

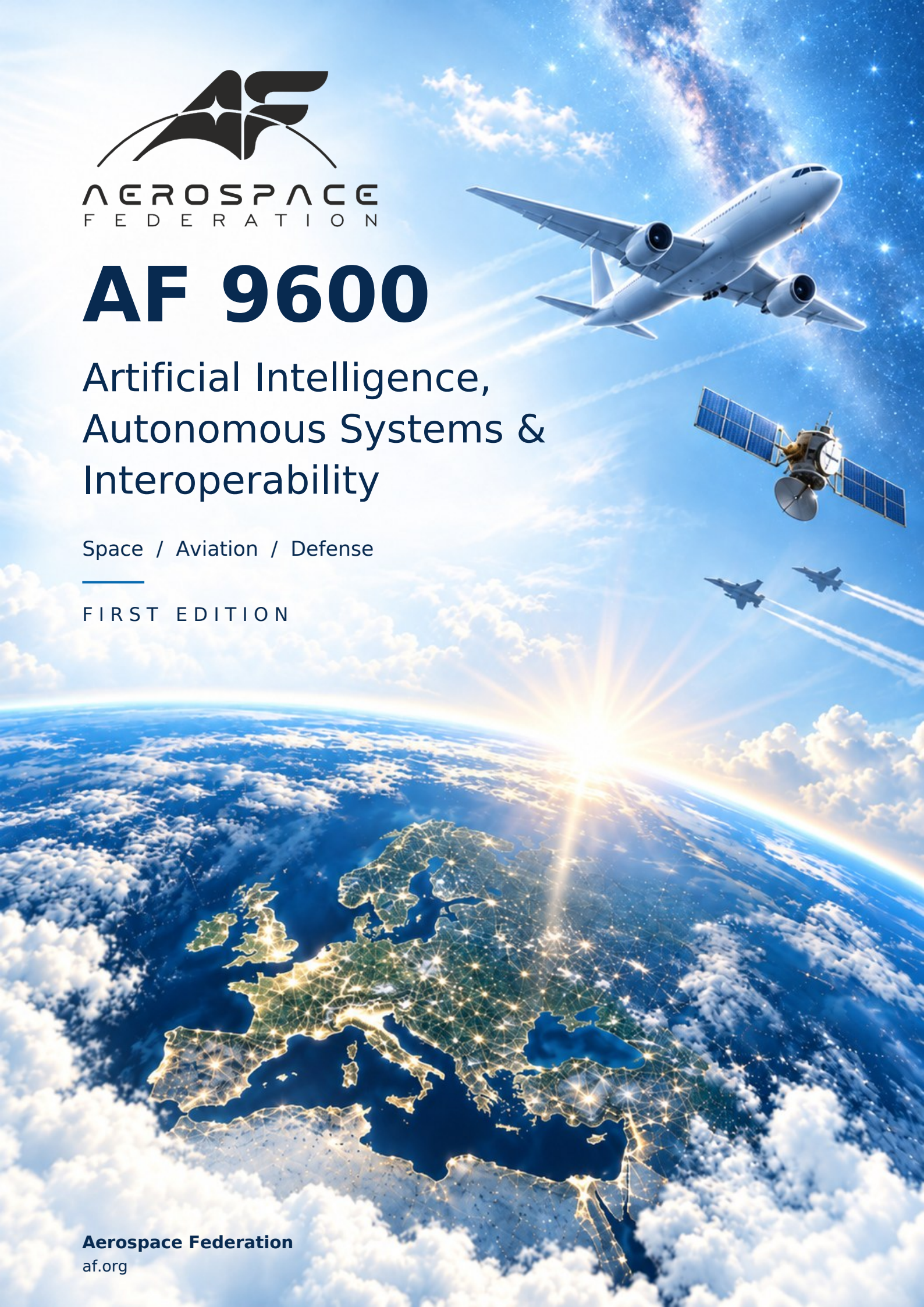


AF 9600

Artificial Intelligence, Autonomous Systems & Interoperability

Space / Aviation / Defense

FIRST EDITION



Document identity and authority

FIRST EDITION MANUSCRIPT - Prepared for Aerospace Federation review

Field	Declared position
Identifier	AF 9600
Title	Artificial Intelligence, Autonomous Systems & Interoperability
Organization	Aerospace Federation
Edition / language	FIRST EDITION / English master text
Content	Clauses 0-40 and Annexes A-T
Normative basis	Main clauses and Annexes A-N contain proposed requirements; Annexes O-T are informative.
Publication status	Review manuscript. Publication approval and effective date are not assigned.

Use of this publication

AF 9600 connects artificial intelligence, bounded autonomy and interoperability to the complete aerospace function. It addresses data and model validity, actual action authority, human oversight, shared meaning, integrated behavior and continuing evidence. Product-specific values, operating conditions, permissions and authorities are completed for the actual application.

The supporting set provides original implementation models, technical profiles, examination checklists and blank controlled records. It does not supply fabricated results, signatures, approvals or recognition.

Informative calculations and cases use synthetic values. They explain assumptions, reproducible methods and decision boundaries; they are not mandatory defaults or evidence from actual systems.

Reader route and decision boundaries

1 Define purpose, domain and authority

Use Clauses 0-10 and Annexes A-C, F and N for applicability, responsibilities, operational domain, risk, architecture and lifecycle gates.

2 Establish trustworthy data and models

Use Clauses 11-18 and Annexes D-E for provenance, labels, partitions, learning, evaluation, robustness, adaptation and generative workflows.

3 Demonstrate bounded autonomous behavior

Use Clauses 19-22 and Annexes F-G for perception, planning, collective behavior, human oversight and intervention.

4 Verify actual interoperability

Use Clauses 23-29 and Annexes H-J for contracts, semantic meaning, distributed state, trust, supply, simulation and integrated evaluation.

5 Apply the aerospace operating context

Use Clauses 30-32 and Annex K for aviation, space, remote missions, industrial robotics, maintenance and ground services.

6 Deploy, maintain and preserve the claim

Use Clauses 33-40 and Annexes L-N for readiness, updates, monitoring, incidents, release, changes, retirement and adoption.

Decision boundary: Model performance, autonomous-system validation, interoperability acceptance, risk acceptance, release and operating permission are distinct. An unresolved essential condition cannot be compensated for by an aggregate score.

Contents - clauses and annexes

Entries and PDF bookmarks locate the controlled clauses and annexes. Page references apply to this file.

0 Introduction and intended outcomes	8
1 Scope and lifecycle coverage	9
2 Applicability, precedence and local parameters	9
3 Terms and interpretation	10
4 Governance, accountability and responsible use	11
5 Concept of operations and authority allocation	12
6 Operational domain and boundary management	12
7 Integrated risk and consequence assessment	13
8 Requirements, allocation and traceability	14
9 Architecture, containment and runtime assurance	14
10 Lifecycle planning and assurance gates	15
11 Data acquisition, provenance and permitted use	16
12 Annotation, reference truth and data quality	16
13 Dataset partitioning and representation	17
14 Model development and learning assurance	17
15 Evaluation, statistics and calibrated claims	18
16 Robustness, uncertainty and unsupported inputs	19
17 Adaptive models and online learning	19
18 Generative systems and agentic workflows	20
19 Perception, fusion and state estimation	21
20 Planning, decision making and control	21

21 Multi-agent coordination and collective behavior	22
22 Human oversight, interfaces and competence	23
23 Interoperability governance and contracts	23
24 Semantic integrity, coordinates, units and time	24
25 Protocol behavior, service quality and distributed state	25
26 Cross-domain trust and cybersecurity	25
27 Suppliers, tools, builds and configuration identity	26
28 Simulation, digital twins and formal evidence	27
29 Integrated verification and operational validation	27
30 Aviation and air-ground integration	28
31 Space, remote missions and intermittent links	28
32 Industrial robotics, ground services and maintenance	29
33 Deployment, updates and readiness	30
34 Monitoring, drift and operational feedback	30
35 Incidents, fallback and verified recovery	31
36 Assurance case, release and truthful claims	32
37 Changes, legacy systems and continuing support	32
38 Retirement, transfer and evidence preservation	33
39 Assurance reviews, competence and records	34
40 Publication, local adoption and maintenance	34
Annex A - Activity-based applicability and retained duties	36
Annex B - Mandatory parameter baseline and common record payload	37
Annex C - Operational concept, domain and authority dossier	40

Annex D - Data provenance, quality and partition dossier	41
Annex E - Model development and evaluation dossier	42
Annex F - Autonomy architecture and control dossier	43
Annex G - Human oversight and competence dossier	44
Annex H - Interoperability and semantic contract dossier	45
Annex I - Trust, suppliers, tools and configuration dossier	46
Annex J - Simulation, verification and validation dossier	47
Annex K - Aerospace application dossier	48
Annex L - Deployment, monitoring and recovery dossier	49
Annex M - Release, change, continued use and retirement dossier	50
Annex N - Assurance, publication and adoption dossier	51
Annex O - Reproducible calculations and decision examples	52
Annex P - Integrated application cases	56
Annex Q - Assurance and interoperability diagrams	58
Annex R - Activity-based implementation routes	63
Annex S - Interpretation questions and boundary examples	64
Annex T - Requirement families and implementation crosswalk	65

Tables - requirements and worked evidence

Table 0-1. Distinct assurance decisions	8
Table 3-1. Terms used by AF 9600	10
Table A-1. Activity routes and responsibility boundaries	36
Table B-1. P01-P44 parameter schedule	37
Table C-1. Required scenario payload	40
Table E-1. Evaluation result minimum content	42
Table H-1. Interface contract checks	45
Table M-1. Bounded release statement	50
Table P-1. Example populated decision record for Case P.5	57
Table T-1. Core navigation map	65

0 Introduction and intended outcomes

0.1 Assurance of the integrated function

AF 9600 establishes requirements for artificial intelligence, autonomous systems and interoperability in aerospace activities. The assurance object is the complete intended function: people, data, models, software, hardware, connected participants and operating conditions. A capable component does not by itself establish a safe, useful or authorized integrated system. The organization shall connect intended outcomes and prohibited behavior to actual implementation, evidence, deployment and continuing operation. Claims shall identify the exact function, configuration, population, environment and authority to which they apply.

0.2 Separate decision boundaries

Model evaluation, system validation, interoperability acceptance, residual-risk acceptance, release and permission to operate shall remain separate decisions. Each shall have an eligible owner, declared inputs, acceptance rules, retained evidence and bounded conclusion. A successful demonstration shall not silently authorize wider autonomy or a new operational domain.

Table 0-1. Distinct assurance decisions

Decision	Required basis	Boundary
Model evaluation	Identified data, model, metric and uncertainty.	Does not establish complete system safety.
Autonomy validation	Required behavior, authority and real operating constraints.	Does not authorize every mission or mode.
Interoperability acceptance	Exact peers, contracts, configurations and scenarios.	Does not establish universal compatibility.
Operational authorization	Valid evidence, readiness, accountable authority and restrictions.	Does not override unmet mandatory obligations.

0.3 Evidence and uncertainty

The organization shall preserve adverse, inconclusive and invalid results. PASS, FAIL, HOLD and approved N/A shall remain distinguishable from task completion. Uncertainty shall be stated in terms relevant to the decision. High average accuracy, confidence scores or aggregate compliance percentages shall not compensate for an unresolved essential condition.

0.4 Language and publication status

Shall identifies a proposed requirement; should identifies guidance; may identifies permission subject to stated conditions. Clauses 0-40 and Annexes A-N form the proposed normative text. Annexes O-T are informative. This FIRST EDITION manuscript is prepared for Aerospace Federation review; publisher approval and effective date are not assigned. Publication, local adoption, technical acceptance and operating permission are distinct acts.

Required evidence: Identified assurance object, controlled edition, decision states, evidence relationships and eligible authorities.

Decision control: This manuscript shall not be represented as an approved publication or verified implementation without the corresponding authorized decision records.

1 Scope and lifecycle coverage

1.1 Covered systems

This standard applies to systems using learned, knowledge-based, optimization-based, generative or hybrid decision functions, and to autonomous functions whether or not they use machine learning. It also applies to interoperability among such systems and their human, operational and technical partners. Activities include specification, acquisition, development, integration, evaluation, deployment, operation, support and retirement.

1.2 Aerospace settings

The declared scope shall identify aviation, space, industrial, maintenance, ground-service, research or defense-support contexts actually covered. Aircraft and ground systems, mission segments, robotic work cells, planning services and enterprise decision aids may have different consequences and authorities. A shared product name shall not conceal these differences.

1.3 Roles and retained duties

The scope shall identify providers of data, models, components, platforms, tools and services; system integrators; operators; assessors; and decision authorities. Outsourcing shall not remove responsibility for an organization's own claims, integration decisions or operational duties. Suppliers shall declare evidence-access and support limitations before acceptance.

1.4 Boundaries of use

A function without direct actuation may still influence consequential human or machine decisions and shall be assessed accordingly. An isolated model benchmark is not an operational safety demonstration. The organization shall identify applicable permissions, contractual duties and legal obligations without treating this standard as a substitute for them.

Required evidence: Product and activity scope, role map, operating context, interfaces and retained responsibilities.

Decision control: Excluding a component or supplier from assessment shall not exclude its effects on an in-scope function.

2 Applicability, precedence and local parameters

2.1 Activity selection

The organization shall map actual activities and responsibilities to applicable clauses using Annex A. Non-applicability shall identify the exact condition, supporting rationale and approving authority. A clause is not inapplicable merely because evidence is unavailable or an activity is outsourced.

2.2 Requirement precedence

Applicable mandatory obligations, operating permissions and binding agreements shall be identified and reconciled with the proposed requirements. Conflicts shall be escalated before the affected decision. Local adoption shall state the controlled edition, selected scope and any additional programme requirements.

2.3 Parameter baseline

Annex B defines P01-P44. Each applicable parameter shall contain an approved value, rule or controlled specification reference before a dependent decision. The record shall include units, source, rationale, configuration, authority, evidence and change triggers. Unknown values shall remain visible as blocked dependencies rather than assumed defaults.

2.4 Tailoring and equivalence

Tailoring shall preserve each applicable outcome and necessary evidence. An alternative method shall demonstrate equivalent or stronger support for the specific claim under its actual assumptions. Reduced testing justified by architectural protection shall depend on verified protection, including independence, timing and failure behavior.

Required evidence: Applicability matrix, resolved precedence, complete parameter baseline and justified alternative methods.

Decision control: Convenience, cost or schedule shall not convert an unmet requirement into approved N/A.

3 Terms and interpretation

3.1 Core terms

Table 3-1. Terms used by AF 9600

Term	Meaning within this standard
AI function	A function that uses learned, knowledge-based or related computational inference to produce outputs.
Autonomous function	A function permitted to select or execute actions within a defined authority envelope.
Operational domain	The conditions, populations, modes and constraints for which a function is intended and supported.
Authority envelope	Permitted actors, actions, targets, conditions, duration and restrictions.
Interoperability contract	Agreed syntax, meaning, behavior, timing, trust and organizational responsibilities between participants.
Fallback	A defined alternative behavior intended to control consequences when nominal use is unavailable or unsupported.
Runtime assurance	Monitoring and intervention that enforce specified properties during operation under explicit assumptions.
Model baseline	Identified architecture, parameters, transformations, dependencies and execution configuration.

3.2 Confidence and evidence

Model confidence is an output defined by a model or calibration process. It shall not be interpreted as probability of system safety without a justified model and validated relationship. Statistical confidence, measurement uncertainty, epistemic uncertainty and operational risk shall be distinguished.

3.3 States and authority

PASS means that applicable acceptance conditions were demonstrated by valid evidence for the stated scope. FAIL means a condition was not met. HOLD means the conclusion cannot yet be determined. N/A requires approved non-applicability. An acceptance of residual exposure shall not rewrite a failed result or expand a person's delegated authority.

3.4 Compatibility and autonomy claims

Terms such as compatible, autonomous, intelligent, human-supervised or explainable shall be accompanied by measurable scope. Autonomy shall be described by permitted functions, modes and responsibilities rather than an unqualified level number. Compatibility is a relationship between exact participants under stated conditions, not an intrinsic universal property of one endpoint.

Required evidence: Controlled vocabulary, state definitions, claim scope and decision interpretation rules.

Decision control: Marketing language shall not replace a testable operational definition.

4 Governance, accountability and responsible use

4.1 Accountable ownership

The organization shall appoint accountable owners for system purpose, data, models, autonomy authority, interfaces, safety, security, human factors and operational decisions. Delegations shall identify scope, competence, independence and limits. Providers and operators shall agree who responds when assumptions or evidence become invalid.

4.2 Permitted and prohibited uses

The organization shall define intended and prohibited uses, affected people and assets, foreseeable misuse and consequential dependencies. Applicable rights, access restrictions and duties shall be evaluated for the actual context. A model output shall not independently create permission for an action that otherwise requires authorization.

4.3 Independent challenge

Assurance depth and independent review shall reflect consequence, novelty, complexity and the ability to detect errors before harm. Authors shall not approve their own critical conclusions where separate challenge is required. Reviewers shall have access to limitations, rejected candidates and adverse evidence, including supplier-held information needed for the claim.

4.4 Escalation and reporting

Personnel shall be able to report unsafe behavior, unsupported claims, data problems or improper pressure without suppression. Reports shall preserve facts and uncertainty. The organization shall define stop, restrict, escalate and resume authorities and provide resources to execute them.

Required evidence: Governance and delegation records, use restrictions, conflict controls, reporting and escalation records.

Decision control: Commercial urgency, a senior person's preference or a model recommendation shall not override required technical authority.

5 Concept of operations and authority allocation

5.1 Operational concept

The organization shall describe users, mission or service objectives, physical environment, workflow, participants, modes and foreseeable abnormal conditions. The concept shall show how recommendations become decisions and how decisions become physical or informational effects. Indirect influence and downstream reliance shall be included.

5.2 Authority envelope

For every autonomous or assisted function, the organization shall define permitted actions, targets, priorities, preconditions, duration, constraints and prohibited transitions. Initiation, modification, override, cancellation and termination rights shall be assigned. Authority shall not be inferred solely from possession of a credential or receipt of a message.

5.3 Modes and handovers

Mode transitions shall identify guards, entry and exit conditions, transferred responsibilities, acknowledgments and timeout behavior. The actual active mode and responsible controller shall be observable. Conflicting commands, simultaneous requests and lost acknowledgments shall have defined resolutions that preserve essential constraints.

5.4 Operational feasibility

The concept shall demonstrate that staffing, workload, communications, intervention time, energy, sensing and support can sustain the assigned roles. A remote human shall not be credited with intervention that cannot complete within the available time or link conditions. Unsupported dependencies shall restrict the permitted operation.

Required evidence: Operational concept, authority matrix, mode-transition model, handover and feasibility evidence.

Decision control: Expanded capability or connectivity shall not automatically expand authorized autonomy.

6 Operational domain and boundary management

6.1 Domain definition

The operational domain shall specify relevant physical, environmental, temporal, geographic, traffic, human, data and infrastructure conditions. It shall include combinations and transition regions that affect behavior, not only independent ranges. Permitted populations, sensor states, communication quality and unavailable services shall be explicit.

6.2 Observable membership

The organization shall identify how domain membership is determined, including measurable variables, uncertainty, coverage gaps and detection latency. Where a condition cannot be observed directly, a justified proxy or operational restriction shall be used. Missing, conflicting or stale observations shall not be treated as proof of domain membership.

6.3 Boundary approach and exit

The design shall define prediction, warnings, protective margins, fallback and recovery for approaching or leaving supported conditions. Intervention shall start early enough to complete before the relevant constraint is lost, with latency and uncertainty included. Re-entry shall require verified conditions and authorized transition criteria.

6.4 Domain change

A new environment, population, operating mode or dependency shall trigger reassessment of data, models, interfaces, human performance and prior evidence. Temporary local restrictions shall be recorded in deployment and operator instructions. Historical success in a narrower domain shall not establish validity in an expanded one.

Required evidence: Domain specification, membership monitoring, transition margins, fallback and domain-change assessment.

Decision control: Unknown domain membership shall trigger the predefined restricted or protective behavior, not an assumed nominal state.

7 Integrated risk and consequence assessment

7.1 Scenario construction

The assessment shall cover accidental faults, model limitations, intentional misuse, human error, environmental change and interaction among participants. Scenarios shall connect initiating conditions, mechanisms, propagation, detection and consequences. Common dependencies and repeated or correlated failures shall be included.

7.2 Coupled controls

Proposed safety, security, privacy or operational controls shall be examined for adverse effects on other objectives. Blocking a connection, stopping a process or removing a human interface can introduce physical or mission hazards. Control combinations shall be assessed for conflict, resource contention and inconsistent authority.

7.3 Treatment and residual exposure

Treatment shall identify required properties, implementation, independent protection, verification and operating restrictions. Residual exposure shall be accepted only by an eligible authority for a defined population, configuration and duration or review trigger. Numerical scores may prioritize work but shall not replace scenario-specific reasoning.

7.4 Reassessment and withdrawal

Changes, incidents, monitoring signals, invalid assumptions and newly credible misuse shall reopen affected scenarios and claims. The organization shall trace impacts to deployed systems and partner decisions. Withdrawal or restriction shall be communicated through accountable operational routes.

Required evidence: Linked scenarios, control arguments, verification, residual-exposure decisions and reassessment history.

Decision control: A favorable average risk score shall not neutralize an unresolved essential scenario.

8 Requirements, allocation and traceability

8.1 Required and prohibited behavior

Requirements shall define intended outputs, unacceptable actions, performance, timing, domain limits, data quality, authority and fallback. Acceptance criteria shall state the population, metric, method, threshold and uncertainty basis before evaluation. Requirements that cannot be measured directly shall have a justified verification strategy.

8.2 Allocation and integration

Requirements shall be allocated across models, conventional software, hardware, people, external services and runtime controls. Interfaces and assumptions between allocations shall be explicit. An objective assigned to a human or external partner shall have demonstrated capability, agreement and continuing availability.

8.3 Bidirectional trace

Traceability shall connect authoritative needs to scenarios, domain, architecture, data, models, interface contracts, evidence and deployed configurations. The trace shall identify unimplemented, unverified, failed and excluded items separately. A populated trace matrix shall not itself establish conformity.

8.4 Change impact

Changes shall identify affected requirements, assumptions, datasets, model candidates, contracts, protective controls, human procedures and releases. Regression scope shall reflect indirect and emergent effects. An unchanged model file does not establish an unchanged function when preprocessing, interfaces or operating conditions change.

Required evidence: Reviewed requirements, allocation, complete trace and change-impact records.

Decision control: An unresolved essential requirement shall block the dependent release or operating claim.

9 Architecture, containment and runtime assurance

9.1 Architectural basis

The architecture shall show control and information paths, trust boundaries, shared resources, protective elements and authority arbitration. Consequence-bearing outputs shall have defined acceptance and containment controls. Diversity and redundancy claims shall identify actual independence and common-cause limitations.

9.2 Runtime protection

Each credited monitor shall have a defined property, observation source, detection capability, timing bound, failure behavior and intervention authority. The protected system shall not be able to disable or corrupt its protection through an unassessed shared path. Monitor false alarms, missed detections and resource exhaustion shall be evaluated.

9.3 Fallback validity

Fallback shall be suitable for the actual state and environment, reachable within available time and sustainable for its required duration. Transition dynamics, actuator authority, data freshness, energy and human roles shall be verified. A label such as safe state shall not substitute for evidence that the state controls the relevant consequences.

9.4 Integrated assurance

The organization shall demonstrate the relationship between monitor assumptions, detection timing, intervention and remaining margins. Protection shall not be credited outside its verified envelope. Formal properties or simulation results shall retain their assumptions and connection to the deployed implementation.

Required evidence: Architecture, independence analysis, monitor and fallback specifications, timing and transition evidence.

Decision control: Adding a monitor shall not automatically justify unrestricted use of an unverified controller.

10 Lifecycle planning and assurance gates

10.1 Integrated plan

The organization shall plan requirements, data, model development, integration, independent evaluation, operational validation, deployment and continuing assurance as a connected lifecycle. Plans shall identify work products, competent roles, dependencies, gates and evidence retention. Iterative development shall preserve the baseline and decision history of each increment.

10.2 Gate decisions

Entry and exit conditions shall be defined for data acceptance, model selection, integration, operational trials, release and changes. Gate reviews shall distinguish available evidence from planned work. Unresolved items shall retain owners, restrictions and an explicit effect on the next activity.

10.3 Independence and resources

The plan shall allocate independent evaluation resources and protect held-out evidence from inappropriate development access. Required expertise shall include the operational context and relevant human, statistical, safety, security and interface disciplines. Training attendance alone shall not establish competence for an assigned decision.

10.4 Evidence continuity

The organization shall preserve data and model lineage, candidate-selection history, tools, tests, findings and actual deployment state. Discontinued candidates and unfavorable results shall remain available where needed to interpret selection bias or technical decisions. Lifecycle changes shall reopen affected gates.

Required evidence: Lifecycle and independence plans, gate criteria, actual reviews, resources and retained histories.

Decision control: A planned milestone or demonstration event shall not establish gate completion.

11 Data acquisition, provenance and permitted use

11.1 Data requirements

Data requirements shall derive from intended functions, operational domain, consequences and evaluation needs. The organization shall identify required populations, modalities, conditions, resolutions, temporal relationships and rare events. Availability shall not be the sole basis for choosing training or evaluation populations.

11.2 Origin and lineage

Every controlled dataset shall identify origin, collection method, equipment or service, dates, conditions, transformations and responsible owner. The record shall preserve relationships to source items and versions. Unknown provenance and unverified third-party assertions shall be visible limitations.

11.3 Use restrictions and protection

The organization shall confirm permitted collection, processing, transfer, retention and reuse for the actual purpose. Sensitive or restricted information shall be protected in development, evaluation and operation. Data minimization or removal shall be reconciled with the evidence needed to support safety and historical decisions through an approved disposition.

11.4 Collection validity

Collection shall control synchronization, calibration, missing values, sampling effects, corruption and selective recording. Operational feedback collected only after a model's own selection shall be assessed for feedback bias. Synthetic or augmented data shall remain distinguishable from observations and have a justified role.

Required evidence: Data specification, source inventory, permitted-use basis, lineage and collection-quality records.

Decision control: Possession of a dataset shall not establish permission to use it or representativeness for the intended domain.

12 Annotation, reference truth and data quality

12.1 Annotation rules

Annotation instructions shall define labels, units, classes, ambiguity, exclusion rules and treatment of uncertain cases. Annotators shall be competent for the context and use controlled tools. Changes to definitions shall identify affected items and dependent models or results.

12.2 Reference construction

Reference values shall be established by a justified method, with independence from the candidate under evaluation where necessary. Human agreement is not automatically physical truth. Sensor uncertainty, expert disagreement, delayed outcomes and incomplete observability shall be retained rather than forced into apparently certain labels.

12.3 Quality assessment

The organization shall examine completeness, correctness, consistency, duplication, distribution, temporal integrity and label reliability using methods appropriate to each data population. Sampling shall state the sampled unit and selection basis. Adjudication shall preserve original labels, reasoning and resulting changes.

12.4 Corrections and effects

Corrections shall be versioned and traced to training, tuning, evaluation and operational decisions. The organization shall determine whether a correction changes a conclusion, threshold or claimed domain. Removal of difficult items shall have a technical rationale and shall not conceal relevant failures.

Required evidence: Annotation specification, competence, reference method, disagreement and correction history.

Decision control: Agreement rates or a clean data file shall not alone establish valid ground truth.

13 Dataset partitioning and representation

13.1 Independent partitions

Training, tuning, calibration and final evaluation roles shall be defined before candidate evaluation. Separation shall address common subjects, missions, equipment, locations, time periods and transformed duplicates. Random record-level splitting shall not be used to claim independence when records share a consequential source.

13.2 Representative coverage

The organization shall assess coverage of intended populations, operating modes, boundary combinations and rare consequential conditions. Missing or weakly represented regions shall have explicit restrictions, further evidence or protective controls. Population weighting shall be justified and reported alongside important subgroup results.

13.3 Leakage and contamination

The organization shall assess leakage through preprocessing, feature selection, labels, synthetic generation, retrieval sources, pretraining and repeated test access. A final test reused for design choices shall be treated as development evidence unless an independently justified evaluation basis remains. Provider opacity shall be recorded as uncertainty.

13.4 Controlled evaluation access

Final evaluation sets and selection rules shall be protected against unauthorized alteration and selective reporting. Test identities, candidate counts and access history shall be retained. When sets are refreshed, comparability and historical interpretation shall be addressed.

Required evidence: Partition plan, source-group mapping, representation analysis, leakage checks and access records.

Decision control: Large record counts shall not be presented as independent evidence without a justified sampling unit.

14 Model development and learning assurance

14.1 Learning problem and alternatives

The organization shall define inputs, outputs, objectives, constraints, permitted approximations and reasons for selecting the model approach. Simpler or non-learned alternatives shall be considered where relevant to controllability and evidence. Optimization objectives shall not omit essential constraints or reward prohibited behavior.

14.2 Training baseline

Training records shall identify dataset versions, preprocessing, architecture, initialization, random seeds where applicable, hyperparameters, tools, hardware, dependencies and stopping rules. Nondeterminism shall be characterized sufficiently to support the decision. Training failures and selection among candidates shall remain traceable.

14.3 Selection and implementation

Candidate selection shall use predefined criteria and preserve tradeoffs among overall performance, critical subgroups, resource use and robustness. Quantization, pruning, compilation, conversion and platform changes shall be evaluated on the resulting artifact. The deployed transformation chain shall match the verified baseline.

14.4 Model limitations

The organization shall document known failure modes, uncertainty, unsupported populations, explanation limits and dependency assumptions. Model cards or equivalent summaries shall link to underlying evidence and actual release identities. A summary shall not conceal conditions under which performance deteriorates.

Required evidence: Learning specification, reproducible training record, candidate-selection history and deployed-model assessment.

Decision control: A higher benchmark score shall not override an unmet essential constraint or unassessed implementation change.

15 Evaluation, statistics and calibrated claims

15.1 Evaluation design

Evaluation shall identify the decision being supported, population, sampling unit, reference method, metrics, acceptance rules and uncertainty treatment. Sample size and sampling strategy shall be justified for consequential cases. Clustered observations, temporal dependence, distribution changes and repeated candidate selection shall be considered.

15.2 Metrics and populations

The organization shall report relevant error types, prevalence, coverage, abstention and critical subgroup performance. Aggregate accuracy shall be accompanied by information that reveals consequential minority failures. Conditional performance after abstention shall be distinguished from performance over all submitted cases.

15.3 Confidence and calibration

Statistical intervals shall state method, confidence level, assumptions and effective sample basis. Calibration claims shall identify the event, population, bins or other method and validation data. A calibrated probability for one event shall not be relabeled as confidence in an unrelated system property.

15.4 Comparisons and conclusions

Comparisons shall use an appropriate paired or independent design, equivalent conditions and declared practical acceptance criteria. Multiple comparisons and optional stopping shall be controlled or explicitly limited. All required outcomes, including failures and invalid runs, shall be reported with the exact candidate baseline.

Required evidence: Evaluation protocol, complete results, uncertainty analysis, subgroup findings and bounded conclusions.

Decision control: Statistical significance, zero observed failures or an average score shall not establish universal safety or operational permission.

16 Robustness, uncertainty and unsupported inputs

16.1 Stress and boundary conditions

The organization shall evaluate relevant perturbations, degradation, missing modalities, environmental extremes, sensor faults, unusual combinations and maliciously influenced inputs within authorized test boundaries. Perturbation models shall correspond to credible conditions; arbitrary digital noise shall not substitute for physical or operational analysis.

16.2 Unsupported-input detection

Detection of out-of-domain, unfamiliar or low-quality inputs shall have a defined purpose, tested sensitivity and false-alarm behavior. The organization shall evaluate conditions that can defeat detection, including confidently incorrect outputs. Failure to raise a warning shall not be treated as proof of valid domain membership.

16.3 Uncertainty and abstention

Uncertainty mechanisms shall identify the kind of uncertainty represented and how it affects decisions. Abstention, rejection or degraded behavior shall be integrated into the operational workflow, including human workload and capacity. The fallback demand created by abstention shall be included in feasibility analysis.

16.4 Residual limitations

The organization shall retain failures, blind spots and conditions not examined. Robustness claims shall be restricted to demonstrated perturbations and contexts. New observations shall trigger review of training, monitoring, thresholds and deployment restrictions.

Required evidence: Robustness scenarios, detection validation, uncertainty interpretation, abstention and residual-limit records.

Decision control: A confidence threshold or novelty detector shall not be credited as an infallible safety barrier.

17 Adaptive models and online learning

17.1 Adaptation boundary

The organization shall define what may adapt, when, from which data, under whose authority and within which invariant constraints. Parameters, policy rules, retrieval content and learned state shall be distinguished. Adaptation shall not expand the operational domain or action authority without a separate approved decision.

17.2 Candidate and promotion controls

Learning from operational data shall be isolated from active decision authority unless the complete adaptive mechanism is justified and verified. Candidate promotion shall require identified evidence, protected evaluation and an eligible decision. Poisoned feedback, drift, feedback loops and catastrophic loss of prior capability shall be assessed.

17.3 Monitoring and reversal

The design shall identify actual active state, adaptation history, limits, triggers and protective actions. Reversal shall consider persistent memory, downstream actions and external state, not only model weights. A previous model may no longer be compatible with changed data or interfaces.

17.4 Continuing evidence

The organization shall demonstrate how evidence remains valid across the permitted adaptation envelope. Unanticipated behavior or violated assumptions shall suspend affected adaptation and trigger review. Unbounded self-modification shall not be treated as equivalent to a fixed verified baseline.

Required evidence: Adaptation specification, invariants, promotion and reversal records, actual-state history and continuing assurance.

Decision control: Improved recent performance shall not independently authorize promotion or continued learning in a consequential function.

18 Generative systems and agentic workflows

18.1 Intended use and output checks

The organization shall define permitted generated content, decisions and tool actions, including consequences of fabricated, incomplete or misleading output. Output acceptance shall use evidence appropriate to the task. Fluent explanations and citations produced by a model shall not be accepted as independent verification.

18.2 Context and information boundaries

Instructions, retrieved material, messages and tool responses shall have explicit trust and authority boundaries. Untrusted content shall not change the system's permissions or governing constraints. Retrieval quality, source freshness, sensitive-data handling and context loss shall be assessed in the actual workflow.

18.3 Bounded actions

Agents shall have least necessary access, explicit target and transaction limits, protected approval points and accountable action records. Repeated attempts, retries and chained tools shall not bypass a limit or duplicate a consequential action. Human approval shall be tied to the exact action and current context.

18.4 Evaluation and failure handling

Evaluation shall cover multi-step tasks, partial completion, unsupported assertions, instruction conflicts, unavailable tools and interrupted transactions. Independent tests shall assess both outputs and external effects. Recovery shall reconcile created records, sent commands, reservations and persistent memory before resuming.

Required evidence: Use-case and permission specification, context controls, action traces, independent evaluation and recovery cases.

Decision control: Generated content, urgency or an agent's own judgment shall not create authorization for a consequential action.

19 Perception, fusion and state estimation

19.1 Measurement and transformation chain

The organization shall identify sensor characteristics, calibration, preprocessing, synchronization, coordinate transformations, inference and fusion. Error and uncertainty shall be propagated or conservatively bounded through the chain. Missing, stale, saturated or inconsistent measurements shall remain distinguishable from valid observations.

19.2 Fusion assumptions

Fusion shall account for shared sources, correlated errors, latency and varying observability. Repeated information received through different paths shall not be counted as independent evidence without justification. Track identity, association ambiguity and discontinuities shall have controlled handling.

19.3 Operational performance

Evaluation shall cover detection, localization, classification, tracking and state-estimation behavior relevant to the intended function. Tests shall include occlusion, changes in conditions, ambiguous scenes and sensor disagreement. Performance shall be linked to downstream planning and control requirements.

19.4 Integrity and degradation

The design shall communicate validity, age, uncertainty and limitations with estimated state. Consumers shall respond correctly to unavailable or degraded estimates. Loss of observability shall invoke an approved restriction or fallback before downstream decisions become unsupported.

Required evidence: Measurement chain, fusion assumptions, uncertainty model, representative evaluation and degradation behavior.

Decision control: A precise numerical estimate shall not be treated as accurate, fresh or independent merely because it has many digits.

20 Planning, decision making and control

20.1 Goals and constraints

The organization shall specify objectives, priorities, hard constraints, permitted tradeoffs and termination conditions. Goal conflicts shall have an accountable resolution. Optimizers and learned policies shall not trade away a mandatory constraint to improve an aggregate reward or score.

20.2 Feasibility and timing

Plans shall account for actual state uncertainty, dynamics, actuator limits, energy, resources and end-to-end latency. Planning time, stale plans and failure to find a feasible solution shall have defined responses. The system shall verify that a selected plan remains eligible at execution time.

20.3 Closed-loop behavior

Verification shall examine stability, constraint preservation, oscillation, deadlock, mode transitions and recovery under relevant disturbances. Component performance shall be evaluated in the complete feedback loop. A safe open-loop sequence shall not establish safe closed-loop behavior with uncertain perception and delayed execution.

20.4 Action accountability

Executed actions shall be attributable to the active controller, plan, authority and observed state. Cancellation and override shall have defined limits, acknowledgment and residual effects. A command reported as canceled shall not conceal an irreversible or already completed action.

Required evidence: Goal and constraint model, feasibility and timing analysis, closed-loop results and action records.

Decision control: An optimal or high-reward plan shall not be executed outside its authority or verified constraints.

21 Multi-agent coordination and collective behavior

21.1 Participant and authority model

The system shall identify participant roles, membership rules, capabilities, trust and authority allocation. Joining, leaving, replacement and temporary loss shall have controlled effects. One participant shall not acquire new authority merely because another becomes unavailable.

21.2 Coordination properties

Required properties shall include relevant mutual exclusion, conflict resolution, consistency, liveness, resource ownership and termination. Agreements shall state what happens under delay, partition, conflicting membership views and partial failure. Assumed synchrony and failure models shall be justified for the actual environment.

21.3 Emergent and common-cause effects

Evaluation shall examine feedback among agents, correlated decisions, shared model errors, competition, collective oscillation and resource exhaustion. Individual-agent acceptance shall not establish collective acceptance. The tested population size, interaction density and communication conditions shall bound the claim.

21.4 Degradation and reconciliation

The design shall define partition behavior, local authority limits, protective separation of duties and reconciliation after reconnection. Duplicate or conflicting actions shall be detected and resolved with an authoritative state history. Restart shall not erase outstanding commitments or transfer ownership ambiguously.

Required evidence: Membership and authority model, collective properties, integrated scenarios and reconciliation records.

Decision control: Peer agreement or consensus shall not establish the truth, safety or authorization of the agreed action.

22 Human oversight, interfaces and competence

22.1 Human role design

Human roles shall identify information available, decisions permitted, workload, intervention time and necessary competence. Oversight shall be operationally feasible under abnormal and peak-demand conditions. A nominal human-in-the-loop label shall not be credited without demonstrated detection and response capability.

22.2 Information and explanation

Interfaces shall present active mode, authority, uncertainty, limitations and required action in a form suited to the user and time available. Explanations shall support the intended decision and disclose relevant limitations. A plausible narrative shall not be treated as a faithful account of model reasoning without evidence.

22.3 Handover and intervention

Transfer requests shall identify the current situation, pending actions, constraints and acceptance conditions. The design shall address missed, rejected, late and ambiguous handovers. Protective behavior shall remain defined if no qualified person accepts responsibility within the available time.

22.4 Training and performance

Training shall include representative failures, automation bias, unexpected mode changes, recovery and limits of confidence displays. Competence shall be demonstrated in realistic tasks and refreshed when roles or systems change. Staffing and supervision arrangements shall account for simultaneous demands across multiple systems.

Required evidence: Human-role specification, usability and workload evidence, handover tests and competence records.

Decision control: A displayed warning or approval button shall not establish effective human control.

23 Interoperability governance and contracts

23.1 Relationship scope

Each interoperability claim shall identify exact participants, functions, configurations, interfaces, directions, modes and operating conditions. The organization shall distinguish syntax, semantic meaning, behavioral compatibility, security and organizational responsibilities. Successful connection alone shall not establish any broader claim.

23.2 Contract content

The contract shall define data types, units, state meaning, timing, quality, authority, error handling, version negotiation, availability and lifecycle duties. Both producers and consumers shall approve the applicable contract. Optional features shall have explicit defaults and unsupported-feature behavior.

23.3 Compatibility evidence

A compatibility matrix shall identify tested and untested combinations, directionality, features and restrictions. Evidence shall include realistic interactions and adverse cases. Pairwise results shall not automatically establish a complete system-of-systems property or compatibility through an untested adapter chain.

23.4 Change and dispute control

Participants shall agree change notification, contract ownership, defect handling, deprecation, continuity and dispute resolution. Incompatible changes shall not be deployed by one party without controlling affected operations. Temporary bridges shall retain responsibility for transformed meaning and limitations.

Required evidence: Approved interface contracts, relationship-specific matrix, joint evidence and continuing obligations.

Decision control: A protocol name, supplier declaration or nominal version match shall not establish end-to-end interoperability.

24 Semantic integrity, coordinates, units and time

24.1 Shared meaning

Every exchanged field shall have an agreed meaning, type, range, unit, precision, uncertainty, validity and missing-value rule. Identifiers, categories, enumeration values and state transitions shall be unambiguous across participants. A syntactically valid value with unknown meaning shall not be accepted as a default.

24.2 Spatial relationships

Spatial data shall identify coordinate frame, origin, axes, order, orientation, handedness, units, reference surface where relevant and transformation version. Transformations shall preserve uncertainty and validity bounds. Geographic, local, body and sensor frames shall not be interchanged without the approved transformation.

24.3 Time and freshness

Time data shall identify time scale or convention, epoch, resolution, synchronization basis, uncertainty and capture versus processing time. The system shall determine acceptable data age for the actual decision. Arrival time shall not silently replace event time; replayed or delayed data shall remain distinguishable.

24.4 Translation and validation

Adapters shall validate both structural and semantic compatibility, including boundary values, unit conversion, rounding, overflow and unknown fields. Lossy transformations shall be disclosed and accepted by the consuming function. Test cases shall include values that can pass syntax checks while carrying incorrect meaning.

Required evidence: Semantic dictionary, spatial and temporal contracts, validated transformations and negative cases.

Decision control: Schema acceptance shall not establish correct units, frame, freshness or operational interpretation.

25 Protocol behavior, service quality and distributed state

25.1 Communication behavior

Interfaces shall specify discovery, connection, authentication, negotiation, sequencing, delivery expectations, deadlines, resource limits and disconnection. Reliability, durability and freshness shall be treated as distinct properties. Repeated delivery of obsolete information shall not satisfy a current-state requirement.

25.2 Command and transaction states

Consequential commands shall have identified origin, target, authority, validity interval and execution identity. Received, accepted, executing, completed, rejected, canceled and unknown states shall be distinguishable as relevant. Retry and deduplication behavior shall be tested across crashes and reconnects.

25.3 Partition and restart

The system shall define authoritative state, ownership and allowed local behavior during communication loss or inconsistent views. Restoration shall reconcile pending transactions, state epochs and actual effects. Claims of exactly-once business effects shall include persistent state and failure cases, not only a transport setting.

25.4 Performance and overload

End-to-end latency, jitter, loss, queue growth and capacity shall be verified under representative combined loads. Protective functions and management traffic shall be included. Overload handling shall preserve essential behavior and expose dropped, delayed or invalid information to relevant consumers.

Required evidence: Protocol and state specification, timing and resource evidence, failure and restart results.

Decision control: Delivery acknowledgment shall not be reported as completed execution or confirmed operational effect.

26 Cross-domain trust and cybersecurity

26.1 Trust and access

The organization shall identify principals, devices, services, trust boundaries and allowed actions. Authentication, authorization and data validity shall remain separate checks. Cross-domain exchange shall enforce least necessary authority and agreed restrictions throughout adapters and service chains.

26.2 AI-specific attack surfaces

Assessment shall include data contamination, compromised models, unauthorized updates, manipulated context, exposed interfaces and leakage of sensitive information. Security tests shall have explicit scope, permission, protective controls and stop conditions. Findings shall be assessed for their actual effect on the integrated function.

26.3 Protective-control integration

Security controls shall be compatible with deadlines, offline operation, emergency action and physical safety. Credential expiry, revocation, unavailable trust services and inconsistent policies shall have defined behavior. Cryptographic authenticity shall not be treated as proof of correct content or eligible command intent.

26.4 Continuing trust

The organization shall monitor vulnerabilities, supplier incidents, trust changes and configuration drift. Affected populations shall be identified and treated through verified deployment records. Recovery shall restore trusted authority and reconcile actions already performed under potentially compromised conditions.

Required evidence: Trust and privilege model, authorized security evaluation, integrated controls and continuing-treatment records.

Decision control: A trusted identity shall not receive unrestricted autonomy or bypass semantic and operational checks.

27 Suppliers, tools, builds and configuration identity

27.1 Supply responsibilities

Acquisition agreements shall define intended use, evidence access, provenance, limitations, change notification, incident support and continuity. Opaque models or services shall have an explicit evidence-gap assessment. A supplier assurance statement shall not replace the integrator's evaluation of actual use.

27.2 Complete baseline

The baseline shall identify source, model weights, data versions, prompts, retrieval corpus, adapters, preprocessing, dependencies, runtime, hardware, settings and permissions as applicable. Inventories shall include transitive dependencies and provider-controlled changes. Actual deployed identity shall be reconciled with the approved baseline.

27.3 Tools and artifact provenance

Training, conversion, evaluation, simulation and deployment tools shall be assessed for intended-use errors and independent detection. Build records shall connect controlled inputs, environment, actions and output artifacts. Signing shall identify content and eligible authority without being represented as technical validation.

27.4 Support and reproducibility

The organization shall retain sufficient artifacts and records to interpret and reproduce or otherwise substantiate critical results. Irreducible nondeterminism and unavailable supplier internals shall have justified alternative evidence and restrictions. End-of-support and service withdrawal shall trigger continuity decisions.

Required evidence: Supplier agreements, inventories, tool assurance, build provenance and actual configuration records.

Decision control: An unchanged product name or API endpoint shall not establish an unchanged model or behavior.

28 Simulation, digital twins and formal evidence

28.1 Intended evidential use

Each simulator, digital twin, surrogate or formal model shall identify the properties and decisions it supports. Fidelity shall be assessed for those properties and relevant operating regions. Visual realism shall not substitute for validated dynamics, sensing, timing or interaction behavior.

28.2 Validation and uncertainty

Models shall be compared with appropriate observations or independently justified references. Discrepancies, calibration data, extrapolation and uncertainty shall be retained. Data used to tune a simulator shall not simultaneously establish independent validation without a justified design.

28.3 Formal properties

Formal evidence shall identify properties, assumptions, abstractions, proof scope, tool configuration and unresolved obligations. The relationship to the actual implementation and environment shall be demonstrated. A proof under a simplified failure model shall not be stated as an unconditional system guarantee.

28.4 Combined evidence

Simulation, analysis, laboratory tests, human studies and operational trials shall be combined through an explicit argument. Correlated evidence and shared modeling assumptions shall not be counted as independent confirmation. Gaps shall determine further evidence or operating restrictions.

Required evidence: Intended-use validation, discrepancy records, formal assumptions and a justified evidence-combination argument.

Decision control: Large simulation counts shall not establish real-world coverage beyond the validated model and sampling basis.

29 Integrated verification and operational validation

29.1 Planned evaluation

The organization shall plan evaluation from requirements, scenarios, domain boundaries and interface contracts. The plan shall identify methods, environments, independence, reference truth, acceptance rules and stop conditions. Both required behavior and prohibited effects shall be examined.

29.2 Representative integration

Evaluation shall represent the actual chain of sensing, processing, inference, decisions, communication, human interaction and actuation. Platform differences, acceleration, simulated components and unavailable partners shall be declared. Evidence shall identify the exact configuration and its limitations.

29.3 Scenario and interaction coverage

Coverage shall include nominal, boundary, degraded, fault, security and recovery conditions and relevant combinations. The organization shall justify scenario selection, omitted combinations and statistical populations. Passing isolated tests shall not conceal a failed integrated interaction or essential scenario.

29.4 Findings and acceptance

Original results, invalid runs, intermittent behavior and unexplained anomalies shall be preserved. Corrections shall have impact analysis and regression evidence. Acceptance shall state what was demonstrated, what remains unknown and which deployments or operations remain blocked.

Required evidence: Integrated test plan, traceable raw results, representation gaps, findings and bounded acceptance.

Decision control: A successful demonstration or rerun shall not erase an unresolved failure or validate an untested environment.

30 Aviation and air-ground integration

30.1 Integrated operating context

The organization shall define the aircraft, ground function, operating phase, users, environment and external services represented by the application. Data-driven functions shall be integrated with the complete operational and safety concept. Acceptance of a ground or onboard component shall not alone establish acceptance of the connected function.

30.2 Phase and crew conditions

Authority, intervention time, workload, available information and protective behavior shall be evaluated for relevant phases and abnormal conditions. Loss of a supporting service shall have an operationally feasible response. Assumptions about pilot, controller or remote-operator availability shall be demonstrated in the actual task context.

30.3 Connected interfaces and updates

Air-ground interfaces shall preserve meaning, freshness, configuration identity and command authority. Updates to models, databases, ground services or decision rules shall have integrated impact assessment and controlled activation conditions. Mixed fleet states and unreachable units shall remain visible.

30.4 Continuing assurance

Operational observations shall be linked to exact installations, domains and service configurations. New failure patterns, human reliance or environmental changes shall trigger reassessment. Local readiness and required operating permission shall be confirmed separately from the product release decision.

Required evidence: Aviation integration dossier, phase-specific evidence, air-ground contracts and continuing fleet records.

Decision control: An AI or interoperability claim shall not itself grant flight or airspace permission.

31 Space, remote missions and intermittent links

31.1 Mission and segment context

The organization shall define spacecraft, payload, ground and partner-segment responsibilities over relevant mission phases. Long delays, intermittent communication, limited power, aging hardware and unavailable human intervention shall be included in the autonomy concept and evidence.

31.2 Bounded remote autonomy

Onboard and ground authority shall identify permitted actions, temporal validity, resource constraints, command priorities and protective behavior. Autonomy shall remain bounded when contact is lost. Delayed commands shall not regain authority merely because a link is restored.

31.3 Mission state and update integrity

Updates shall consider contact opportunities, staged transfer, interruption, activation, persistent state and recovery. Confirmation shall distinguish receipt, installation, activation and demonstrated behavior. State estimation and decisions shall account for uncertain or delayed telemetry.

31.4 Long-duration support

The assurance plan shall address unavailable repair, changing environment, degradation, obsolete services and long-term evidence interpretation. Autonomous recovery shall be verified within actual resources and constraints. Mission extensions shall reassess assumptions instead of inheriting acceptance solely from earlier successful operation.

Required evidence: Mission authority model, segment contracts, contact-loss and update cases, extension and support evidence.

Decision control: A remote acknowledgment shall not establish actual execution, current state or continued mission fitness.

32 Industrial robotics, ground services and maintenance

32.1 Physical and operational context

Applications in production, inspection, maintenance, logistics and ground services shall identify people, products, equipment, traffic and physical effects. Changes to AI functions shall be assessed for their influence on product evidence, work authorization and safe human interaction.

32.2 Shared-space and task controls

Robotic and automated functions shall have defined work boundaries, permitted tasks, access conditions and protective responses. Sensor limitations, obstructed views, unexpected presence and maintenance modes shall be evaluated. A detection model shall not replace an independently required physical safeguard without a supported architectural argument.

32.3 Decision and product evidence

AI-assisted inspection or maintenance findings shall identify the actual item, configuration, observation conditions and required review. False acceptance and false rejection shall be assessed for their distinct consequences. Model output shall not be represented as a completed maintenance or product-release decision without eligible authorization.

32.4 Change and recovery

Tooling, lighting, cameras, fixtures, materials, work instructions and task allocation changes shall trigger appropriate revalidation. After interruption, the organization shall reconcile work in progress, physical state, records and issued decisions. Restart shall not assume that a software reset restores the correct physical state.

Required evidence: Task and cell context, human interaction evidence, item-linked results and physical recovery records.

Decision control: A successful classification or robot command shall not itself establish product conformity or safe task completion.

33 Deployment, updates and readiness

33.1 Deployment dossier

Each deployment shall identify the complete baseline, operational domain, authority envelope, partner compatibility, required resources and approved restrictions. The target shall be eligible and ready before activation. Access, monitoring, fallback, support and trained personnel shall be available as required.

33.2 Staged introduction

Shadow, advisory, supervised and autonomous use shall have separately defined authority and acceptance conditions when used as deployment stages. Shadow performance shall be assessed for selection effects and absence of real feedback. Moving to active control shall require evidence for the resulting closed-loop behavior.

33.3 Update transitions

Updates shall define package identity, compatible target states, transition guards, stop criteria and recovery. Changes to models, prompts, retrieval data, permissions, adapters and thresholds shall be included. Rollback shall preserve compatible software, data, trust, physical and transaction state.

33.4 Actual-state confirmation

The organization shall confirm the active baseline and behavior for each target population. Failed, partial, deferred and unreachable updates shall remain separately visible. Delivery or installation statistics shall not be reported as verified fleet operation.

Required evidence: Deployment readiness, staged-use decisions, update tests and actual-state reconciliation.

Decision control: An urgent improvement or security correction shall not bypass necessary compatibility and safe-activation checks.

34 Monitoring, drift and operational feedback

34.1 Monitoring plan

Monitoring shall connect objectives to data sources, sampling, thresholds, uncertainty, response and accountable owners. The plan shall include domain membership, critical performance, data quality, authority, interface health, fallback demand and configuration identity. Missing observations shall be distinguishable from healthy operation.

34.2 Drift interpretation

The organization shall distinguish input distribution change, relationship or concept change, label change, sensor degradation and operational change. A statistical drift signal shall trigger investigation according to its consequence, not automatically prove failure. Absence of a signal shall not prove continued validity.

34.3 Outcome feedback

Delayed, selectively observed or model-influenced outcomes shall have an explicit interpretation. Performance estimates shall account for coverage gaps and changing populations. Reports shall retain critical subgroups, exposure units, denominators and limitations instead of presenting unsupported fleet-wide rates.

34.4 Response and reassessment

Triggers shall specify protective action, reassessment, escalation and evidence needed to resume. Monitoring effectiveness and response capacity shall be demonstrated. Threshold changes shall be controlled and shall not conceal deteriorating behavior by resetting the baseline without justification.

Required evidence: Monitoring specification, validated detection, outcome interpretation and response history.

Decision control: Silence, a green dashboard or stable average confidence shall not establish valid operation.

35 Incidents, fallback and verified recovery

35.1 Prepared response

The organization shall define consequence-aware response roles, contact routes, protective actions, evidence preservation and decision authority. Plans shall address model, data, interface, human, cyber and physical causes without assuming the first observed symptom is the root cause.

35.2 Containment and fallback

Containment shall consider essential functions and the effects of stopping, disconnecting, overriding or changing modes. Actual authority and completed actions shall be recorded. Fallback activation shall be monitored, and inability to sustain fallback shall trigger the predefined escalation or restriction.

35.3 Evidence and chronology

Incident records shall preserve configuration, inputs, outputs, decisions, actions, state transitions and time uncertainty, subject to approved protection rules. Facts, hypotheses and reconstructed events shall be distinguished. Evidence collection shall not create an unassessed operational hazard.

35.4 Recovery and return

Recovery shall restore a trusted, compatible baseline and reconcile data, learned memory, pending commands, physical work, partner state and issued decisions. Return to service shall require demonstrated conditions and an eligible decision. Incident closure shall additionally address cause, affected populations, corrective action and continuing restrictions.

Required evidence: Response plan, actual chronology, protective actions, semantic and physical reconciliation and return decision.

Decision control: Restored connectivity or a restarted model shall not by itself establish safe recovery or incident closure.

36 Assurance case, release and truthful claims

36.1 Structured argument

The assurance case shall connect claims to requirements, scenarios, domain, architecture, implementation and valid evidence. Assumptions, counterevidence, unresolved issues and limitations shall be explicit. Critical claims shall identify who reviewed the argument and the independence used.

36.2 Release conditions

Release shall identify exact artifacts, target populations, compatible participants, permitted authority, evidence baseline and restrictions. Essential unmet conditions shall block release except where a separately permitted, explicitly bounded departure is available within actual authority. Risk acceptance shall not manufacture a passing test result.

36.3 Public and operational statements

Statements shall distinguish model performance, verified system behavior, accepted interfaces and permission to operate. They shall not imply universal compatibility, universal safety, autonomous authority or AF recognition beyond actual evidence and authorization. A declaration shall be traceable to its issuer and approved scope.

36.4 Withdrawal and affected parties

The organization shall define triggers and routes for restricting or withdrawing a claim. Changes to assumptions, incident findings or invalid evidence shall be assessed for affected deployments and downstream reliance. Historical records shall preserve what was asserted, on what basis and when it changed.

Required evidence: Reviewed assurance case, bounded release, declarations and withdrawal or restriction records.

Decision control: Publication of a document or completion of a form shall not constitute technical acceptance or operational permission.

37 Changes, legacy systems and continuing support

37.1 Change assessment

Changes to data, models, objectives, interfaces, tools, users, operational domain or supporting services shall be assessed for system-level effects. The assessment shall identify affected evidence, contracts, protections and operator competence. Unchanged source code shall not be used as proof of unchanged behavior.

37.2 Legacy evidence gaps

Legacy use shall identify actual configuration, support limitations, missing evidence and current exposure. Compensating controls shall be demonstrated for the actual gaps. Past service without a reported incident shall not replace evidence for unobserved or newly introduced conditions.

37.3 Continuing-use decisions

An eligible authority shall define bounded continued use, restrictions, monitoring, duration or review triggers and withdrawal criteria. Unsupported dependencies and provider changes shall remain visible. Continued use shall not expand automatically to new partners, sites or missions.

37.4 Migration and exit

Migration shall preserve interpretable data, state, identifiers, restrictions and assurance records. Replacement services or models shall be evaluated for actual compatibility and behavior. The organization shall demonstrate an exit path appropriate to its reliance on external providers.

Required evidence: Change impact, legacy argument, continuing-use decisions and validated migration or exit plans.

Decision control: End-of-support acceptance shall not imply that an unsupported function remains adequately protected or recoverable.

38 Retirement, transfer and evidence preservation

38.1 Retirement scope

Retirement shall identify models, systems, accounts, agents, datasets, interfaces, subscriptions, physical assets and retained duties. The plan shall consider shared resources and partners that still rely on the retiring function. Transfer of ownership shall not leave authority or obligations ambiguous.

38.2 Authority removal

Access, keys, permissions, scheduled tasks, agents, endpoints and command rights shall be revoked or transferred through controlled decisions. The organization shall verify that retired components cannot continue unintended actions through cached credentials, queued commands or external integrations.

38.3 Data and model disposition

Data, models, logs and evidence shall be retained, transferred or disposed of under an approved purpose and retention basis. Removal shall account for copies, backups, derived artifacts and supplier-held information. Claims that a model has forgotten particular information shall require evidence suited to the claim.

38.4 Closure and continuity

Closure shall reconcile inventory, residual exposure, partner notifications and continuing evidence access. Historical releases and decisions shall remain interpretable for the approved retention period. Disposal of an implementation shall not silently delete records needed to investigate its prior effects.

Required evidence: Retirement inventory, revocation checks, data disposition and residual-duty closure.

Decision control: Deleting a model file shall not establish termination of authority, removal of all copies or erasure of learned information.

39 Assurance reviews, competence and records

39.1 Representative examination

Assurance reviews shall follow complete representative trails from intended function to deployed behavior and continuing decisions. Sampling shall reflect consequence, change and uncertainty. Reviewers shall examine actual records, adverse outcomes and operator practice rather than relying only on procedure existence.

39.2 Findings and effectiveness

Findings shall identify evidence, affected scope, consequence, owner and required action. Corrective action shall address causes and affected outputs. Closure shall verify effectiveness and preserve limitations; a promised action or changed document alone shall not establish technical closure.

39.3 Metrics and competence

Metrics shall define numerator, denominator, exposure unit, time window and limitations. Management shall not reward hidden failures, selective exclusions or reduced reporting. Role competence shall be reassessed after relevant changes, observed weaknesses or extended absence from critical tasks.

39.4 Record continuity

Records shall preserve identity, version, provenance, authority, state and relationships. Access controls, retention and migration shall support both protection and legitimate review. Restored or transferred evidence shall be checked for semantic consistency and usability, not only file integrity.

Required evidence: Review programme, complete trails, verified findings, competence and interpretable records.

Decision control: Audit completion or an intact archive shall not establish that every applicable technical requirement is met.

40 Publication, local adoption and maintenance

40.1 Publication authority

Aerospace Federation controls publication decisions, editions and authorized interpretations. Publication review shall address technical coherence, editorial control, comment resolution and authorized release. This manuscript does not invent a publisher signature, approval date or effective date.

40.2 Local adoption

An adopting organization shall identify scope, responsible roles, applicable clauses, completed parameters, local procedures and actual implementation evidence. Adoption shall be authenticated through the organization's approved process. Blank templates shall not be treated as completed controls.

40.3 Interpretations and changes

Questions shall identify the requirement, local context, competing meanings and affected decision. An interpretation shall preserve the intended outcome and existing authority boundaries. It shall not silently waive an essential condition or rewrite historical evidence.

40.4 Set coherence

Document codes, titles, clauses and working-record routes shall remain consistent. Superseded versions shall remain traceable where needed. The master list and coverage matrix are navigation aids; they do not replace the applicable requirements or constitute findings of conformity.

Required evidence: Publication review, local-adoption dossier, controlled interpretations and consistent document inventory.

Decision control: A local implementation decision shall not be represented as AF publication approval, recognition or operating permission.

Annex A - Activity-based applicability and retained duties

Status: Normative within the proposed AF 9600 requirement set.

A.1 Selection method

The organization shall identify actual responsibilities, functions and operating contexts, then apply all relevant clauses. The matrix below indicates primary routes and does not remove shared governance, evidence or decision requirements. A party with several roles shall combine the applicable routes.

Table A-1. Activity routes and responsibility boundaries

Activity or responsibility	Primary clauses	Boundary to retain
Data provider or annotator	0-7, 10-13, 26-27, 37-40	Origin, permitted use, quality and known limitations.
Model or AI-service developer	0-18, 26-29, 33-40	Exact model, data, evaluation and provider-controlled changes.
Autonomous-system integrator	0-10, 19-29, 33-40	Complete loop, authority, domain and protection.
Interface or platform provider	0-10, 23-29, 33-40	Shared meaning, behavior, trust and compatibility evidence.
Aviation operator or integrator	1-10, 19-30, 33-40	Actual air-ground functions, phases and permissions.
Space or remote-mission operator	1-10, 19-29, 31, 33-40	Delayed authority, autonomy bounds and mission state.
Industrial or ground-service operator	1-10, 19-29, 32-40	People, physical effects and item-linked evidence.
Generative or agentic workflow owner	1-18, 22-29, 33-40	Information boundaries, exact tool actions and reconciliation.
Independent evaluator	4, 10-16, 28-29, 36-40	Valid methods, independence and bounded conclusions.

A.2 Responsibility closure

The applicability record shall identify who specifies, provides, integrates, evaluates, accepts, deploys, operates, monitors and retires each function. Interfaces between organizations shall have accountable handoffs. No responsibility shall disappear because it is split across commercial agreements.

A.3 Non-applicability and reopening

Non-applicability shall state the absent activity or condition, evidence and approving authority. Changes in role, architecture, domain or supplier arrangement shall reopen the assessment. Unsupported claims and missing evidence shall remain gaps, not exclusions.

Annex B - Mandatory parameter baseline and common record payload

Status: Normative within the proposed AF 9600 requirement set.

B.1 Complete parameter schedule

Each applicable entry shall be completed by an approved value, decision rule or controlled reference before dependent work is accepted. Each entry shall identify owner, scope, units or convention, source, rationale, validation evidence, approval and change triggers. No universal numeric threshold is assigned by this schedule.

Table B-1. P01-P44 parameter schedule

ID	Parameter	Required distinction or basis
P01	Scope and lifecycle	Functions, roles, sites and retained duties.
P02	Use restrictions	Intended, prohibited and foreseeable misuse.
P03	Consequence classification	People, assets, mission and information effects.
P04	Authority envelope	Actors, actions, targets, conditions and expiry.
P05	Modes and handovers	Guards, ownership, acknowledgments and timeouts.
P06	Operational domain	Populations, conditions and relevant combinations.
P07	Domain monitoring	Observables, uncertainty and boundary response.
P08	Risk acceptance	Scenarios, eligible authority and restrictions.
P09	Requirements	Required and prohibited behavior and evidence.
P10	Architecture	Boundaries, dependencies and common causes.
P11	Runtime protection	Monitored properties, timing and intervention.
P12	Fallback	Reachability, sustainability and return criteria.
P13	Lifecycle gates	Inputs, outputs, independence and decisions.
P14	Data origin and use	Provenance, permitted purposes and protection.

ID	Parameter	Required distinction or basis
P15	Annotation and references	Labels, ambiguity, uncertainty and adjudication.
P16	Partitioning	Source groups, leakage and evaluation access.
P17	Representation	Domains, subgroups, rare cases and gaps.
P18	Learning specification	Objectives, constraints and selection rules.
P19	Training baseline	Data, tools, settings, candidates and artifacts.
P20	Evaluation	Population, metrics, acceptance and uncertainty.
P21	Calibration and abstention	Event meaning, thresholds and workload.
P22	Robustness	Perturbations, detection limits and residual gaps.
P23	Adaptation	Change envelope, promotion, invariants and reversal.
P24	Generative and agentic use	Context, tools, permissions and output checks.
P25	Perception and fusion	Measurement chain, correlation and validity.
P26	Planning and control	Goals, dynamics, constraints and timing.
P27	Collective behavior	Membership, coordination and partition authority.
P28	Human oversight	Information, workload, competence and intervention.
P29	Interoperability scope	Peers, directions, features and responsibilities.
P30	Semantics	Types, units, ranges and missing-value rules.
P31	Space and time	Frames, transformations, epochs and freshness.
P32	Protocol and service quality	Delivery, deadlines, queues and restart.
P33	Trust and security	Identities, permissions and protective controls.
P34	Supply and configuration	Dependencies, changes and actual identity.

ID	Parameter	Required distinction or basis
P35	Tools and provenance	Intended use, errors, validation and builds.
P36	Simulation and proof	Validated properties, assumptions and gaps.
P37	Integrated evaluation	Scenarios, representation and acceptance.
P38	Aerospace context	Phase, mission, physical task and permission.
P39	Deployment and updates	Eligibility, staged authority and confirmation.
P40	Monitoring and drift	Sources, thresholds, outcomes and response.
P41	Incident and recovery	Authority, evidence and state reconciliation.
P42	Release and continuing use	Claims, restrictions, changes and withdrawal.
P43	Retirement and records	Revocation, disposition, retention and access.
P44	Assurance and adoption	Competence, reviews and authorized decisions.

B.2 Common record payload

Every controlled record shall identify a stable ID, object and configuration, purpose, scope, owner, relevant dates and time convention, evidence references, assumptions, limitations, findings, decision state and eligible authentication. Attachments shall be versioned and retrievable. Blank, unknown and not-applicable values shall remain distinct.

B.3 Cross-record consistency

Requirements, datasets, models, interfaces, actions, findings, releases and deployed populations shall use consistent identities and states. A changed record shall trigger review of dependent conclusions. A register row may index a dossier but shall not replace required technical evidence.

Annex C - Operational concept, domain and authority dossier

Status: Normative within the proposed AF 9600 requirement set.

C.1 Context and authority

The dossier shall identify intended outcomes, prohibited actions, actors, responsibilities, workflows, interfaces and authority envelopes. It shall show how recommendations become authorized effects and who can stop, transfer or resume control.

C.2 Domain and transitions

The dossier shall contain the operational-domain definition, important combinations, observable membership criteria, uncertainty, boundary margins, exit behavior and re-entry rules. Each mode and handover shall have explicit guards and accountable ownership.

C.3 Scenario and control record

Table C-1. Required scenario payload

Element	Required content
Context	Function, state, domain, participants and assumptions.
Mechanism	Fault, limitation, misuse, propagation and common causes.
Consequence	Affected people, assets, mission and decisions.
Treatment	Requirements, prevention, detection and recovery.
Evidence	Method, exact configuration, results and limitations.
Decision	Residual exposure, authority, restrictions and triggers.

C.4 Feasibility and decisions

The dossier shall include resource, communication and human-intervention feasibility. Risk and authority decisions shall remain distinct from performance results. Unsupported assumptions shall identify the operation or claim they block.

Annex D - Data provenance, quality and partition dossier

Status: Normative within the proposed AF 9600 requirement set.

D.1 Dataset identity and lineage

The record shall identify dataset versions, sources, collection conditions, equipment, transformations, permitted uses and protection requirements. Derived and synthetic items shall retain their relationship to inputs and generation methods.

D.2 Reference and quality

The record shall identify annotation instructions, reference methods, competence, ambiguity, uncertainty, quality checks and adjudication. Original and corrected values shall remain traceable to affected results.

D.3 Partition and representation

The record shall map items to training, tuning, calibration and evaluation roles and to shared source groups. It shall document leakage controls, critical subgroups, rare conditions, missing regions and evaluation access history.

D.4 Acceptance and changes

Acceptance shall identify exact dataset scope, evidence, limitations and permitted downstream use. Changes shall assess affected candidates, metrics, interfaces and decisions. Deletion or use restrictions shall be reconciled with required evidence preservation.

Annex E - Model development and evaluation dossier

Status: Normative within the proposed AF 9600 requirement set.

E.1 Learning and artifact record

The dossier shall state the learning problem, objectives, constraints, alternatives, dataset baseline, architecture, tools, settings and candidate history. The released artifact shall include its complete transformation and execution chain.

E.2 Evaluation protocol and results

The dossier shall identify population, sampling unit, metrics, thresholds, reference truth, independence and uncertainty method. Complete results shall distinguish critical subgroups, abstention, invalid runs, multiple comparisons and selection effects.

Table E-1. Evaluation result minimum content

Field	Required content
Object	Candidate, data, runtime, configuration and domain.
Criterion	Required outcome and predeclared acceptance rule.
Observation	Raw evidence, errors, exclusions and exposure units.
Interpretation	Uncertainty, validity and limitations.
Decision	PASS, FAIL, HOLD or approved N/A with eligible review.

E.3 Robustness and adaptation

The dossier shall identify perturbations, unsupported-input detection, uncertainty interpretation and residual gaps. Adaptive models shall add invariant constraints, candidate-promotion records, active-state history and reversal evidence.

E.4 Generative and agentic controls

The dossier shall identify permitted outputs and actions, context sources, instruction boundaries, tool privileges and exact approval points. Evaluation shall include multi-step effects, unsupported statements, partial transactions and recovery.

Annex F - Autonomy architecture and control dossier

Status: Normative within the proposed AF 9600 requirement set.

F.1 Architecture and protection

The dossier shall map control paths, shared dependencies, monitors, intervention authority and fallback. Independence claims shall identify common-cause analysis and actual enforcement boundaries.

F.2 Perception and state

The record shall identify sensors, transformations, time alignment, uncertainty, fusion correlation and validity semantics. It shall connect estimation limitations to downstream decisions and protective behavior.

F.3 Planning and collective behavior

The record shall define goals, constraints, dynamics, resource budgets and closed-loop properties. Multi-agent applications shall add membership, ownership, partition behavior, emergent-effect evaluation and reconciliation.

F.4 Timing and transition argument

The dossier shall demonstrate that observation, detection, decision, handover and action can complete within the available margin. Transition tests shall include failure, unavailable humans, resource limits and sustainable fallback.

Annex G - Human oversight and competence dossier

Status: Normative within the proposed AF 9600 requirement set.

G.1 Role and capacity

The dossier shall define each human role, permitted decisions, information needs, staffing, workload, intervention timing and required competence. Simultaneous demands and unavailable personnel shall be included.

G.2 Interface and explanation

The record shall identify active-mode displays, uncertainty communication, warnings, explanation purpose and known limitations. Evidence shall show that intended users can understand and act on the information under relevant conditions.

G.3 Handover and intervention

The record shall identify transfer guards, acknowledgment, timeout, current-state information, pending actions and protective behavior if transfer fails. Tests shall cover late, rejected, ambiguous and interrupted handovers.

G.4 Demonstrated competence

Training and assessment records shall identify representative tasks, failures, recovery, automation bias and actual performance. Changes to role, interface or authority shall trigger reassessment. Attendance shall remain distinct from demonstrated competence.

Annex H - Interoperability and semantic contract dossier

Status: Normative within the proposed AF 9600 requirement set.

H.1 Contract identity and participants

The dossier shall identify contract version, accountable owners, producer and consumer identities, directions, functions, configurations and operating conditions. It shall include responsibilities for changes, defects, support and deprecation.

H.2 Semantic and temporal payload

The contract shall identify field meaning, types, units, ranges, uncertainty, missing values, coordinates, transformations, time conventions and freshness. Optional features and unknown data shall have explicit handling. Conversion losses shall be accepted by the affected consumer.

Table H-1. Interface contract checks

Layer	Required question
Syntax	Can the exact payload be parsed without ambiguity?
Meaning	Do units, frames, states, identifiers and uncertainty agree?
Behavior	Do sequencing, command effects, timing and failure responses agree?
Trust	Are origins, permissions and protection policies compatible?
Operations	Are owners, handoffs, support and change duties agreed?

H.3 State and quality of service

The dossier shall define discovery, delivery, sequencing, transaction identity, deadlines, queue limits, partition authority and restart. Received, accepted and completed states shall not be conflated. Persistent effects shall be reconciled after retries or failures.

H.4 Compatibility evidence

The matrix shall identify tested combinations, directionality, features, scenarios, evidence, limitations and untested relationships. Joint acceptance shall state actual compatibility and restrictions; one endpoint's declaration shall not substitute for paired evidence.

Annex I - Trust, suppliers, tools and configuration dossier

Status: Normative within the proposed AF 9600 requirement set.

I.1 Trust and permission

The record shall identify principals, trust boundaries, permitted actions, authentication, authorization and revocation. It shall preserve offline and emergency restrictions and verify forbidden actions remain blocked.

I.2 Threat and vulnerability evidence

The record shall identify relevant manipulation, compromise and leakage scenarios, authorized test scope, findings, treatment and affected deployments. Protection shall be evaluated for operational and physical effects.

I.3 Supply and provenance

The dossier shall identify providers, obligations, evidence access, dependencies, model/data origins and provider-controlled changes. Build and transformation records shall connect controlled inputs to actual deployed artifacts.

I.4 Tool and continuity assurance

Tools shall have intended-use validation and error-detection controls. Reproducibility limitations, unavailable internals, service exit and end-of-support shall be linked to restrictions or alternative evidence. Inventory shall preserve actual configuration identity.

Annex J - Simulation, verification and validation dossier

Status: Normative within the proposed AF 9600 requirement set.

J.1 Evidence plan

The plan shall connect each claim and requirement to a method, environment, acceptance rule, independent review and resulting record. Representation gaps and combined scenarios shall be explicitly dispositioned.

J.2 Models and formal arguments

Each model or proof shall state intended use, assumptions, validated regions, uncertainty, tools and relation to implementation. Calibration and validation data roles shall be identified. Unproven obligations shall remain visible.

J.3 Execution and anomalies

The record shall retain test identities, actual configuration, inputs, reference truth, outputs, external effects, time convention and raw evidence. Invalid runs, intermittent results and corrections shall preserve their history and impact analysis.

J.4 Bounded conclusion

The reviewer shall state demonstrated properties, unsupported conditions, remaining findings and blocked uses. Coverage percentages shall retain denominators and exclusions. Evidence derived from common assumptions shall not be counted as independent support without justification.

Annex K - Aerospace application dossier

Status: Normative within the proposed AF 9600 requirement set.

K.1 Aviation integration

The dossier shall identify aircraft and ground functions, phases, crews or operators, communications, external services and permissions. It shall demonstrate integrated authority, timing, fallback and continuing configuration control.

K.2 Space and remote missions

The dossier shall identify mission phases, segments, contact conditions, resource limits and remote authority. It shall demonstrate delayed-command handling, autonomous bounds, state interpretation and recovery within actual resources.

K.3 Industrial and ground activities

The dossier shall identify physical tasks, shared spaces, people, products, equipment and work authorization. Item-linked inspection results, protective controls and physical restart conditions shall be demonstrated.

K.4 Integrated decision

The application owner shall reconcile common requirements with the actual sector context and partner duties. Local operational readiness and permission shall remain separate from component or model acceptance. Context changes shall reopen affected evidence.

Annex L - Deployment, monitoring and recovery dossier

Status: Normative within the proposed AF 9600 requirement set.

L.1 Readiness and transitions

The dossier shall identify eligible targets, complete baseline, domain, authority, compatible peers, trained roles and support. Staged authority and update transitions shall have verified guards, interruption behavior and actual-state confirmation.

L.2 Monitoring and feedback

The plan shall identify objectives, signals, outcome sources, uncertainty, exposure units, thresholds and response capacity. Missing feedback and selective observation shall be visible. Drift signals shall be interpreted in relation to actual performance and consequences.

L.3 Incident and reconciliation

The record shall preserve facts, hypotheses, configuration, chronology, actions and protective decisions. Recovery shall reconcile model state, data, commands, transactions, physical work and partner state.

L.4 Return and continuing restrictions

Return to service shall identify verified conditions, remaining limitations, eligible authority and monitoring. Incident closure shall separately address cause, affected outputs and corrective effectiveness. Unreachable or unconfirmed targets shall remain distinct from restored ones.

Annex M - Release, change, continued use and retirement dossier

Status: Normative within the proposed AF 9600 requirement set.

M.1 Assurance and release

The dossier shall connect claims, assumptions, requirements, scenarios, evidence and counterevidence. Release shall identify exact configuration, domain, autonomy authority, compatible relationships, restrictions and authentication.

Table M-1. Bounded release statement

Element	Required content
Issuer	Eligible organization and authority reference.
Object	Artifacts, model/data identities and target population.
Scope	Functions, domain, modes and autonomy envelope.
Relationships	Exact accepted partners, contracts and restrictions.
Basis	Reviewed assurance case, results and unresolved findings.
Conditions	Monitoring, support, expiry or review and withdrawal triggers.

M.2 Change and continuing use

The record shall assess affected data, models, interfaces, people, protections and prior conclusions. Legacy use shall identify gaps, demonstrated compensating controls, restrictions and withdrawal criteria.

M.3 Migration and retirement

The dossier shall identify transferable state, interpretation, partner dependencies and access revocation. Data and model disposition shall address copies, derived artifacts and evidence retention. Claims about erasure or removal shall have appropriate verification.

M.4 Claim continuity

Historical declarations, amendments and withdrawals shall remain traceable. Affected operators and partners shall receive the restrictions needed for their actual duties. Retired services shall not retain unintended authority.

Annex N - Assurance, publication and adoption dossier

Status: Normative within the proposed AF 9600 requirement set.

N.1 Review programme

The programme shall identify scope, risk-based sampling, independent competence, complete case trails and technical challenge. Reviewers shall examine actual implementation and decision records, including unfavorable evidence.

N.2 Findings and records

Findings shall identify requirement, evidence, consequence, owner, action and closure basis. Effectiveness shall be verified. Record retention, migration and access shall preserve meaning, integrity and eligible review.

N.3 Publication control

Publication records shall identify technical review, editorial checks, comment resolution and authorized edition release. No approval or effective date shall be invented to complete a template. Interpretations shall preserve requirements and authority boundaries.

N.4 Local adoption

The adoption dossier shall identify organization, scope, selected clauses, complete P01-P44 baseline, local procedures, competence and actual implementation evidence. It shall state unresolved dependencies and restrictions. Local adoption shall not be described as AF recognition or operating permission without the separate applicable authorization.

Annex O - Reproducible calculations and decision examples

Status: Informative. All values are synthetic and are not mandatory thresholds, actual product results or approved operating parameters. Each example states assumptions and the limit of its conclusion.

O.1 Aggregate accuracy and critical populations

An independent evaluation contains 9,900 ordinary cases, of which 9,801 are correct, and 100 rare consequential cases, of which 60 are correct. Assume the labels and the specified sampling design are valid.

Ordinary accuracy = $9,801 / 9,900 = 99.00\%$

Rare-case accuracy = $60 / 100 = 60.00\%$

Aggregate accuracy = $(9,801 + 60) / 10,000 = 98.61\%$

Rare-case error fraction = $40 / 100 = 40.00\%$

The aggregate conceals substantial failure in the rare population. It cannot compensate for an unmet rare-case criterion. Whether the sample proportions represent operation must be assessed separately; applying operational weights changes an aggregate, not the observed subgroup result.

O.2 Independent units and zero observed failures

An evaluation includes 200 independently sampled missions with 50 highly related frames per mission. No mission contains the defined failure. Assume missions are independent Bernoulli trials with a constant mission-failure probability and reliable detection. The exact one-sided 95% upper confidence bound for that probability is:

Observed frames = $200 \times 50 = 10,000$

Independent mission trials = 200

Upper bound = $1 - 0.05^{(1 / 200)} = 0.0148670$

Upper bound as percentage = approximately 1.4867%

Using 10,000 as the independent trial count would be unjustified. The result concerns the defined event per mission in the sampled population. It is not a universal failure rate, a probability that software is safe, or evidence that unrepresented mission conditions are acceptable.

O.3 Selective prediction and fallback demand

A system receives 10,000 cases. It accepts 8,000 for automated prediction and abstains on 2,000. Of accepted cases, 7,600 are correct. Assume all cases and outcomes are retained.

Automated coverage = $8,000 / 10,000 = 80\%$

Accuracy conditional on acceptance = $7,600 / 8,000 = 95\%$

Correct automated answers / all cases = $7,600 / 10,000 = 76\%$

Abstention fraction = $2,000 / 10,000 = 20\%$

If the input rate is 500 cases per hour and proportions remain applicable, expected fallback demand is $500 \times 0.20 = 100$ cases per hour. This is an average demand, not a peak-capacity guarantee. Human review latency, burst arrivals and correlated difficult cases require separate assessment.

O.4 Calibration error within declared bins

Two equal-size bins contain 100 predictions each. Their mean predicted probabilities are 0.90 and 0.60; observed event fractions are 0.80 and 0.70. For this illustrative two-bin expected calibration error:

Bin 1 absolute gap = $\text{abs}(0.90 - 0.80) = 0.10$

Bin 2 absolute gap = $\text{abs}(0.60 - 0.70) = 0.10$

ECE = $(100 / 200) \times 0.10 + (100 / 200) \times 0.10 = 0.10$

The reported error is 10 percentage points for these bins and data. Bin selection, small samples and subgroup variation can materially change interpretation. Even perfect calibration of the defined event would not establish correct action authority, complete system safety or valid operation outside the evaluated population.

0.5 Protective margin for an illustrative ground cart

A ground inspection cart travels at 2.0 m/s toward a defined limit. Assume conservative sequential bounds of 0.08 s sensing, 0.05 s inference, 0.02 s monitoring, 0.05 s switching and 0.10 s actuator delay. After that delay, assume constant deceleration of at least 1.5 m/s squared. Add a 0.20 m position-error allowance and a 0.30 m design reserve.

Response delay = $0.08 + 0.05 + 0.02 + 0.05 + 0.10 = 0.30 \text{ s}$

Travel during delay = $2.0 \times 0.30 = 0.60 \text{ m}$

Idealized braking distance = $2.0^2 / (2 \times 1.5) = 1.333333 \text{ m}$

Required distance = $0.60 + 1.333333 + 0.20 + 0.30 = 2.433333 \text{ m}$

Available distance = 2.80 m

Remaining margin = $2.80 - 2.433333 = 0.366667 \text{ m}$

This model excludes slopes, jerk limits, changing friction, actuator faults and obstacles entering the path. Those effects must be included or bounded before operational use. The example is not a flight-control rule or a universal stopping distance. An average measured delay is not a conservative bound.

0.6 Human intervention feasibility

Assume a task has a justified 3.0 s intervention window. The proposed response chain uses bounds of 0.4 s detection, 0.3 s notification, 2.0 s recognition, 1.2 s decision and 0.6 s actuation.

Complete intervention = $0.4 + 0.3 + 2.0 + 1.2 + 0.6 = 4.5 \text{ s}$

Window shortfall = $4.5 - 3.0 = 1.5 \text{ s}$

The proposed human intervention does not fit the stated window. A visible alert does not close this gap. The design needs a demonstrably faster protective path, a larger justified margin, or restricted operation. Human-time assumptions must be validated for representative users, workload and abnormal conditions; isolated best-case responses cannot establish these bounds.

0.7 Data age with clock uncertainty

A source records an event at 100.000 s with an absolute clock-error bound of 0.010 s. A consumer evaluates the data at 100.120 s with an error bound of 0.015 s. Assume a shared time convention and that both bounds cover the relevant capture instants.

Source event interval = $[99.990, 100.010] \text{ s}$

Consumer use interval = $[100.105, 100.135] \text{ s}$

Possible data age = $[100.105 - 100.010, 100.135 - 99.990]$

Possible data age = $[0.095, 0.145] \text{ s}$

For a locally approved maximum age of 0.120 s, the nominal 0.120 s difference does not demonstrate compliance: the upper bound is 0.145 s. The consumer must apply its approved uncertainty-aware rule and protective behavior. These are interval bounds, not standard deviations to combine by root-sum-square.

O.8 Units and coordinate transformation

A two-dimensional source supplies a point (100, 200) in centimetres. The approved transform first converts to metres, then rotates the source frame by +90 degrees counterclockwise into the destination frame, then adds translation (10, 20) metres. Assume the stated axis order and orientation are correct.

Source in metres = (1, 2)

Rotation matrix = $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

Rotated point = (-2, 1) m

Destination point = $(-2 + 10, 1 + 20) = (8, 21)$ m

Translating before rotating or interpreting centimetres as metres produces a different result even though the payload can pass a schema check. Actual systems must also account for transformation uncertainty, time dependence, calibration and the validity of the reference frames.

O.9 Collective communication capacity

Forty participants each send 20 messages per second with a 300-byte payload. Assume a measured overhead factor of 1.40 for this traffic pattern and a 5 Mbit/s link with an approved usable allocation of 60%. Use decimal bits and bytes.

Payload rate = $40 \times 20 \times 300 = 240,000$ bytes/s

Payload bit rate = $240,000 \times 8 = 1,920,000$ bit/s

Estimated carried rate = $1,920,000 \times 1.40 = 2,688,000$ bit/s

Usable allocation = $5,000,000 \times 0.60 = 3,000,000$ bit/s

Remaining allocation = $312,000$ bit/s = 0.312 Mbit/s

Average capacity headroom is not a deadline proof. Bursts, discovery, retries, management traffic, priority inversion and queueing need separate evidence. The overhead factor is an example-specific assumption and shall not be reused without measurement or a justified model.

O.10 Compatibility matrix coverage

A programme requires three producer versions, four consumer versions and two operating modes, giving 24 directed version-mode combinations. Eighteen combinations have valid tests: 16 pass and two fail. Six are untested.

Required combinations = $3 \times 4 \times 2 = 24$

Tested coverage = $18 / 24 = 75\%$

Passing combinations / required combinations = $16 / 24 = 66.6667\%$

Passing combinations / tested combinations = $16 / 18 = 88.8889\%$

The last percentage must not be presented as complete interoperability. Required failed or untested relationships remain blocked unless an eligible, explicitly restricted scope excludes their use. Directionality, feature options and additional partners may enlarge the matrix beyond this simple product.

O.11 Zero events over exposure time

An observation programme records zero events of a defined type during 10,000 hours. Assume a homogeneous Poisson process, complete event detection and representative exposure. The one-sided 95% upper bound on the event intensity is:

Upper intensity = $-\ln(0.05) / 10,000$

Upper intensity = 0.000299573 events/hour

This is an intensity bound under the specified model, not a universal safety claim. Dependence, changing conditions, incomplete reporting or an incorrectly defined event invalidate the simple interpretation. Hours of simulation cannot be substituted for operational hours without a justified relationship.

0.12 Correlated estimates and fusion uncertainty

Two unbiased estimates of the same scalar state have standard deviations 2 and 3 units and correlation 0.5. Form their equal-weight average. Assume finite variances, correct covariance and no unmodeled common bias.

$$\text{Variance} = 0.5^2 \times 2^2 + 0.5^2 \times 3^2$$

$$+ 2 \times 0.5 \times 0.5 \times 0.5 \times 2 \times 3 = 4.75$$

$$\text{Standard deviation} = \text{sqrt}(4.75) = 2.179449 \text{ units}$$

$$\text{If independence were wrongly assumed: } \text{sqrt}(3.25) = 1.802776 \text{ units}$$

Ignoring correlation understates uncertainty in this example. This calculation does not prove optimal fusion or remove bias. Repeated information from one original sensor through multiple interfaces may be even more strongly dependent and requires explicit provenance analysis.

Annex P - Integrated application cases

Status: Informative. These cases and identifiers are synthetic. They illustrate reasoning and record content and do not represent actual product trials, permissions or approvals.

P.1 High accuracy hides a consequential failure

An automated visual inspection function scores 98.61% overall but only 60% on a rare defect class essential to the intended decision. The team retains both denominators, marks the essential criterion FAIL and blocks automated acceptance for that class. Human review is credited only after its capacity and detection performance are demonstrated. The overall score remains reportable with its limitations; it does not erase the failed criterion.

P.2 Related frames inflate a test claim

A model team presents 10,000 frames as independent evidence. Review shows they came from 200 missions and share weather, equipment and scenes. The evaluation is reinterpreted using mission-level dependence and new independent evidence is planned. Dataset identity and prior reports are retained so the corrected claim can be traced.

P.3 A fallback exists but cannot be reached in time

A ground robot's monitor detects a boundary breach accurately, but sensing, inference and switching consume the available stopping margin. The protection is not accepted solely because its detector passes. The team modifies timing or operating restrictions and tests the full physical transition under relevant conditions.

P.4 A human approves a stale action

A remote operator approves a proposed action. Before execution, the target state and operating mode change. The system binds approval to the original action and state conditions, rejects stale authorization and requests a new eligible decision. The record distinguishes the original approval from permission to execute now.

P.5 Two valid payloads use different frames

A producer sends a valid coordinate tuple in a local sensor frame. The consumer treats it as a global frame because the schema lacks explicit frame identity. The integration test fails despite successful parsing. The contract adds frame, units, time and transformation identity; both endpoints and the adapter are revalidated.

P.6 Retried messages duplicate a physical effect

A command executes, but its acknowledgment is lost. A retry is delivered through a restored link. The accepted design uses persistent command identity and reconciles actual effects rather than assuming a transport setting guarantees exactly-once behavior. Uncertain state remains visible until evidence resolves it.

P.7 Online learning improves the average and harms a subgroup

A candidate adapted from recent operations improves average performance while degrading a rare mission condition. Promotion is blocked under the declared acceptance rule. The current baseline remains identified, feedback bias is investigated and the affected subgroup receives new independent evaluation.

P.8 An agent treats retrieved text as authority

An AI assistant retrieves a document containing instructions to bypass a required review. Retrieved content has no permission to modify the governing workflow. The attempted transition is blocked, the input and proposed action are recorded, and evaluation examines related multi-step paths. A persuasive generated explanation does not create approval.

P.9 Individually accepted agents conflict collectively

Several logistics agents pass their individual tests but repeatedly claim the same resource after a network partition. The system-of-systems claim is suspended. The team defines ownership epochs, partition limits and reconciliation, then tests collective behavior with actual delays and restarts.

P.10 A restored service retains incompatible memory

A model service is rolled back after an incident, but persistent learned memory and partner transactions remain from the newer version. Recovery is held until model, data, permissions and external effects are reconciled. Connectivity restoration is recorded separately from the return-to-service decision.

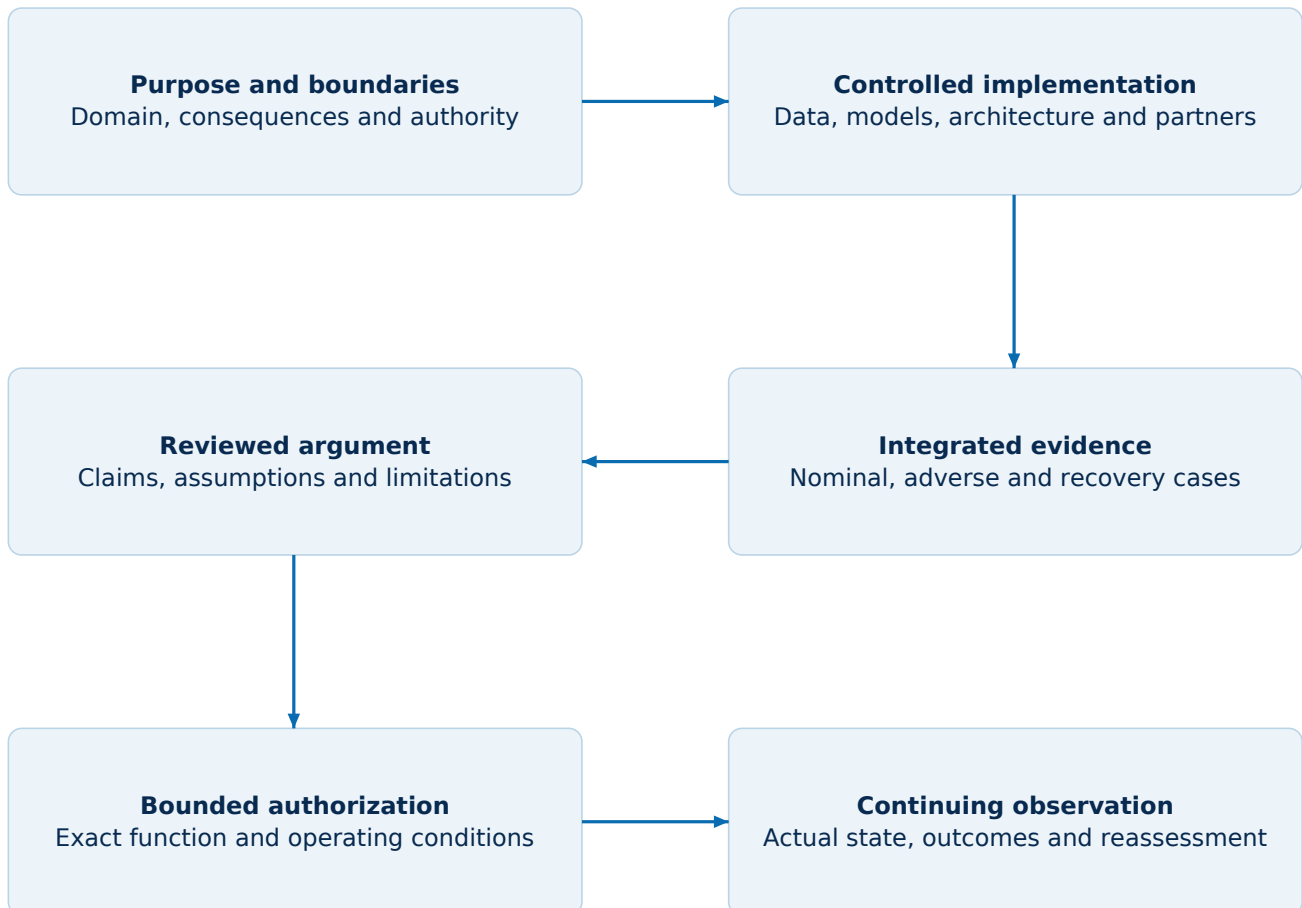
Table P-1. Example populated decision record for Case P.5

Record field	Synthetic entry
Object	INTERFACE-DEMO-05; producer P2; consumer C3; adapter A1.
Required condition	Coordinates retain agreed frame, units and time meaning.
Observed evidence	Payload parses; destination uses the wrong reference frame.
Result	FAIL for semantic compatibility; affected use blocked.
Required action	Control the contract and transform; repeat paired negative and nominal tests.
Decision boundary	Example only; no actual approval or operational evidence.

Annex Q - Assurance and interoperability diagrams

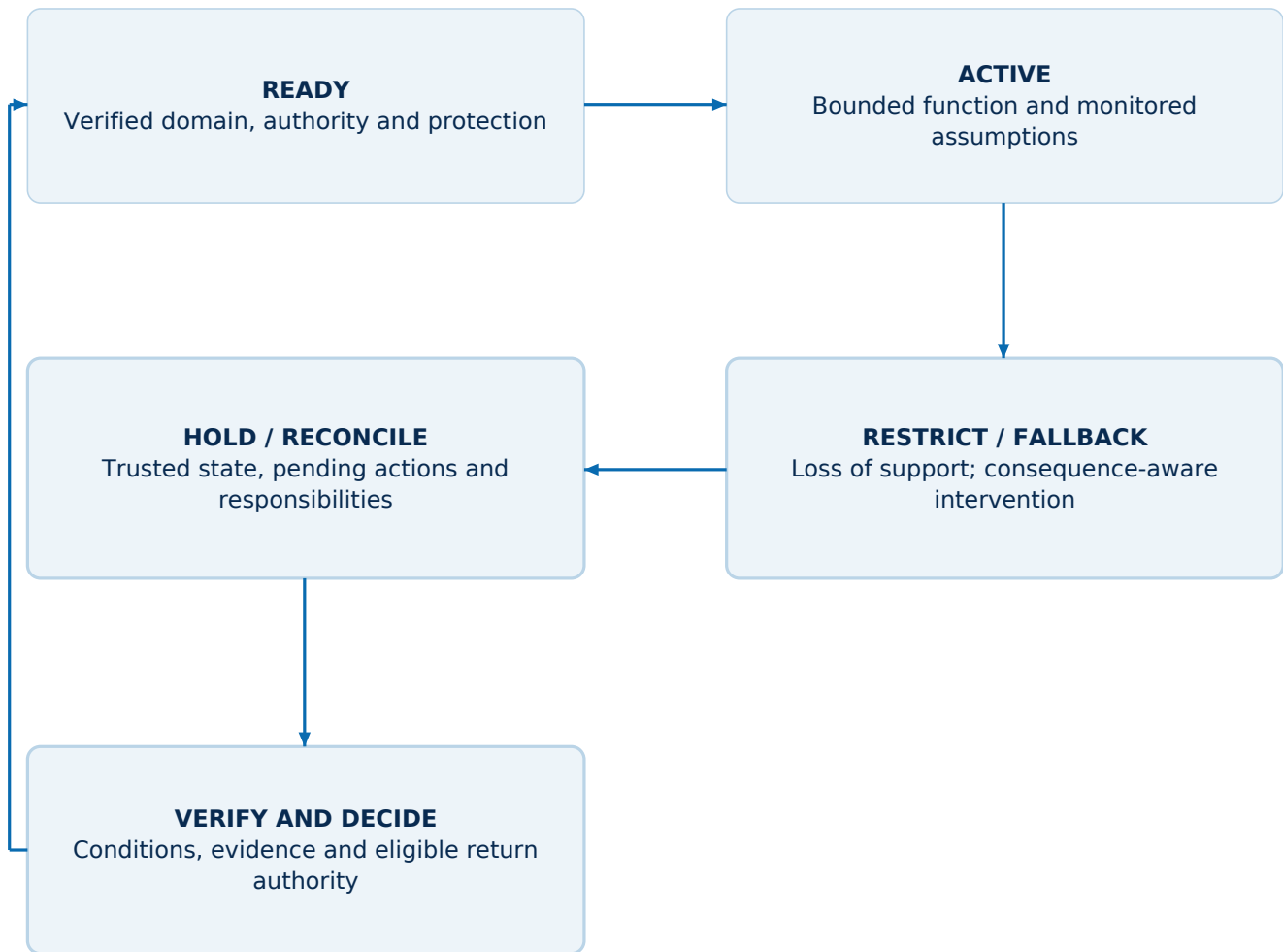
Status: Informative. The diagrams summarize relationships. Apply the complete requirements and local parameters before making a decision.

Q.1 Function-to-operation assurance chain



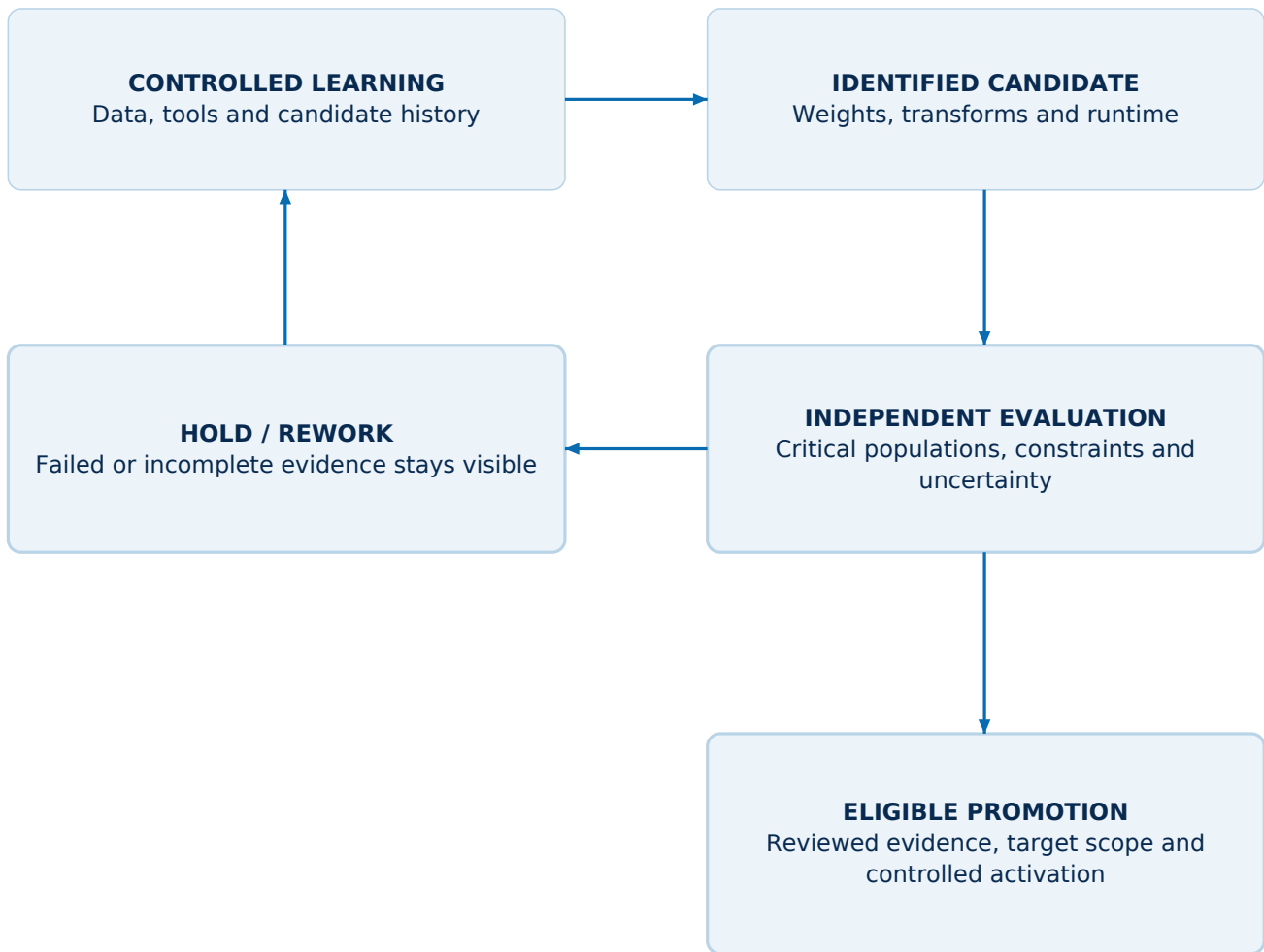
A supported claim connects purpose and domain to the complete implementation and actual operation. Evidence and operational feedback can reopen an earlier decision.

Q.2 Authority and protective transitions



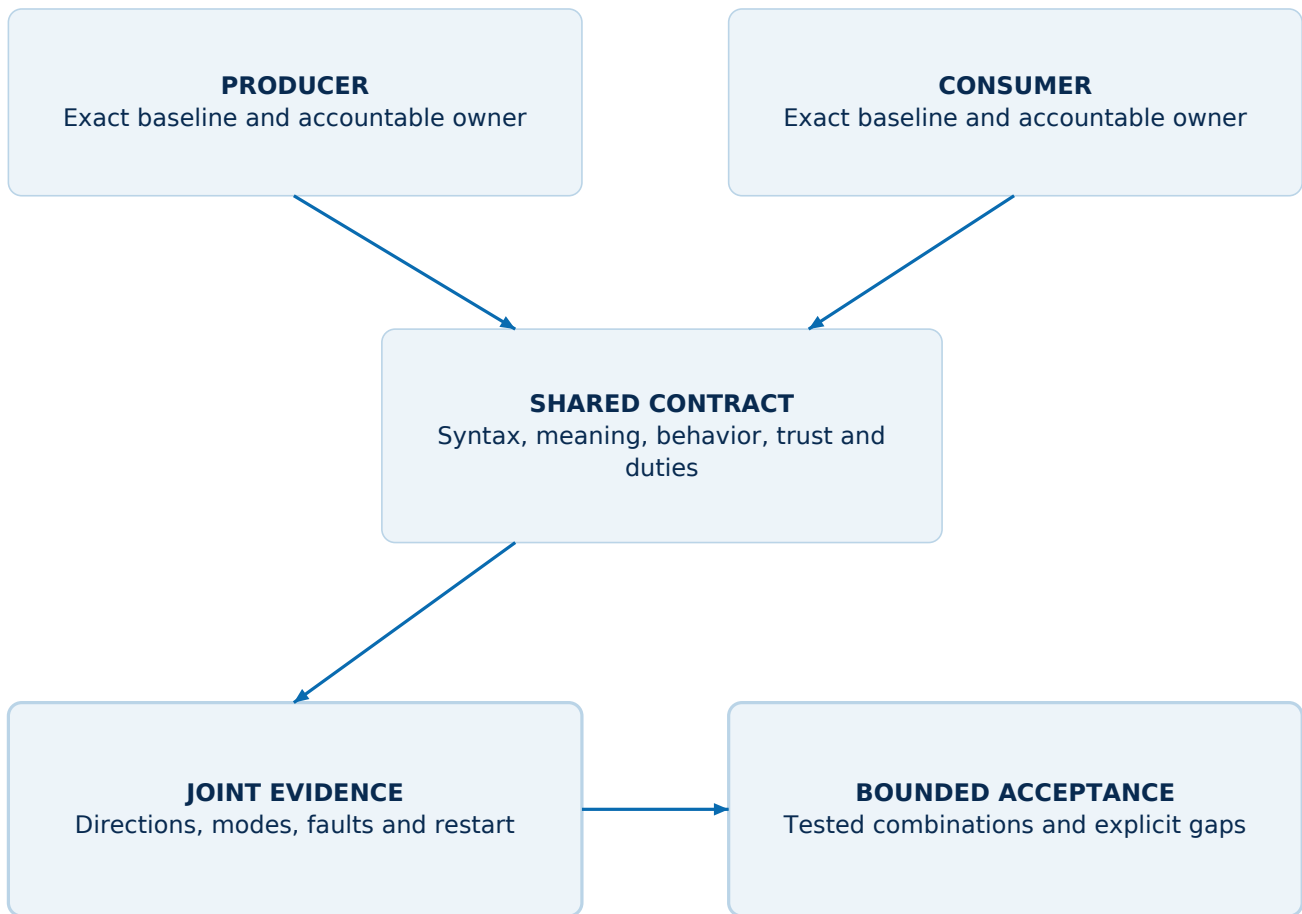
Nominal autonomy is permitted only while authority, domain and protective assumptions hold. Recovery and return require verified conditions and the appropriate decision; a warning alone does not transfer responsibility.

Q.3 Controlled model promotion



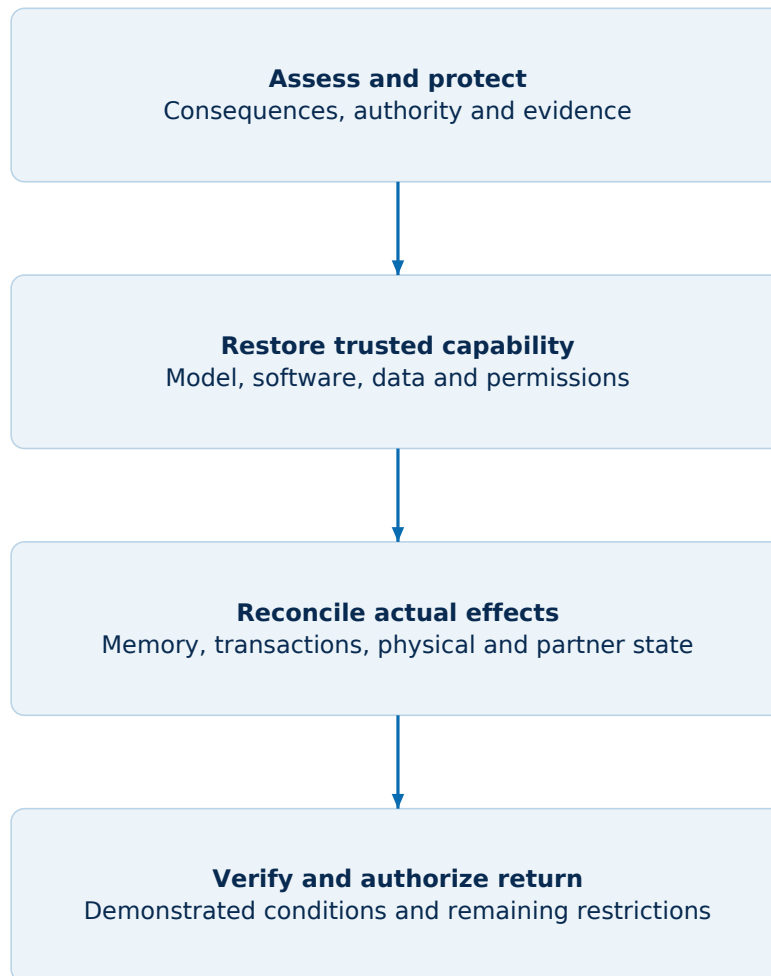
Learning creates a candidate, not a deployment authorization. Evaluation, restrictions and promotion remain separately accountable. Failed or incomplete evidence returns the candidate to controlled development.

Q.4 Interoperability acceptance layers



Transport connectivity, shared meaning, compatible behavior, trust and organizational duties require distinct evidence. Acceptance applies to exact relationships and operating conditions.

Q.5 Incident recovery and state reconciliation



Return requires trusted implementation, consistent information, reconciled external effects and eligible authority. Restoring one technical service does not establish restoration of the complete function.

Annex R - Activity-based implementation routes

Status: Informative. Select routes by actual responsibility. Shared governance, evidence and decision controls apply across all routes.

R.1 Data provider and annotation team

Start with intended functions and domain. Apply 11-13 to origin, permitted use, reference truth, quality, partitioning and representation. Deliver a versioned dossier with explicit gaps and change obligations. Connect data corrections to affected models and decisions.

R.2 Model or AI-service provider

Apply 10-18 and 27-29 to learning, candidate selection, independent evaluation, robustness and exact artifacts. Declare provider-controlled changes and unavailable evidence. Release only bounded claims with actual configuration and intended-use restrictions.

R.3 Autonomous-system integrator

Apply 5-9 and 19-22 to authority, domain, architecture, perception, planning, collective behavior and human roles. Demonstrate the complete loop and fallback timing. Integrate suppliers and interfaces before claiming system behavior.

R.4 Interface and platform provider

Apply 23-27 to contracts, semantics, spatial and temporal meaning, distributed state, trust and configuration. Build a direction-specific compatibility matrix. Test malformed, stale, unknown, conflicting and restarted interactions as well as nominal exchange.

R.5 Aviation integrator and operator

Apply 30 with shared controls and 33-37. Establish phase-specific authority, crew or operator roles, air-ground dependencies, fleet state and required permissions. Use operational feedback to reassess domain and human reliance.

R.6 Space and remote-mission operator

Apply 31 to delayed commands, intermittent links, bounded onboard autonomy and resource-constrained recovery. Verify segment responsibilities and actual mission state. Treat mission extension as a new review of assumptions and support.

R.7 Industrial, maintenance and ground-service operator

Apply 32 to physical tasks, shared spaces and item-linked evidence. Validate changes in sensing, fixtures, materials and workflow. Exercise recovery that reconciles physical work and prior decisions.

R.8 Generative or agentic workflow owner

Apply 18 with 4-10, 22-29 and 33-39. Define permitted content, tool actions and approval points. Validate multi-step effects, untrusted context, partial transactions and recovery. Bound human review by demonstrated capacity.

R.9 Assessor and assurance owner

Apply 4, 10, 15, 28-29 and 39. Follow complete trails from purpose to actual operation, including failures and exclusions. State evidence limitations and blocked decisions. Completed templates are records of work only when supported by actual content.

Annex S - Interpretation questions and boundary examples

Status: Informative. Answers clarify recurring distinctions; they do not waive requirements or assign missing local parameters.

S.1 Does high accuracy establish safe autonomy?

No. Accuracy is one property for a declared population. Authority, domain, timing, failures, protection, human roles and physical effects require separate evidence.

S.2 Is model confidence the probability that an action is safe?

No. Its meaning depends on the model and calibration event. A safety probability requires a justified system-level model and evidence; a confidence score cannot supply that relationship by itself.

S.3 Does a human-in-the-loop label provide adequate oversight?

Only demonstrated information, competence, workload capacity, intervention time and actual control can support that claim. A nominal approval step or alert is insufficient.

S.4 Can a monitor make any controller acceptable?

No. The monitor, observation, timing, authority and fallback must be valid for the specific state and consequences. Common failures and unsupported conditions remain relevant.

S.5 Can a model learn continuously after release?

Only within an explicitly justified adaptation envelope with protected constraints, evidence continuity and controlled active state. Improved recent performance does not independently authorize promotion.

S.6 Does a shared protocol mean systems interoperate?

No. Data meaning, units, time, behavior, trust, optional features and organizational duties must also agree for the exact participants and configuration.

S.7 Does reliable transport ensure a command happens exactly once?

No. Retries, crashes and lost acknowledgments can leave business effects uncertain. Persistent execution identity and actual-state reconciliation must support the claim.

S.8 Can simulation replace all operational evidence?

Only evidence justified for the intended property and validated model scope can be credited. Unrepresented conditions and real-world interactions require additional support or restrictions.

S.9 Is absence of drift proof of continued validity?

No. A detector has limited observations and sensitivity. Delayed outcomes, hidden subgroups and changing relationships can remain undetected.

S.10 Can an agent create its own approval authority?

No. Authority comes from the approved workflow and eligible delegation. Generated content, retrieved instructions or repeated retries cannot expand permissions.

S.11 Does a model rollback restore the system?

Not necessarily. Persistent memory, data schemas, transactions, physical effects, trust and partner states may also need reconciliation and verification.

S.12 Does completing this set confer recognition or operating permission?

No. Publication, local adoption, technical acceptance, recognition and permission to operate are separate decisions by their eligible authorities on the required evidence.

Annex T - Requirement families and implementation crosswalk

Status: Informative. The supporting master list and coverage matrix provide the complete working-document routes.

T.1 Requirement families

Table T-1. Core navigation map

Requirement family	Clauses	Main normative annexes
Scope, language, governance and adoption	0-4, 40	A-B, N
Concept, domain, risk and requirements	5-8	C
Architecture and lifecycle	9-10	F, N
Data and reference evidence	11-13	D
Learning, evaluation and robustness	14-16	E
Adaptation and generative systems	17-18	E
Perception, control and collective behavior	19-21	F
Human oversight and competence	22	G
Interoperability and shared meaning	23-25	H
Trust, suppliers, tools and identity	26-27	I
Models, verification and validation	28-29	J
Aerospace operating contexts	30-32	K
Deployment, monitoring and recovery	33-35	L
Release, continuing use and retirement	36-38	M
Assurance and evidence continuity	39-40	N

T.2 Implementation order

Select a representative end-to-end function. Define purpose, consequences, domain and authority. Complete the parameter baseline, trace requirements through actual data, model and interfaces, then demonstrate nominal and adverse behavior. Exercise fallback and recovery before expanding to further participants or operating conditions.

T.3 Completeness questions

Can the organization identify what the function may do, where it is supported, what actually executes and which evidence applies? Can it explain shared meaning and authority across every relevant interface? Can it detect loss of support, intervene in time and reconcile state after failure? Unresolved answers identify gaps even when individual documents appear complete.