



WEKID™ OPEN STANDARD

Sector Calibration and Authority Alignment Protocol

The governed process for adapting epistemic maturity and AI decision authority by sector

WEKID-STD-002 | Version 1.0

Ratified - Normative Standard

James Madigan Judge and Members of the WEKID Community

September 2026

Status. WEKID-STD-002 v1.0 is the ratified normative companion standard to WEKID-STD-001 v1.1. It governs the development, validation, ratification, publication, and maintenance of sector calibration profiles.

Document Control

Field	Value
Standard identifier	WEKID-STD-002
Title	Sector Calibration and Authority Alignment Protocol
Version	1.0
Status	Ratified - Normative
Normative baseline	WEKID-STD-001 v1.1
Publication class	Normative Open Standard
Steward	WEKID LLC / WEKID Standards Council
Approved change record	WK-CAL-RFC; registry identifier assigned by the RFC record
Date	September 2026

Version Control Log

Version	Date	Disposition	Summary
1.0	September 2026	Ratified	Initial normative release establishing the thirteen-phase calibration lifecycle, protected invariants, participation, evidence sufficiency, constrained expected-loss analysis, validation, evaluator assurance, public comment, ratification, and continuous recalibration.

Governing Authority

This Standard operationalizes WEKID-STD-001 Sections 8.3, 14, and 15.3. If this Standard conflicts with WEKID-STD-001, the current ratified WEKID-STD-001 provision governs unless a later ratified Standard or RFC expressly supersedes it. The WEKID Master Whitepaper and Engine Whitepaper are informative and do not establish or modify conformance requirements.

Contents

Section	Page
Document Control and Version Log	2
Contents	3
Normative Integration Provisions	4-6
Part A - Calibration Scope and Invariants	7
Part B - Sector Working Group	8
Part C - Thirteen-Phase Calibration Process	9-12
Part D - Profile Assurance Requirements	13
Part E - Governance, Versioning and Recalibration	14
Appendix - Quick Reference	15-16

Normative Integration Provisions

These provisions form part of WEKID-STD-002 v1.0. They clarify and strengthen the thirteen-phase protocol reproduced in the following parts. Where wording differs, these provisions govern.

1. Normative Terms and Scope

SHALL and SHALL NOT identify mandatory requirements. SHOULD and SHOULD NOT identify recommended practices for which a deviation requires documented justification and risk assessment. MAY identifies a permitted option. Phase objectives, mandatory inputs, required outputs, evidence standards, exit gates, invariants, version-pinning requirements, and ratification controls are normative.

2. Precedence

Where sources conflict, the following order SHALL apply: applicable law and binding regulation; WEKID-STD-001; a later ratified Standard or RFC that expressly modifies it; this Standard; the ratified sector profile; conforming local policy; and informative guidance. A lower-precedence source SHALL NOT weaken a higher-precedence requirement.

3. Protected Core

A profile MAY rebalance weights, raise thresholds, tighten or add gates, define decision classes, set monitoring requirements, and reserve decisions to AL-4. It SHALL NOT remove a required layer, permit substitution across the hierarchy, reverse Trust Bridge ordering, relax a core hard gate, lower a stakes floor, weaken fixed precedence, authorize AI at AL-4, or promote an action above the maturity ceiling.

4. Constrained Calibration Objective

Expected-loss optimization SHALL operate only within applicable legal, rights-based, safety, and human-dignity constraints. It SHALL NOT trade away a protected right, legal prohibition, categorical safety requirement, or AL-4 reservation because aggregate expected benefit is greater.

5. Working-Group Legitimacy

A sector working group SHALL include domain practitioners, a regulator or standards liaison, a risk or assurance auditor, an evaluation lead, an ethics or civil-society representative, a Standards Council liaison, and representation reasonably capable of speaking to materially affected people. Its charter SHALL document selection, voting rights, conflicts, recusals, quorum, and formal dissent.

For AL-4 reservations involving binding law or material rights, the regulator or legal liaison and the ethics, civil-society or affected-party representative SHALL record concurrence or a formal dissent for Standards Council adjudication.

6. Evidence Sufficiency

The corpus SHALL cover material decision classes and document source quality, representativeness, severity, frequency, recency, disputed facts, and known gaps. Numeric minimums MAY support planning but SHALL NOT alone establish sufficiency.

7. Validation

Validation SHALL include complete detection of categorical hard-gate and AL-4 cases, defined thresholds for other failure classes, an independently held-out sample, measurement of over-restriction, and documented residual risk. Failure to satisfy a categorical protection SHALL block ratification.

8. Reproducibility and Evaluator Assurance

A profile SHALL distinguish evaluator reproducibility from governance determinism. A probabilistic evaluator SHALL report material variability, uncertainty, and validation limits. Given accepted scores, gates, profile version, policy state, and decision context, the authority-resolution function SHALL produce the same Authority Level.

9. Public Comment and Dissent

Every material comment SHALL receive a documented disposition and reason. Unresolved material objections SHALL be presented to the Standards Council with the working group response and any minority report. Completion requires transparent adjudication, not unanimity.

10. Monitoring and Recalibration

A material new incident, decision class, regulatory requirement, or evidence source SHALL reopen calibration at Phase 1. A change limited to interpretation, weighting, or thresholds MAY reopen at Phase 2 only when the existing corpus remains sufficient. Each recalibration SHALL create a new version and preserve the prior profile and linked verdict history.

11. Formal and Local Profiles

A profile represented as a ratified WEKID sector profile SHALL complete all thirteen phases and receive Standards Council ratification. A locally governed profile MAY use an equivalent process when equivalence is documented, but it SHALL NOT be represented as formally ratified by WEKID.

12. Relationship to Evaluator Assurance

Validation of a calibration profile SHALL NOT substitute for validation of the evaluator applying it. A sound profile applied by an unvalidated evaluator does not establish conformance. The evaluator identity and assurance version SHOULD be pinned to each consequential verdict.

13. Initial Validation Pilot

Before the first sector profile is represented as fully ratified under this Standard, the complete calibration lifecycle SHOULD be exercised through an independently reviewed pilot. Findings SHALL be recorded and SHALL trigger an RFC if they reveal a material defect in this Standard.

Normative decision

Context may change how WEKID is applied. Context SHALL NOT weaken the protections required to make the resulting authority decision defensible.

Part A — What calibration is (and what it must never touch)

Calibration makes WEKID industry specific without weakening it. A sector working group re-balances how a compliant output is scored so the model rewards the epistemic qualities that actually matter in that domain, and it aligns the delegation of authority to the sector's stakes. It does this against two objects and inside five invariants.

The two calibration objects

- The maturity profile — the five layer weights (Data, Information, Knowledge, Experience, Wisdom) plus the tunable gate and threshold settings that cap or constrain a score.
- The authority alignment — the maturity-to-Authority-Level bands, the stakes-floor factors, and the decision classes the sector reserves to a human (AL-4).

The calibration objective

Threshold choices are not merely documented; they are derived. For each decision class the group expresses expected loss as a function of demonstrated maturity, decision stakes, reversibility, and overseer engagement, and chooses the Authority Level that minimizes expected loss. A regulator can then be answered not only “what level, under what calibration?” but “under what objective, and why that threshold?”

The five invariants — no working group may change these

- Weights tune the reward surface; they never override a hard gate or an AL-4 reservation.
- Trust Bridge ordering: maturity is assessed first, authority second — never the reverse.
- Two-determinant conservatism: the assigned level is the more conservative of the maturity ceiling and the stakes floor.
- AL-4 (Human Authority) decision classes stay reserved to a human even at a perfect maturity score.
- Every ratified profile is versioned and pinned to every verdict it produces.

Part B — The sector working group

Calibration is a governed, evidence-based activity carried out by a balanced group, not a solo tuning exercise. The group's legitimacy comes from who is in the room and from the discipline of the process below.

Composition (7-12 members)

Seat	Why it is on the group	Draw from
Chair (domain leader)	Owens the profile end-to-end (Accountable).	Sector professional body / senior practitioner
Domain practitioners ×2–3	Know how the sector's decisions actually fail.	Front-line experts, sector institutions
Regulator / standards liaison	Keeps the profile defensible and aligned to mandates.	Sector regulator or standards body
Risk / assurance auditor	Owens gates, back-test, and audit trail.	Assurance/audit body (e.g., For Humanity, internal audit)
Data-science / evaluation lead	Turns signals into scores; runs the math.	Eval/standards body (e.g., MLCommons, NIST)
Ethics / civil-society seat	Guards the stakes floor and public legitimacy.	Accountability/civil-society org
Standards Council liaison	Shepherds the RFC and ratification path.	WEKID Standards Council
Scribe / editor	Keeps the evidence log and versioning.	Working-group support

Operating rules. Decision rule: consensus, with unresolved items escalated to the Standards Council. Quorum: ≥ 60% of voting members including the chair or the Council liaison. Conflicts of interest are declared at charter and per agenda item. Target cadence: a v1.0 profile in 8–12 weeks.

RACI key. A = accountable (one owner) R = responsible (does the work) C = consulted = informed. Each phase below names its owner(s).

Part C — The calibration process, step by step

Thirteen phases run in order. Each phase has an explicit exit gate that must be met before the next begins; the whole sequence is re-entered from Phase 2 whenever the profile is recalibrated. Each phase produces a named artifact that lives in the companion workbook, so the process and the evidence stay together.

Phase 0 • Charter & convene the working group

Objective	Stand up an accountable, balanced group with a clear mandate to produce one versioned sector profile.
Inputs	Sector scope; the WEKID outreach list (candidate member organizations); Standards Council appointment of a liaison.
Key activities	Recruit 7–12 members across the required seats; adopt the charter; declare conflicts of interest; set cadence and a target ratification date.
Method & rigor	Balanced composition — domain practitioners plus a regulator/standards liaison, a risk/assurance auditor, a data-science/evaluation lead, and an ethics/civil-society seat; consensus decision rule with Standards-Council escalation; quorum ≥ 60%.
Evidence standard	Signed charter; declared COIs.
Output artifact	Charter (Workbook tab 1).
Exit gate	Standards Council accepts the charter and chair.
Owner (RACI)	Chair (A); Council liaison (A/C).

Phase 1 • Assemble the sector evidence corpus

Objective	Build the documented, sector-specific failure record that the whole calibration is derived from and tested against.
Inputs	The WEKID public failure library filtered to this sector; sector incident reports, regulator actions, court and inquiry findings, post-mortems.
Key activities	Collect ≥ 15–20 cases; for each, record what failed, the WEKID layer(s) implicated, any hard gate tripped, the real-world outcome, and the authority that was (wrongly) exercised.
Method & rigor	Every case tagged to Data / Information / Knowledge / Experience / Wisdom and to a decision-authority outcome; sources cited; disputed facts flagged rather than asserted.
Evidence standard	Every case is traceable to a primary or reputable secondary source — no anecdotes without a citation.
Output artifact	Case-evidence log (seeds Workbook tab 5).
Exit gate	Auditor confirms the corpus covers the sector's material decision classes.
Owner (RACI)	Auditor (R); practitioners (R); scribe (R).

Phase 2 • Establish the baseline & failure-mode hypothesis

Objective	State, before touching numbers, where this sector's AI decisions actually fail — the hypothesis the weights will encode.
Inputs	The corpus (Phase 1); the global baseline weights (25/20/20/15/20); the illustrative sector profile (reference only).
Key activities	Cluster the corpus by dominant failure layer; write a one-paragraph, falsifiable hypothesis (e.g., 'clinical decisions fail at Data and Wisdom — fluent but wrong, or confident without uncertainty').
Method & rigor	The hypothesis must be falsifiable and tied to counts of corpus cases per layer, not to intuition.
Evidence standard	Each claimed emphasis backed by a count of corpus cases.
Output artifact	Failure-mode hypothesis memo.
Exit gate	Working-group consensus on the hypothesis.
Owner (RACI)	Chair (A); practitioners (R).

Phase 3 • Calibrate the layer weights

Objective	Re-balance the five weights to the hypothesis; produce a proposed profile that totals 100%.
Inputs	The hypothesis; the baseline; the corpus.
Key activities	Move weight toward the layers where the sector fails; link every deviation from baseline to ≥ 3 corpus cases; record the rationale for each move.
Method & rigor	Sum-to-100 constraint; a sensitivity check (does a ± 3 -point change flip any back-test outcome?); documented trade-offs. Weighting is the primary lever, but it only changes how a compliant output is scored.
Evidence standard	A deviation-to-cases mapping; no unsupported weight moves.
Output artifact	Weight table (Workbook tab 2).
Exit gate	Weights total 100% and every deviation is evidenced.
Owner (RACI)	Data-science lead (R); chair (A).

Phase 4 • Calibrate gates & tunable thresholds

Objective	Confirm the hard gates and set the tunable layer floors for the sector.
Inputs	The baseline gates (Data Integrity $< 2 \rightarrow$ cap 50; Fabrication $\rightarrow$ cap 40; High-Stakes Judgment $\rightarrow$ cap 60); the corpus.
Key activities	Keep or tighten the hard gates (never weaken); set the layer floor and the Wisdom/Experience minimums; add any sector-specific gate (e.g., 'no autonomous adverse determination').
Method & rigor	Invariant: weights never override gates. A stricter gate must be justified by demonstrated corpus harm; a gate may never be relaxed below the framework baseline.
Evidence standard	Each gate and threshold are tied to a documented failure mode.
Output artifact	Gate & threshold register (Workbook tab 3).
Exit gate	Auditor confirms no gate has been weakened.
Owner (RACI)	Risk/assurance auditor (R).

Phase 5 • Align the Decision Authority levels

Objective	Map maturity to Authority Levels for the sector and reserve the decisions that must stay human.
Inputs	The decision-matrix defaults; the corpus; regulatory mandates for human decision-makers.
Key activities	Set the maturity bands (AL-3 autonomous / monitored, AL-1 remediation); enumerate the sector's AL-4 policy-reserved decision classes; select the stakes-floor factors in force; set overseer-engagement requirements for AL-3.
Method & rigor	Two-determinant conservatism: assign the more conservative of the maturity ceiling and the stakes floor. AL-4 decision classes bind regardless of score; the stakes floor may raise required human authority but never lower it.
Evidence standard	Each AL-4 reservation tied to irreversibility, a legal mandate, or safety-criticality.
Output artifact	Authority alignment + AL-4 register (Workbook tab 4).
Exit gate	Ethics seat and regulator liaison endorse the reservations.
Owner (RACI)	Ethics seat (C); regulator liaison (C); chair (A).

Phase 6 • Derive & document the calibration objective

Objective	Make the thresholds defensible by derivation, not merely by record.
Inputs	The bands (Phase 5); estimates of consequence and reversibility per decision class.
Key activities	For each decision class, express expected loss as a function of maturity, stakes, reversibility and (where in force) overseer engagement; choose the threshold that minimizes expected loss.
Method & rigor	The explicit calibration objective (Master §4.5): calibrating a threshold is the auditable act of choosing the level that minimizes expected loss for a decision class, given its evidence.
Evidence standard	A short expected-loss rationale attached to each threshold choice.

Output artifact	Calibration-objective memo.
Exit gate	Council liaison accepts the derivation.
Owner (RACI)	Data-science lead (R); auditor (R).

Phase 7 • Back-test against the corpus

Objective	Prove the proposed profile catches the known failures without over-restricting benign cases.
Inputs	The proposed profile (tabs 2-4); the corpus with WG-assigned layer scores.
Key activities	Run every corpus case through the workbook; confirm each known failure resolves its expected outcome (Rejected / Constrained / AL-4); run a benign sample to measure over-restriction.
Method & rigor	Deterministic re-run — identical inputs always resolve to the same Authority Level. Report catches rate and false-positive / over-restriction rate; investigate every mismatch.
Evidence standard	100% of reserved and known-failure cases resolve correctly; over-restriction within the agreed tolerance.
Output artifact	Back-test scorecard (Workbook tab 5).
Exit gate	Pass rate meets the acceptance criteria.
Owner (RACI)	Auditor (R); data-science lead (R).

Phase 8 • Independent red-team / adversarial review

Objective	Try to break the profile before the public — or a real incident — does.
Inputs	The passed profile; reviewers independent of the drafting subgroup.
Key activities	Construct edge cases and gaming attempts; probe borderline scores; confirm that a fluent, confident, wrong output cannot buy authority.
Method & rigor	Reviewers must be independent of the authors; every finding is logged and closed before ratification.
Evidence standard	All red-team findings dispositioned; no open critical findings.
Output artifact	Red team report.
Exit gate	No open critical findings remain.
Owner (RACI)	Auditor (A); external reviewer (R).

Phase 9 • Public comment (RFC)

Objective	Expose the draft profile to the wider community for challenge and legitimacy.
Inputs	The draft profile with its evidence; a defined comment window.
Key activities	Publish the draft; collect comments; adjudicate each with a disposition and reason.
Method & rigor	Every comment logged with a disposition (accepted / rejected / deferred) and a rationale; legitimacy comes from the challenge surviving, not from avoiding it.
Evidence standard	All comments dispositioned; no unresolved objections carried into ratification.
Output artifact	RFC comment log (Workbook tab 6).
Exit gate	Comment window closed; objections resolved.
Owner (RACI)	Scribe (R); chair (A).

Phase 10 • Standards Council ratification

Objective	Convert the validated draft into a ratified standard.
Inputs	The completed workbook; the ratification checklist.
Key activities	The Council reviews the checklist, confirms the invariants held, and assigns a version number.
Method & rigor	All checklist items YES; precedence and invariants independently verified by the Council, not the working group.
Evidence standard	A signed ratification record.

Output artifact	Ratification checklist (Workbook tab 7).
Exit gate	Council sign-off.
Owner (RACI)	Standards Council (A).

Phase 11 · Publish & pin the versioned profile

Objective	Release the profile and make every future verdict traceable to it.
Inputs	The ratified profile and its version number.
Key activities	Publish 'WEKID for <Sector> vX.0'; pin the profile version to every verdict it produces; notify implementers.
Method & rigor	Version pinning is mandatory and auditable (Master §8.11): an audit can always reconstruct which sector weighting produces a given decision.
Evidence standard	Verdict records carry the profile version.
Output artifact	Published, versioned profile.
Exit gate	Profile live and referenced by verdicts.
Owner (RACI)	Council (A); chair (R).

Phase 12 · Monitor, diagnose & recalibrate

Objective	Keep the profile true as evidence accumulates.
Inputs	Outcome telemetry; new incidents; drift signals.
Key activities	Track real-world outcomes against assigned authority; watch for drift and new failure modes; schedule a periodic review and define the triggers that re-open the profile (a new failure class, a regulatory change, a demonstrated miss).
Method & rigor	Recalibration re-enters at Phase 2 and is itself a versioned, evidence-backed activity — not an ad-hoc edit.
Evidence standard	A documented review and a trigger register.
Output artifact	Monitoring report → next version.
Exit gate	Review completed, or a new RFC opened.
Owner (RACI)	Auditor (R); chair (A).

Part D — What makes a profile bulletproof

A profile is defensible only if each of the following holds. These are the questions a regulator, auditor or adversary will ask; the process above is designed so the answer is always yes.

- Evidence, not intuition. Every deviation from the baseline weights, and every gate or threshold, is tied to documented cases in the corpus — a minimum of three per weight move.
- It catches the known failures. The back-test resolves every reserved and known-failure case to its expected outcome; a profile that misses one of its own corpus cases is not ratifiable.
- It does not over-restrict. A benign sample measures the false-positive burden, so the profile is defensible on both sides — it neither under- nor over-governs.
- It is derived, not asserted. Each threshold minimizes an explicit expected-loss objective, so the numbers can be justified as well as recorded.
- It is independently challenged. Red-team review and public RFC surface weaknesses before deployment does.
- It is deterministic and reproducible. Identical inputs always resolve to the same Authority Level, so any verdict can be replayed and audited.
- It is versioned and pinned. Every verdict records the exact profile version that produced it, so decisions remain reconstructable years later.
- The floor is never lowered. Hard gates and AL-4 reservations survive calibration untouched; the group can only make the framework stricter, never weaker.

Part E — Governance, versioning & recalibration

The core framework stays stable while domain communities encode their risk priorities into calibrated, auditable profiles. Governance of that encoding follows one path: propose → evidence → back-test → red-team → public comment → Standards-Council ratification → publish & pin → monitor. Recalibration is not an edit; it re-runs the process from Phase 2 and produces a new numbered version. Triggers to re-open a ratified profile include a new failure class in the corpus, a regulatory change, a demonstrated miss in production, or a scheduled periodic review.

Appendix — Quick reference

A1 • The five Authority Levels

Level	Who decides	AI may execute	Typical profile
AL-0 No Authority	No one (rejected/informational)	Nothing	Immature or gated outputs
AL-1 Constrained	Policy (pre-approved bounds)	Whitelisted, reversible, impact-capped actions	Low-stakes, recoverable tasks
AL-2 Supervised	Human, per action	Proposals / staged actions only	Operational or high-impact decisions
AL-3 Augmented	AI within scope; human over process	Autonomous execution within scope	Repetitive/time-critical, proven maturity
AL-4 Human	Accountable human	Analysis & decision support only	Irreversible, safety-critical, legally reserved

A2 • Hard gates (invariant — may be made stricter, never weaker)

Gate	Trigger	Effect
Data Integrity	Data score < 2	Cap total maturity at 50/100
Fabrication	Fabricated citation, source, or tool output	Cap at 40/100; force rejection
High-Stakes Judgment	High-stakes output lacking uncertainty / risk language	Cap at 60/100

A3 • Maturity-to-Authority decision matrix (defaults; bands are tunable)

Maturity outcome	Condition	Authority / action
Approved (autonomous)	Score ≥ 80 and no gate	AL-3 — autonomous, subject to AL-4 reservations
Approved w/ monitoring	Score 70–79 and no gate	AL-3 — enhanced sampling review
Constrained	Any layer < floor, or Wisdom/Experience < min	AL-2 — human-in-the-loop
Remediation	Score 50–69 or soft gate	AL-1 — block & remediate
Insufficient	Score < 50 and no hard gate	AL-0 — reject & resubmit
Rejected	Hard gate tripped	AL-0 — reject; escalate if repeated
Policy-reserved	Decision class reserved by policy	AL-4 — human decides regardless of score

A4 • Illustrative sector weighting profiles (starting hypotheses only)

Profile	D	I	K	E	W	Emphasis
Global baseline	25	20	20	15	20	Balanced default
Defense	25	10	15	22	28	Judgment under adversarial uncertainty
Government	25	22	18	13	22	Transparency & policy-aligned judgment
Healthcare	30	12	18	15	25	Factual integrity & risk-calibrated judgment
Legal	28	15	25	10	22	Verifiable authority & sound reasoning
Education	22	25	23	10	20	Correct, clearly communicated understanding
Insurance	27	20	20	13	20	Numeric accuracy & explainable adverse decisions
Technology	22	18	25	20	15	Sound reasoning & procedure that executes

Note. These are hypotheses, not settled values — each becomes real only when its working group has tested it against documented cases and the Standards Council has ratified the resulting profile.

A5 • Templates (in the companion workbook)

Artifact	Workbook tab
Working-group charter & RACI	1 • Charter
Layer weight calibration (sum-to-100)	2 • Weights
Gate & Threshold register	3 • Gates & Thresholds
Authority alignment & AL-4 register	4 • Authority
Back-test scorecard (live decision matrix)	5 • Back-test
RFC / public-comment log	6 • RFC Log
Ratification checklist & sign-off	7 • Ratification