foundation / SPDX - SPDX Specification
Role: Standard for software bill of materials, licensing, security, and supply-chain data.
Verified: 2026-07-26
Official source: https://spdx.github.io/spdx-spec/

09. OWASP Foundation - CycloneDX Specification
Role: Software, services, ML, and cryptographic bill-of-materials standard.
Verified: 2026-07-26
Official source: https://cyclonedx.org/specification/overview/

Authority application and refresh triggers

Authority class	Required library action
Protocols and specifications	Recheck interoperability, security considerations, defaults, and migration guidance when a referenced specification changes.
Security and assurance frameworks	Update control mappings, evidence expectations, risk language, and assessment procedures after material framework revision.
Languages, runtimes, and projects	Re-run examples, tests, deprecation checks, and operational recommendations against supported releases.
Benchmarks and evaluations	Re-execute claims when workloads, methods, models, compilers, drivers, hardware, or scoring rules change.
Operational evidence	Incorporate validated incidents, failure tests, reader corrections, and measured implementation outcomes into the next revision.

Limitations, freshness, and revision control

Boundary	Statement
Publication limitation	This guide synthesizes broad engineering practice. It does not replace product documentation, legal advice, safety certification, or environment-specific design review.
Evidence limitation	Not every pattern is benchmarked in every environment. Examples illustrate engineering methods and must be validated against actual versions, workloads, hardware, policy, and failure conditions.
Freshness limitation	Vendor behavior, software libraries, AI models, security threats, and emerging standards can change quickly. Volatile claims require source revalidation before operational adoption.
Adoption limitation	Frontier content is not an implicit adoption recommendation. Adoption requires defined value, qualification gates, controlled proof, rollback, ownership, and residual-risk acceptance.
Status boundary	Candidate Focused Field Guide describes the publication, not certification of a product, implementation, engineer, or organization.

Revision record

Version	Revision statement
1.0.0-candidate	Initial candidate living-guide edition. Published 2026-07-26; scheduled review 2026-10-26.

Search and relationship metadata

Dimension	Engineering position
Primary domain	MYTHOS-SEC - Cybersecurity, Security Engineering & Cryptography
Secondary domain	MYTHOS-SWE; MYTHOS-AI; MYTHOS-CLD
Publication class	Field Guide / Canonical Living Overview Guide
Skill levels	L1 Foundational; L2 Practitioner; L3 Advanced; L4 Frontier
Tags	software supply chain, AI supply chain, SLSA, provenance, Sigstore, SBOM, VEX, CI/CD, model security, artifact signing
Canonical route	/library/field-guides/mythos-fg-sec-0014/

Living publication

Corrections, expanded laboratories, improved diagrams, authority changes, and new engineering evidence should revise this permanent publication rather than create disconnected replacements.