

Sensor AI for Transport and Industrial Safety: From Prototype to Safety Case

An Evidence-Gated Framework for Sensing Conditions, Hardware Reliability, Human Response and Certification Readiness

Authored by Chennakeshav Adya, Independent Researcher

September 2026

Abstract

Sensor AI can prevent collisions, detect unsafe proximity, identify defects, monitor machinery and support autonomous transport, yet a successful prototype does not establish a defensible safety case. Field outcomes depend on sensing conditions, synchronization, perception, fusion, decision logic, actuation, hardware reliability, human response, operating procedures and certification evidence. This paper develops an evidence-gated framework for moving sensor AI from prototype to controlled deployment in transport and industrial safety. Forty modules connect the safety outcome, sensor-to-action chain, operational design domain, hazards, safety functions, integrity targets, sensor diversity, calibration, timing, environmental conditions, ground truth, long-tail sampling, perception, fusion, scenario understanding, uncertainty, consequence thresholds, false negatives, false positives, safe states, hardware reliability, diagnostics, actuation, human factors, foreseeable misuse, cybersecurity, traceability, safety-case construction, simulation, controlled trials, end-to-end outcomes, comparative evidence, operational economics, supplier control, software updates, in-service monitoring, event investigation, revalidation, certification and scale gates. Five figures, five tables, eight frequently asked questions and twenty-six primary or authoritative references support use-, condition-, hazard-, system-, jurisdiction- and period-specific review. The framework requires evidence that the complete control loop produces bounded risk across the approved operating domain and fails safely outside it. It does not certify any product or establish that a system is safe, compliant, investable or suitable for deployment and does not substitute for authorised engineering, functional-safety, transport, machinery, legal, cybersecurity, insurance or regulatory advice.

JEL Classification: L92, L62, L64, C45, O32, K32, G24

Keywords: sensor AI, transport safety, industrial safety, functional safety, SOTIF, safety case, operational design domain, autonomous systems, certification, physical AI

1. Define the safety outcome

State the protected people, assets, hazards, operating domain, action and tolerated residual risk.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an approved safety-outcome charter. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when sensor accuracy is treated as the safety objective. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Sensor AI should be judged on the controlled safety outcome across sensing, interpretation, action, human response and recovery.

The decision pack for define the safety outcome should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

2. Map the sensor-to-action chain

Trace sensing, synchronization, perception, fusion, decision, actuation, human response and recovery.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an end-to-end safety architecture. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when teams validate components without the complete control loop. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Performance depends on the operating domain, exposure, hardware condition, timing, uncertainty and interaction among people, machines and infrastructure.

The decision pack for map the sensor-to-action chain should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

3. Define the operational design domain

Specify geography, weather, lighting, speed, traffic, machinery state, connectivity and human presence.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an operational-domain specification. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when a prototype is exposed to conditions absent from its design basis. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Reviewers need traceable hazard evidence, consequence-based thresholds, credible adverse testing and authority to withhold deployment or require a safe state.

The decision pack for define the operational design domain should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

4. Build the hazard register

Connect credible hazards, initiating events, exposure, severity, controls and owners.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a controlled hazard log. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when rare severe events disappear inside average performance. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Scalable value requires safety integrity, operational availability, maintainability, certification readiness and monitored field performance.

The decision pack for build the hazard register should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

5. Allocate safety functions

Assign detection, decision, warning, intervention, safe state and recovery across hardware, software and people.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a safety-function allocation. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when responsibility falls between subsystems. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Sensor AI should be judged on the controlled safety outcome across sensing, interpretation, action, human response and recovery.

The decision pack for allocate safety functions should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

6. Set integrity targets

Translate risk into required performance, independence, diagnostic coverage and evidence.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an integrity-target matrix. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when commercial targets substitute for safety integrity. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Performance depends on the operating domain, exposure, hardware condition, timing, uncertainty and interaction among people, machines and infrastructure.

The decision pack for set integrity targets should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

7. Engineer sensor diversity

Combine appropriate camera, radar, lidar, acoustic, thermal, inertial and environmental sensing.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a sensor-diversity design. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when correlated blind spots defeat apparent redundancy. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Reviewers need traceable hazard evidence, consequence-based thresholds, credible adverse testing and authority to withhold deployment or require a safe state.

The decision pack for engineer sensor diversity should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

8. Control calibration

Manage intrinsic, extrinsic, temporal and environmental calibration with traceable tolerances.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a calibration-control plan. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when calibration drift silently moves the decision boundary. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Scalable value requires safety integrity, operational availability, maintainability, certification readiness and monitored field performance.

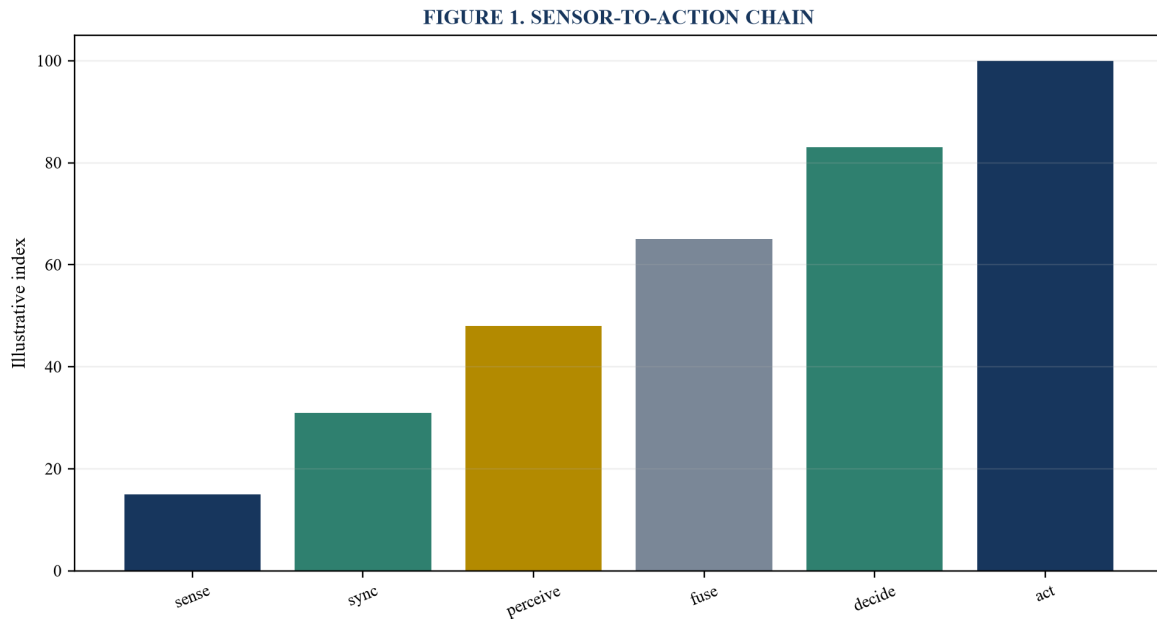
The decision pack for control calibration should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

Table 1. Sensor-to-action assurance layers

Layer	Primary question	Required evidence
sense	what was observed	condition benchmark
interpret	what the scene means	scenario validation
decide	which response is required	hazard logic
act	did the control respond	end-to-end test

Illustrative controls require use-, consequence-, jurisdiction-, firm- and period-specific approval.

Figure 1. Sensor-to-action chain



Values are illustrative indices and require replacement with approved validation evidence.

9. Synchronise time and position

Bound timestamp, latency, jitter, clock drift, localization and coordinate transforms.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a synchronization assurance pack. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when correct observations are fused into the wrong scene. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Sensor AI should be judged on the controlled safety outcome across sensing, interpretation, action, human response and recovery.

The decision pack for synchronise time and position should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

10. Characterise sensing conditions

Measure range, angle, reflectivity, occlusion, glare, precipitation, dust, vibration, electromagnetic interference and temperature.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an operating-condition envelope. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when laboratory accuracy masks field degradation. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Performance depends on the operating domain, exposure, hardware condition, timing, uncertainty and interaction among people, machines and infrastructure.

The decision pack for characterise sensing conditions should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

11. Construct ground truth

Create independently reviewed event, object, state and action labels with provenance and uncertainty.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a governed reference corpus. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when developer labels become unquestioned truth. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Reviewers need traceable hazard evidence, consequence-based thresholds, credible adverse testing and authority to withhold deployment or require a safe state.

The decision pack for construct ground truth should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

12. Sample the long tail

Stratify normal, boundary, degraded, misuse, near-miss and emergency conditions.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a risk-weighted sampling plan. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when a clean test set omits the events that define risk. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Scalable value requires safety integrity, operational availability, maintainability, certification readiness and monitored field performance.

The decision pack for sample the long tail should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

13. Evaluate perception

Measure detection, classification, tracking, segmentation, distance and velocity by condition.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a segmented perception scorecard. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when one aggregate metric hides safety-critical misses. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Sensor AI should be judged on the controlled safety outcome across sensing, interpretation, action, human response and recovery.

The decision pack for evaluate perception should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

14. Evaluate sensor fusion

Test association, conflict resolution, uncertainty propagation and graceful degradation.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a fusion-validation report. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when fusion amplifies a confident faulty source. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Performance depends on the operating domain, exposure, hardware condition, timing, uncertainty and interaction among people, machines and infrastructure.

The decision pack for evaluate sensor fusion should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

15. Test scenario understanding

Assess intent, trajectory, right of way, exclusion zones, machine states and evolving context.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a scenario-understanding benchmark. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when objects are detected while the hazardous situation is misunderstood. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Reviewers need traceable hazard evidence, consequence-based thresholds, credible adverse testing and authority to withhold deployment or require a safe state.

The decision pack for test scenario understanding should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

16. Calibrate uncertainty

Relate model and sensor confidence to observed error, novelty and consequence.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a calibrated uncertainty model. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when confidence scores are mistaken for probabilities of safety. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Scalable value requires safety integrity, operational availability, maintainability, certification readiness and monitored field performance.

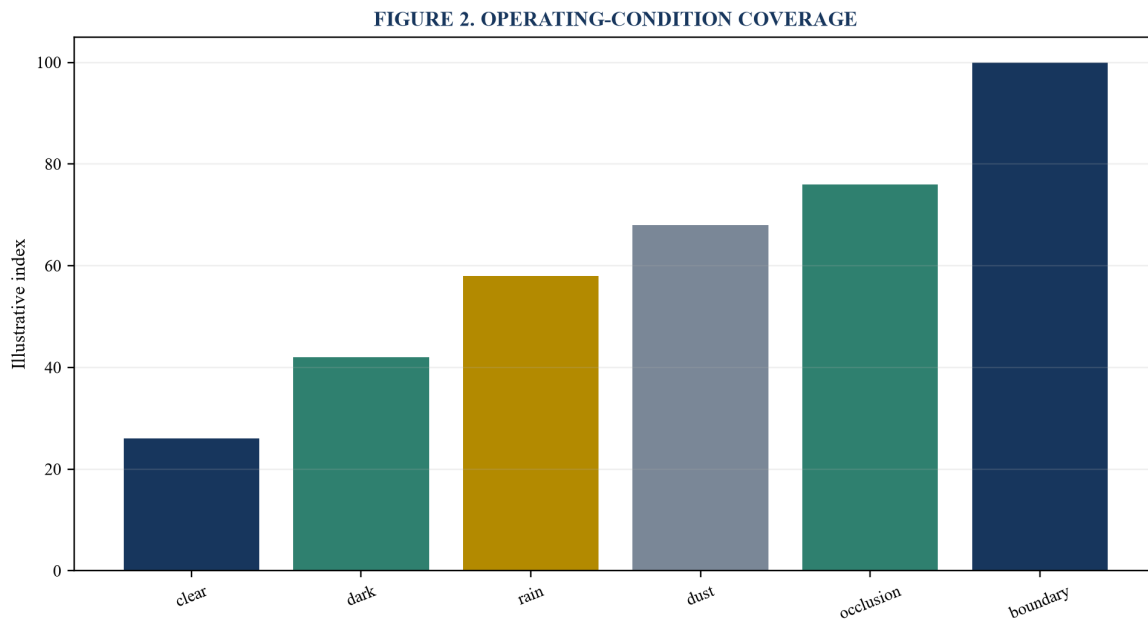
The decision pack for calibrate uncertainty should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

Table 2. Operating-condition test matrix

Dimension	Adverse condition	Decision evidence
environment	rain, dust, glare and darkness	robustness curve
geometry	range, angle and occlusion	coverage boundary
system	latency, loss and drift	degraded-mode result
human	attention, override and misuse	response study

Illustrative controls require use-, consequence-, jurisdiction-, firm- and period-specific approval.

Figure 2. Operating-condition coverage



Values are illustrative indices and require replacement with approved validation evidence.

17. Set consequence thresholds

Vary warning, review, intervention and safe-state thresholds by hazard and operating context.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a consequence-threshold matrix. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when one threshold either misses hazards or stops operations excessively. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Sensor AI should be judged on the controlled safety outcome across sensing, interpretation, action, human response and recovery.

The decision pack for set consequence thresholds should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

18. Control false negatives

Measure missed people, vehicles, obstacles, defects and unsafe states under adverse conditions.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a miss-risk assessment. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when headline precision conceals dangerous omissions. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Performance depends on the operating domain, exposure, hardware condition, timing, uncertainty and interaction among people, machines and infrastructure.

The decision pack for control false negatives should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

19. Control false positives

Measure nuisance alarms, unnecessary braking, production stops and alarm fatigue.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a false-intervention analysis. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when protective automation becomes operationally unusable. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Reviewers need traceable hazard evidence, consequence-based thresholds, credible adverse testing and authority to withhold deployment or require a safe state.

The decision pack for control false positives should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

20. Design the safe state

Specify minimal-risk manoeuvres, controlled stops, isolation, degraded modes and restart authority.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a safe-state specification. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when failure leaves no bounded recovery path. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Scalable value requires safety integrity, operational availability, maintainability, certification readiness and monitored field performance.

The decision pack for design the safe state should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

21. Engineer hardware reliability

Assess sensor failure rates, power, connectors, contamination, thermal stress, vibration and aging.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a hardware reliability case. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when AI performance is evaluated on unreliable physical inputs. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Sensor AI should be judged on the controlled safety outcome across sensing, interpretation, action, human response and recovery.

The decision pack for engineer hardware reliability should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

22. Design diagnostic coverage

Detect obstruction, misalignment, frozen output, implausibility, data loss and latent failure.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a diagnostic-coverage model. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when failed sensors continue to appear healthy. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Performance depends on the operating domain, exposure, hardware condition, timing, uncertainty and interaction among people, machines and infrastructure.

The decision pack for design diagnostic coverage should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

23. Validate actuation

Test brake, steering, shutdown, interlock, alarm and control-command latency and authority.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an actuation assurance report. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when correct perception produces late or ineffective intervention. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Reviewers need traceable hazard evidence, consequence-based thresholds, credible adverse testing and authority to withhold deployment or require a safe state.

The decision pack for validate actuation should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

24. Integrate human response

Define attention, comprehension, reaction time, escalation, override and remote-operation roles.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a human-factors control design. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when the safety case assumes ideal human recovery. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Scalable value requires safety integrity, operational availability, maintainability, certification readiness and monitored field performance.

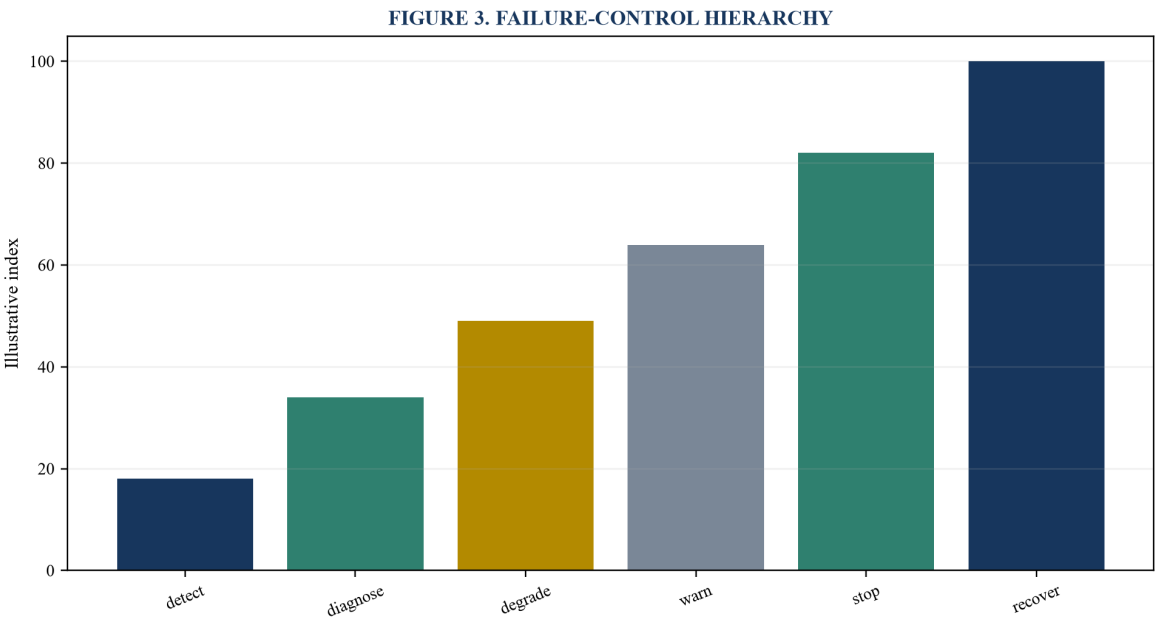
The decision pack for integrate human response should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

Table 3. Safety-function allocation

Function	Primary control	Completion evidence
detect	diverse sensors and diagnostics	coverage results
decide	hazard and threshold logic	validated response
intervene	actuation and safe state	timing evidence
recover	human and system authority	approved procedure

Illustrative controls require use-, consequence-, jurisdiction-, firm- and period-specific approval.

Figure 3. Failure-control hierarchy



Values are illustrative indices and require replacement with approved validation evidence.

25. Control foreseeable misuse

Test bypass, masking, poor placement, ignored alarms, unsafe overrides and workarounds.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a misuse-resistance assessment. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when field behaviour invalidates controlled assumptions. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Sensor AI should be judged on the controlled safety outcome across sensing, interpretation, action, human response and recovery.

The decision pack for control foreseeable misuse should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

26. Test cybersecurity effects

Probe spoofing, jamming, replay, tampering, model manipulation, access and update compromise.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a cyber-safety test pack. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when security failure disables or corrupts a safety function. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Performance depends on the operating domain, exposure, hardware condition, timing, uncertainty and interaction among people, machines and infrastructure.

The decision pack for test cybersecurity effects should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

27. Build traceable evidence

Link requirement, hazard, design, dataset, model, hardware, test, result, exception and approval.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an immutable assurance trail. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when claims cannot be reconstructed for certification or incident review. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Reviewers need traceable hazard evidence, consequence-based thresholds, credible adverse testing and authority to withhold deployment or require a safe state.

The decision pack for build traceable evidence should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

28. Construct the safety case

Organise claims, arguments, evidence, assumptions, limitations and unresolved risks.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a structured safety-case dossier. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when documentation accumulates without proving the required outcome. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Scalable value requires safety integrity, operational availability, maintainability, certification readiness and monitored field performance.

The decision pack for construct the safety case should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

29. Validate in simulation

Use scenario libraries, perturbations, sensor models and fault injection within stated fidelity limits.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a simulation-validity report. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when simulation volume is mistaken for representative evidence. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Sensor AI should be judged on the controlled safety outcome across sensing, interpretation, action, human response and recovery.

The decision pack for validate in simulation should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

30. Validate on controlled sites

Progress through bench, hardware-in-loop, proving-ground and supervised operational trials.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a staged trial record. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when public deployment begins before failure modes are bounded. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Performance depends on the operating domain, exposure, hardware condition, timing, uncertainty and interaction among people, machines and infrastructure.

The decision pack for validate on controlled sites should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

31. Measure end-to-end outcomes

Score collisions, hazardous proximity, interventions, safe stops, near misses and exposure.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an outcome-based validation report. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when component gains fail to improve system safety. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Reviewers need traceable hazard evidence, consequence-based thresholds, credible adverse testing and authority to withhold deployment or require a safe state.

The decision pack for measure end-to-end outcomes should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

32. Demonstrate comparative safety

Use exposure-adjusted baselines, confidence intervals and transparent exclusions.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a comparative evidence model. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when small samples support exaggerated safety claims. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

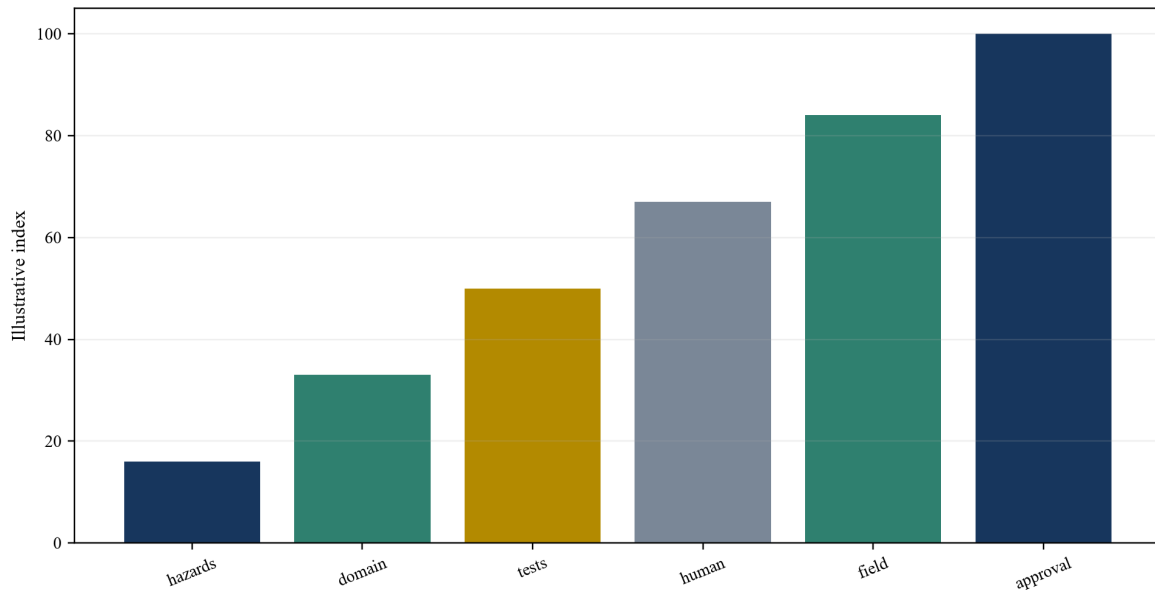
Scalable value requires safety integrity, operational availability, maintainability, certification readiness and monitored field performance.

The decision pack for demonstrate comparative safety should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

Table 4. Safety-case evidence

Claim	Test	Required record
domain	bounded operating conditions	domain specification
integrity	fault and uncertainty control	assurance results
outcome	exposure-adjusted safety	field evidence
lifecycle	change and monitoring	accountable approval

Illustrative controls require use-, consequence-, jurisdiction-, firm- and period-specific approval.

Figure 4. Safety-case maturity**FIGURE 4. SAFETY-CASE MATURITY**

Values are illustrative indices and require replacement with approved validation evidence.

33. Quantify operational economics

Connect availability, nuisance actions, maintenance, downtime, insurance, incidents and productivity.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a safety-adjusted value model. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when the business case excludes the cost of safety controls. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Sensor AI should be judged on the controlled safety outcome across sensing, interpretation, action, human response and recovery.

The decision pack for quantify operational economics should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

34. Control suppliers and versions

Assess component provenance, model releases, firmware, dependencies, change notices, audit and exit.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a supplier-assurance file. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when an upstream change invalidates certified behaviour. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Performance depends on the operating domain, exposure, hardware condition, timing, uncertainty and interaction among people, machines and infrastructure.

The decision pack for control suppliers and versions should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

35. Manage software updates

Require impact analysis, regression testing, rollback, configuration control and release authority.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a controlled update process. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when continuous delivery bypasses safety approval. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Reviewers need traceable hazard evidence, consequence-based thresholds, credible adverse testing and authority to withhold deployment or require a safe state.

The decision pack for manage software updates should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

36. Monitor in service

Track conditions, exposure, errors, interventions, overrides, near misses and incidents.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an in-service monitoring dashboard. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when field performance drifts beyond the approved envelope. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Scalable value requires safety integrity, operational availability, maintainability, certification readiness and monitored field performance.

The decision pack for monitor in service should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

37. Investigate events

Preserve synchronized sensor, model, decision, command and human-action records.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an event-investigation protocol. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when incident evidence is incomplete or overwritten. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Sensor AI should be judged on the controlled safety outcome across sensing, interpretation, action, human response and recovery.

The decision pack for investigate events should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

38. Trigger revalidation

Require review after material domain, sensor, model, hardware, software, workflow or regulatory change.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a revalidation register. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when changed systems inherit obsolete evidence. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Performance depends on the operating domain, exposure, hardware condition, timing, uncertainty and interaction among people, machines and infrastructure.

The decision pack for trigger revalidation should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

39. Plan certification and market access

Map applicable transport, machinery, product, AI, cybersecurity and sector approvals by jurisdiction.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is a certification roadmap. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when technical readiness is confused with legal market access. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Reviewers need traceable hazard evidence, consequence-based thresholds, credible adverse testing and authority to withhold deployment or require a safe state.

The decision pack for plan certification and market access should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

40. Close through control gates

Require bounded domain, controlled hazards, reliable sensing, validated action, human factors, safety-case evidence and monitoring.

The controlled record includes scope; authority; source; passage; fact; calculation; reviewer; decision. The immediate deliverable is an accountable scale decision. Preserve jurisdiction, effective date, version, transformation, reviewer, approval and unresolved exceptions.

The principal failure occurs when a successful prototype is mistaken for a scalable safety system. Reviewers should reconstruct the evidence path, test consequence under adverse conditions and identify who may approve or withhold the recommendation.

Scalable value requires safety integrity, operational availability, maintainability, certification readiness and monitored field performance.

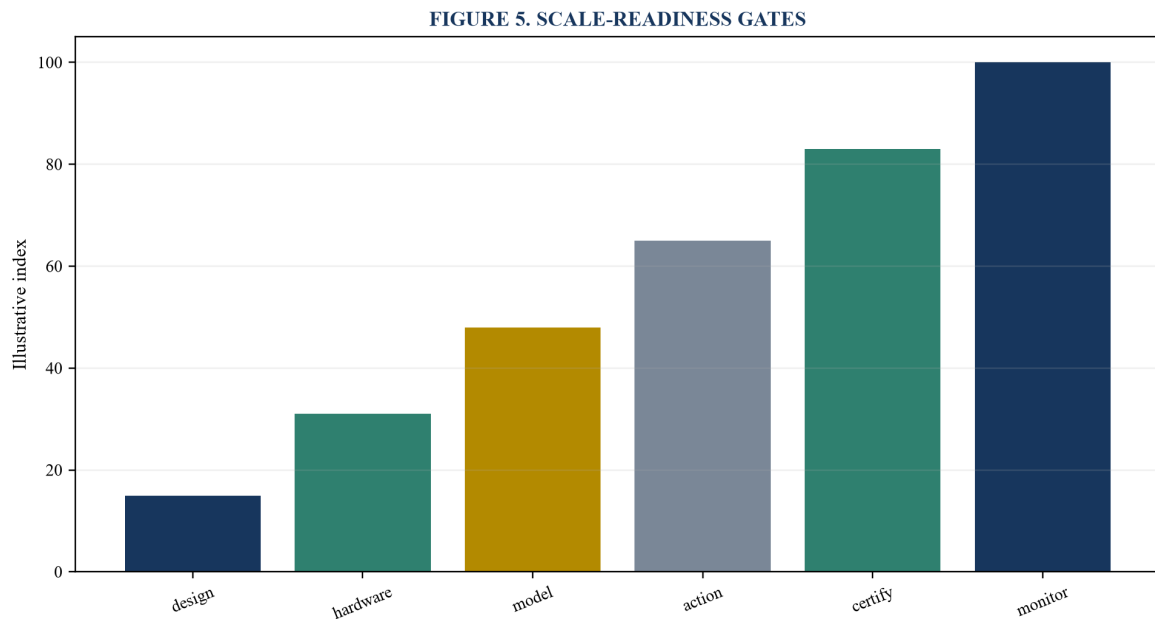
The decision pack for close through control gates should show the prior view, correction, consequence at stake, alternatives, owner, due date, next evidence gate and observed outcome. Exceptions require closure evidence and an explicit effect on retrieval, citation, abstention, review or release.

Table 5. Scale-deployment gates

Gate	Primary work	Completion evidence
hazard	risk and safety functions	approved hazard log
system	hardware, AI and actuation	integrated test
operation	human factors and recovery	trial record
market	certification and monitoring	release decision

Illustrative controls require use-, consequence-, jurisdiction-, firm- and period-specific approval.

Figure 5. Scale-readiness gates



Values are illustrative indices and require replacement with approved validation evidence.

References

- [1] International Maritime Organization, International Code of Safety for Maritime Autonomous Surface Ships, <https://www.imo.org/en/mediacentre/hottopics/pages/autonomous-shipping.aspx>
- [2] International Maritime Organization, Adoption of the MASS Code, <https://www.imo.org/en/mediacentre/pressbriefings/pages/imo-adopts-mass-code.aspx>
- [3] International Maritime Organization, International Convention for the Safety of Life at Sea, [https://www.imo.org/en/about/conventions/pages/international-convention-for-the-safety-of-life-at-sea-\(solas\)-1974.aspx](https://www.imo.org/en/about/conventions/pages/international-convention-for-the-safety-of-life-at-sea-(solas)-1974.aspx)
- [4] United Nations Economic Commission for Europe, Regulation on Automated Driving Systems, <https://unece.org/sustainable-development/press/unece-adopts-first-ever-global-rules-allowing-fully-autonomous>
- [5] United Nations Economic Commission for Europe, ADS Safety Requirements Assessments and Test Methods, <https://unece.org/info/Transport/Vehicle-Regulations/pub/402564>
- [6] United Nations Economic Commission for Europe, Framework Document for Automated and Autonomous Vehicles, <https://unece.org/info/Transport/pub/365097>
- [7] United Nations Economic Commission for Europe, Working Party on Automated Autonomous and Connected Vehicles, <https://unece.org/transport/road-transport/working-party-automatedautonomous-and-connected-vehicles-introduction>
- [8] International Organization for Standardization, ISO 26262 Road Vehicles Functional Safety, <https://www.iso.org/standard/68383.html>
- [9] International Organization for Standardization, ISO 21448 Safety of the Intended Functionality, <https://www.iso.org/standard/77490.html>
- [10] International Organization for Standardization, ISO 13849-1 Safety-Related Parts of Control Systems, <https://www.iso.org/standard/73481.html>
- [11] International Organization for Standardization, ISO 12100 Machinery Risk Assessment and Risk Reduction, <https://www.iso.org/standard/51528.html>
- [12] International Electrotechnical Commission, IEC 61508 Functional Safety, <https://www.iec.ch/functional-safety>
- [13] European Union, Regulation 2023/1230 on Machinery, <https://eur-lex.europa.eu/eli/reg/2023/1230/oj>

- [14] European Union, Regulation 2024/1689 Artificial Intelligence Act,
<https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- [15] European Union Agency for Cybersecurity, Cybersecurity of AI and Standardisation,
<https://www.enisa.europa.eu/publications/cybersecurity-of-ai-and-standardisation>
- [16] National Institute of Standards and Technology, AI Risk Management Framework 1.0,
<https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10>
- [17] National Institute of Standards and Technology, AI Resource Center Testing Evaluation Verification and Validation, <https://airc.nist.gov/>
- [18] National Institute of Standards and Technology, Cybersecurity Framework 2.0,
<https://www.nist.gov/cyberframework>
- [19] International Organization for Standardization, ISO IEC 42001 AI Management Systems,
<https://www.iso.org/standard/81230.html>
- [20] International Organization for Standardization, ISO IEC 23894 AI Risk Management,
<https://www.iso.org/standard/77304.html>
- [21] International Organization for Standardization, ISO 31000 Risk Management,
<https://www.iso.org/iso-31000-risk-management.html>
- [22] International Organization for Standardization, ISO 45001 Occupational Health and Safety,
<https://www.iso.org/standard/63787.html>
- [23] United States Department of Transportation, Automated Vehicles Comprehensive Plan,
<https://www.transportation.gov/av/avcp>
- [24] United States National Highway Traffic Safety Administration, Automated Vehicles for Safety,
<https://www.nhtsa.gov/vehicle-safety/automated-vehicles-safety>
- [25] European Union Aviation Safety Agency, Artificial Intelligence Roadmap 2.0,
<https://www.easa.europa.eu/en/document-library/general-publications/artificial-intelligence-roadmap-20>
- [26] International Civil Aviation Organization, Safety Management,
<https://www.icao.int/safety/safetymanagement/pages/default.aspx>

About the Author

Authored by Chennakeshav Adya, Independent Researcher

Chennakeshav (CK) is a corporate finance and investment banking executive with 25+ years of global experience in deal origination, structuring and execution across M&A, growth capital and corporate strategy. He has led value-creation mandates for founders, corporates and funds — bridging the boardroom view to hands-on execution and close.

His career spans Morgan Stanley, HSBC, Lloyds Banking Group, EWEC, ADQ portfolio companies and Emirates Growth Fund, across TMT, real estate, fintech, deeptech, cleantech, infrastructure and energy. He has partnered with C-suite leaders, private equity and venture funds, sovereign wealth funds and family offices to finance complex fund raises and scale-up ventures, and has led M&A due diligence, post-merger integration and business-transformation initiatives to create value.

At Matchpoint Partners he is Managing Partner, leading the firm's corporate finance, M&A and capital-raising practice. He holds an MBA from London Business School, an engineering degree from VTU and a Master of Laws (LLM, in progress) from UCL London.

An active start-up mentor, CK mentors at Techstars, DIFC FinTech Hive, Startup Grind, Founder Institute and IN5, serves as Entrepreneur Mentor in Residence (EMiR) at London Business School, and judges the Entrepreneurship World Cup.