



VEYRUM ROBOTICS STANDARD

**VRS-GEN-104**

IR Category: Specification Integrity

Scoring basis, specification-integrity criteria SI-1...SI-6, evidence by tier and determination procedure for the spec-integrity intrinsic category.

VEYRUM RESEARCH INSTITUTE

DOCUMENT

VRS-GEN-104

VERSION

1.0

STATUS

Draft

DATE

15 September 2026

Contents

1	Introduction	
2	1 Scope	
3	2 Normative references	
4	3 Terms and definitions	
5	4 What this category measures	
6	5 Scoring basis	
7	6 Criteria	
7.1	6.1 SI-1 Claim declaration quality (0–4)	
7.2	6.2 SI-2 Headline claims honoured (0–26)	
7.3	6.3 SI-3 Secondary claims honoured (0–8)	
7.4	6.4 SI-5 Claims bound to the rated version (0–4)	
7.5	6.5 SI-6 Overclaim conduct (0–5)	
8	7 Evidence by tier	
9	8 Determination procedure	
10	9 Worked example (fictional model)	
11	10 Assessor checklist	
12	Bibliography	
13	Change history	

Foreword. IR Category Standard for **spec_integrity** — the datasheet-honesty signal of the Intrinsic Rating (category weights live in the VRS-GEN-005 master and its data twin, never here). This category scores the honesty of the datasheet — claimed versus independently measured performance — never the performance level itself (that is priced by **safety_incidents**, **reliability** and the sector protocols). This edition scores the category as a **sum of integer points** across five criteria (SI-1, SI-2, SI-3, SI-5, SI-6; 47 points; Clause 6), each earned against a plain-language ladder and bounded by an evidence tier — Unverified, Verified or Certified (Clause 7). It supersedes the rate-based model of the 0.x drafts: the input schema, the gap/effective-shortfall formula, the materiality weighting, the measurement window and small-sample rules, the recency-weighted determination formula and the qualitative level descriptors are all retired. Tolerance and measurement-uncertainty handling survives as the evidence bar of the honoured-claim criteria. Requirements use “shall”; recommendations use “should”.

1 Introduction

A datasheet is a manufacturer’s declaration, and an underwriter or lender who has never seen the robot must price risk from it. Two models with identical field records but different candour are not the same risk: the one whose numbers survive independent measurement can be trusted on the figures no assessor will ever test, while the one that overstates its headline numbers signals that its untested claims are also inflated. This category turns that signal into a repeatable number by asking five questions — are the headline claims published with a tolerance, a test method and the version they apply to; do the headline claims (including runtime) hold up against independent measurement; do the secondary claims hold up; are the claims and their evidence bound to the rated hardware and firmware version; and how does the manufacturer behave when a shortfall is found — and awarding integer points for each. The points sum to a category total of 0–47. What a model can earn is bounded by how well its evidence is substantiated: public measurement alone reaches the **Unverified** ceiling, a manufacturer’s claims and measurement pack reaches the **Verified** ceiling, and a witnessed live demonstration of the claims reaches the **Certified** ceiling. The category reuses evidence manufacturers already produce — the ISO 9283 / ISO 18646 performance test reports behind a CE technical file, the supplier’s declaration of conformity (ISO/IEC 17050), and accredited-lab reports (ISO/IEC 17025) — so adoption costs a mapping, not a new test campaign.

2 1 Scope

1.1 This Standard defines, for the **spec_integrity** intrinsic category: what the category measures (Clause 4); the scoring basis (Clause 5); the five criteria SI-1, SI-2, SI-3, SI-5 and SI-6 with their point ladders and tier ceilings (Clause 6); the evidence admissible at each tier (Clause 7); the determination procedure (Clause 8); a worked example (Clause 9); and the assessor checklist (Clause 10).

1.2 The category shall be assessed at one of three **tiers** — Unverified, Verified, Certified — and each criterion states the maximum points reachable at each tier. The category tier is the lowest tier among the criteria that determined the score (Clause 8). An Application Rating tier shall not exceed the Intrinsic Rating tier.

1.3 This Standard excludes category weighting and aggregation into the Intrinsic Rating (VRS-GEN-005 §6), sector re-weighting (the sector -201 protocols) and the evidence-grade definitions (VRS-GEN-012). It measures the *honesty* of the datasheet, not the level of performance: a modest but truthful datasheet outscores an ambitious but inflated one, and the safety, reliability and serviceability loss drivers are their own categories' concerns and shall not be inferred from the specification-integrity points.

3 Normative references

- **VRS-GEN-001**, Vocabulary and Terminology — robot and performance terms.
- **VRS-GEN-005**, Intrinsic Rating — Methodology — point aggregation, tiers, cold-start parity and the company-wide fallback rule (§6).
- **VRS-GEN-009**, Robot Risk Passport and Registry Schema — the record that carries the category result (Clause 8).
- **VRS-GEN-012**, Evidence Grades and Data Requirements — evidence kinds and the tier evidence bars.
- **ISO 8373:2021**, *Robotics — Vocabulary* — terms for robot performance characteristics.
- **ISO 9283:1998**, *Manipulating industrial robots — Performance criteria and related test methods* — pose accuracy and repeatability, path accuracy, cycle characteristics.
- **ISO 18646-1:2016 / -2:2024 / -3:2021 / -4:2021**, *Robotics — Performance criteria and related test methods for service robots* — locomotion; navigation; manipulation; assembly.
- **ISO/IEC 17050-1:2004** and **ISO/IEC 17050-2:2004**, *Conformity assessment — Supplier's declaration of conformity* — the declaration whose claims this category verifies.
- **ISO/IEC 17025:2017**, *General requirements for the competence of testing and calibration laboratories* — the accredited independent-measurement lane.
- **ISO/IEC Guide 98-3:2008 (GUM)**, *Uncertainty of measurement* — expanded uncertainty **U** ($k = 2$), the measurement-uncertainty credit in the SI-2/SI-3 evidence bar.
- **Regulation (EU) 2023/1230** on machinery (14 June 2023; applicable 20 January 2027) — accurate instructions and technical documentation as the regulatory basis for a declared datasheet.

4 3 Terms and definitions

Terms per VRS-GEN-001 and ISO 8373:2021. Locally:

- **claim** — a quantitative performance figure the manufacturer declares for the model in its datasheet, supplier's declaration of conformity, or technical documentation.
- **headline claim** — a claim a buyer, insurer or lender would rely on for a purchase, pricing or credit decision (typically payload, reach, accuracy/repeatability, speed, uptime, task-success, runtime); all other declared figures are **secondary claims**.
- **measurement** — an independent determination of the same quantity under a stated test method.
- **admissible measurement** — a measurement matched to a claim of the same metric and unit, taken against the rated firmware version, that meets the criterion's evidence bar.

- **per-claim result $\bar{x}$** — the dimensionless, tolerance- and uncertainty-adjusted degree to which one claim is honoured by its measurement (1.0 = within tolerance; lower = greater shortfall), and $\bar{\bar{x}}$ the mean of $\bar{x}$ across the counted claims; the ladders of SI-2 and SI-3 are read against $\bar{\bar{x}}$.
- **tolerance** — the declared $\pm$ band on a claim; credited before any shortfall counts.
- **expanded uncertainty U** — the measurement's uncertainty at coverage factor $k = 2$ (GUM); credited alongside the tolerance so a model is never penalised for a difference inside the measurement's own noise.
- **tier** — the evidence stringency at which the category is assessed: **Unverified** (public evidence only), **Verified** (a claims and measurement pack examined), **Certified** (claims witnessed live).
- **ceiling** — the maximum points a criterion can award at a given tier; points earned are capped at the ceiling of the assessed tier.
- **evidence bar** — the minimum evidence a criterion requires before any points are awarded; below it the criterion scores 0.
- **company-wide fallback** — where model-specific evidence is unavailable, the substitute points a criterion allows from manufacturer-wide evidence (VRS-GEN-005 §6); marked in the criterion row.
- **certificate rung** — a ladder step earned by a publicly verifiable third-party record or measurement in lieu of a witnessed demonstration.

5 4 What this category measures

4.1 The category measures **datasheet honesty** — the agreement between declared and independently measured performance — across five observable properties: the quality of the claim declaration (SI-1), whether the headline claims including runtime hold up against measurement (SI-2), whether the secondary claims hold up (SI-3), whether the claims and their evidence are bound to the rated version (SI-5), and the manufacturer's overclaim conduct (SI-6). It does **not** measure how high the performance is; truthful conservatism and truthful ambition score the same when both hold up.

4.2 The category shall be assessed at **model** granularity (manufacturer + model + declared major variant) against a single declared datasheet edition and firmware version; claims and the measurements behind them shall not be pooled across specification editions that changed the claimed figure.

4.3 Overperformance (measured better than claimed) is not rewarded here and carries no penalty; a measurement better than the claim is honoured in full.

4.4 This category carries the **basis / parameter-risk (data-quality)** driver that conditions the other categories: an inflated datasheet degrades the reliability of every category input drawn from declared figures and inflates the declared capability a lender's collateral valuation keys off. It does not itself carry frequency, severity or the finance residual-value drivers and shall not be read as a performance score.

6 5 Scoring basis

5.1 The category score shall be the **sum of integer points** awarded across the five criteria SI-1, SI-2, SI-3, SI-5 and SI-6 (Clause 6). The maximum is 47 points. There is no rate formula, no gap or effective-shortfall aggregate and no materiality weighting; the point range of each criterion sizes its importance.

Aggregation of this total into the Intrinsic Rating, tier rules and the cold-start parity and company-wide fallback rules are defined once in VRS-GEN-005 §6 and are not restated here.

5.2 Each criterion awards points against a **ladder** of rungs, highest first, decidable by inspection, measurement, test or documented evidence (VRS-GEN-002 §7.1). The assessor shall award the highest rung the evidence supports.

5.3 Each criterion states three **ceilings** — Unverified, Verified, Certified. Points earned shall be capped at the ceiling of the tier at which the criterion's evidence is assessed (5.3). A criterion may reach a higher tier than another; the category tier is recorded per Clause 8.

5.4 Each criterion states an **evidence bar**. Where the bar is not met, or no admissible measurement is presented, the criterion scores **0**; no evidence is 0 points, never a null or a discard.

5.5 A criterion marked **not applicable by class** for the model's form (Clause 6 Applicability row) is excluded from the score and from the totals; the reachable maximum is reduced accordingly and recorded (Clause 8). For the specification-integrity category all five criteria apply to all classes.

5.6 The ladders of SI-2 and SI-3 are read against the mean per-claim result $\bar{x}$ over the counted claims (Clause 6); the tolerance and measurement-uncertainty credit that yields each per-claim result $\bar{x}$ is the evidence bar of those criteria and is stated in Clause 7. Numeric thresholds inside a rung that carry **(prior)** are launch-edition values that change only through the calibration data twin, never in this document.

7 6 Criteria

6.0 Summary. Points and tier ceilings per criterion; totals reproduce by addition.

§	CRITERION	POINTS	UNVERIFIED CEILING	VERIFIED CEILING	CERTIFIED CEILING
6.1	SI-1 Claim declaration quality	0–4	4	4	4
6.2	SI-2 Headline claims honoured	0–26	14	19	26
6.3	SI-3 Secondary claims honoured	0–8	6	6	8
6.4	SI-5 Claims bound to the rated version	0–4	4	4	4
6.5	SI-6 Overclaim conduct	0–5	4	4	5
Total		47	32	37	47

7.1 6.1 SI-1 Claim declaration quality (0–4)

ROW	
What is measured	Headline claims are published with a tolerance, a test standard/method, and the firmware/hardware they apply to.
Ladder	4 = all three for every headline claim · 3 = tolerance + standard for all · 2 = standard for all, tolerances partial · 1 = bare figures with method for some · 0 = bare figures, no method
Unverified — evidence & ceiling	Published datasheet and manual. (fully assessable from public material). Ceiling 4.
Verified — evidence & ceiling	Manufacturer’s claims declaration (claims table with tolerance, standard, version). Ceiling 4.
Certified — evidence & ceiling	None (documentary). Ceiling 4.
Company-wide fallback	n/a
Evidence bar	A dated datasheet or claims declaration for the rated version exists.
Applicability	All classes

7.2 6.2 SI-2 Headline claims honoured (0–26)

ROW	
What is measured	Mean per-claim result $\bar{x}$ across all headline claims including runtime/endurance where the product has on-board energy storage (former SI-4 folded in, F13). Min 3 headline claims for the full range; with 2 counted claims the ceiling is 14, with 1 it is 7 (F14).
Ladder	$26 = \bar{r} = 1.0$ (every counted claim within tolerance) · $19 = \bar{r} \geq 0.75$ · $14 = \bar{r} \geq 0.5$ · $7 = \bar{r} \geq 0.25$ · $0 = \bar{r} < 0.25$ or no admissible measurement
Unverified — evidence & ceiling	Independent published measurements (editorial hands-on tests, peer-reviewed/academic measurements , published third-party reports, regulator filings). Only claims with such a measurement count; 1 counted claim → ceiling 7; 2 → 14. Ceiling 14.
Verified — evidence & ceiling	Manufacturer's own measurement evidence per headline claim (method, sample size, uncertainty) on the rated version; plus any third-party test reports held. Ceiling 19.
Certified — evidence & ceiling	Each headline claim demonstrated live at the declared value with manufacturer-supplied apparatus sized to the claim (payload lift/carry; speed over marked distance; reach to markers; step/slope/gap at declared dimension; repeatability to marks; cycle count; navigation goals; grasp set). Runtime is witnessed as the in-session energy sample (≥ 30 min representative duty, extrapolated). Claims with no witnessable form (e.g. MTBF) score from the Verified evidence at the Verified ceiling for that claim. Ceiling 26.
Company-wide fallback	n/a (model-specific by nature)
Evidence bar	At least one admissible measurement per counted claim, on the rated firmware version.
Applicability	All classes

7.3 6.3 SI-3 Secondary claims honoured (0–8)

ROW	
What is measured	Mean $\bar{x}$ across secondary claims (min 3; fewer → ceiling 3).
Ladder	$8 = \bar{r} = 1.0 \cdot 6 = \bar{r} \geq 0.75 \cdot 3 = \bar{r} \geq 0.4 \cdot 0 = \bar{r} < 0.4$ or no admissible measurement
Unverified — evidence & ceiling	As SI-2. Ceiling 6.
Verified — evidence & ceiling	Manufacturer’s measurement evidence for secondary claims. Ceiling 6.
Certified — evidence & ceiling	Secondary claims demonstrated where witnessable (acoustic level on the manufacturer’s meter shown on camera; power draw on the data feed; dimensions/mass on camera). Ceiling 8.
Company-wide fallback	n/a
Evidence bar	As SI-2.
Applicability	All classes

7.4 6.4 SI-5 Claims bound to the rated version (0–4)

ROW	
What is measured	The claims and the measurements behind them are tied to the hardware revision and firmware version being rated; changes to claims across versions are published.
Ladder	4 = all evidence on rated version and version-tagged · 2 = evidence on an earlier version with a published no-change statement · 0 = untagged or mismatched
Unverified — evidence & ceiling	Public release notes and datasheet edition history. Ceiling 4.
Verified — evidence & ceiling	Version-tagged claims declaration + safety/spec software change history for the rated range. Ceiling 4.
Certified — evidence & ceiling	Version proof on camera/feed at session open matches the declaration. Ceiling 4.
Company-wide fallback	n/a
Evidence bar	Rated hardware revision + firmware version declared.
Applicability	All classes

7.5 6.5 SI-6 Overclaim conduct (0–5)

ROW	
What is measured	Absence of material overclaim and the manufacturer's behaviour when a shortfall is found.
Ladder	5 = no headline claim with $\bar{x} < 0.6$ and a published corrections practice · 4 = evidenced by documents alone (public or pack), not demonstrated · 3 = one headline claim at $\bar{x} = 0.3$ or corrections practice absent · 1 = one headline claim at $\bar{x} = 0$ · 0 = two or more headline claims at $\bar{x} = 0$
Unverified — evidence & ceiling	As SI-2 sources plus published datasheet revisions/corrections. Ceiling 4.
Verified — evidence & ceiling	Manufacturer's measurement evidence + declared corrections practice. Ceiling 4.
Certified — evidence & ceiling	Derived from the SI-2 witnessed results. Ceiling 5.
Company-wide fallback	n/a
Evidence bar	SI-2 has ≥ 2 counted claims.
Applicability	All classes

8 7 Evidence by tier

7.1 Unverified. The assessor works only from evidence anyone can obtain without the manufacturer's cooperation: the published datasheet and manual for the declaration quality and version binding (SI-1, SI-5); and independent published measurements — editorial hands-on tests, peer-reviewed or academic measurements, published third-party test reports and regulator filings — matched to the headline and secondary claims (SI-2, SI-3), together with published datasheet revisions and corrections (SI-6). Only a claim with such an admissible measurement counts; each criterion's Unverified ceiling is the most an honest reading of public evidence can earn (SI-2 → 14, SI-3 → 6, SI-6 → 4), and the higher rungs of SI-2 and SI-3 require Verified or Certified evidence. The category tier is **Unverified**.

7.2 Verified. The assessor additionally examines a claims and measurement pack from the manufacturer: a claims declaration listing each claim with its tolerance, test standard and version (SI-1, SI-5); the manufacturer's own measurement evidence per headline and secondary claim, stating method, sample size and expanded uncertainty on the rated version, plus any third-party test reports held (SI-2 → 19, SI-3 → 6); and the declared corrections practice (SI-6). Verified ceilings apply only to criteria whose pack items are present and examined; the category tier is **Verified**.

7.3 Certified. The Certified tier adds a witnessed session in which each headline claim is demonstrated live at its declared value with manufacturer-supplied apparatus sized to the claim (payload, speed, reach, step/slope/gap, repeatability, cycle count, navigation goals, grasp set), runtime witnessed as an in-session energy sample over at least 30 minutes of representative duty and extrapolated, and secondary claims shown where witnessable (acoustic level on the manufacturer's meter on camera, power draw on the data feed, dimensions and mass on camera). Claims with no witnessable form (e.g. MTBF) carry their Verified points. SI-2 reaches 26, SI-3 reaches 8 and SI-6 reaches 5 from the witnessed results; SI-1 and SI-5 are documentary and carry their documentary ceiling. The category tier is **Certified**.

7.4 Tolerance and measurement-uncertainty bar. For SI-2 and SI-3 a claim is honoured only after its declared tolerance and the measurement's expanded uncertainty U at $k = 2$ (GUM) are credited: the shortfall that counts against a claim is the amount by which the measurement is worse than the claim beyond $\text{tolerance} + U$, so a difference inside the declared band or inside the measurement's own noise is no shortfall. Where the manufacturer declares no tolerance it is taken as zero and only U is credited; where U is absent on an otherwise accredited measurement, the measurement is treated one grade lower and U is taken as the method's published repeatability. A claim with no admissible measurement meeting this bar does not count and, where it leaves a criterion below its minimum counted claims, reduces the ceiling as the criterion row states (SI-2, SI-3); a criterion with no counted claim scores 0 (5.4).

7.5 Admissibility and conflicts. A measurement is admissible for SI-2/SI-3 only where its metric and unit match the claim and it was taken against the rated firmware version (SI-5). Where two admissible measurements of the same claim conflict, the assessor shall use the higher-grade measurement; among equal grades, the more recent; among equal grade and date, the one with more replication. The discarded measurement shall be retained for audit. A field-record window, where a claim is honoured against a running record rather than a single test, is the window stated in the criterion row and nowhere else.

7.6 Certificate rung and public records. Where a claim is honoured by a publicly verifiable third-party record — an accredited-lab report, a peer-reviewed measurement, a regulator filing or a published corrections notice — that record earns the corresponding rung at the Unverified tier without a witnessed demonstration; a record that cannot be verified shall not.

7.7 Company-wide and parent-company evidence. No SI criterion offers a company-wide fallback: claim integrity is model-specific by nature (SI-2), and the declaration, version-binding and conduct criteria are assessed on the model's own datasheet and corrections history. Parent-company corrections practice is admissible as context for SI-6 only where the manufacturer is a declared subsidiary and the parent's published practice governs the model.

9 8 Determination procedure

Step 1 — Fix the tier and applicability. Record the tier at which each criterion is assessed (Clause 7) and, per 5.5, the set of any not-applicable criteria; for the specification-integrity category all five apply.

Step 2 — Pair and admit (SI-2, SI-3). Match each headline and secondary claim to an admissible measurement of the same metric and unit on the rated version, crediting the declared tolerance and the

measurement's expanded uncertainty before any shortfall counts (7.4, 7.5). Record the counted-claim count per criterion; where it falls below the criterion's minimum, apply the reduced ceiling stated in the criterion row.

Step 3 — Award points. For each criterion, the assessor shall award the highest ladder rung the evidence supports (Clause 6) — for SI-2 and SI-3 the rung set by the mean per-claim result $\bar{x}$ — and shall **cap** the award at the ceiling of that criterion's assessed tier (5.3). Where the evidence bar is not met, the award shall be 0 (5.4). Numeric rung thresholds marked **(prior)** are read as written.

Step 4 — Record substitutions. No company-wide fallback applies (7.7); where SI-6 rests on parent-company corrections practice, record that basis.

Step 5 — Sum. The category score shall be the **sum of the capped per-criterion points**. With no not-applicable criteria the reachable maximum is 47; otherwise it is 47 minus the ranges of the excluded criteria, and this reachable maximum shall be recorded alongside the score.

Step 6 — Record the tier reached. The category tier is the lowest tier among the criteria that determined the score (a score resting on any Unverified criterion is an Unverified-tier result).

Step 7 — Output, passport binding and loss-driver mapping. The result shall be recorded in the model's Robot Risk Passport specification-integrity-category record (VRS-GEN-009): the category score, the per-criterion points, the ceiling reached and tier for each criterion, the counted-claim count per criterion, the not-applicable set and reachable maximum, and any parent-company substitution, so that an underwriter or lender can reproduce and audit the score from the passport alone. The per-criterion points carry the **basis / parameter-risk (data-quality)** driver: the honoured-claim points (SI-2, SI-3) and the declaration-quality and version-binding points (SI-1, SI-5) condition the reliability of every other category input drawn from declared figures and the declared capability a lender's collateral valuation keys off; the overclaim-conduct points (SI-6) carry the misrepresentation signal — surfaced never silent. These are underwriting and credit **signals**, not prices; rate-setting is the carrier's act and risk classification the actuary's (ASOP No. 12, Risk Classification). This category conditions the other categories' inputs; it does not itself carry frequency, severity, or the finance residual-value and recoverability drivers, which are single-sourced in VRS-GEN-101, -102, -103 and -107 and shall not be inferred from the specification-integrity points (4.1, 4.4).

10 9 Worked example (fictional model)

"*Cirrus P200*", a collaborative arm, class = manipulator, assessed at the **Verified** tier: the assessor examines the manufacturer's claims declaration and its own measurement evidence per claim on the rated firmware version, with no witnessed live demonstration. All five criteria apply (reachable maximum at Verified = 37).

§	CRITERION	RUNG EARNED	VERIFIED CEILING	POINTS
6.1	SI-1	4 (every headline claim published with tolerance, standard and version)	4	4
6.2	SI-2	4 counted headline claims, $\bar{r} = 0.80 \rightarrow$ rung 19 ($\bar{r} \geq 0.75$)	19	19
6.3	SI-3	3 counted secondary claims, $\bar{r} = 0.78 \rightarrow$ rung 6 ($\bar{r} \geq 0.75$)	6	6
6.4	SI-5	4 (all evidence on the rated version and version-tagged)	4	4
6.5	SI-6	4 (no headline claim with $r < 0.6$; documents only, not demonstrated)	4	4
Total			37	37

The specification-integrity total is **37 of a reachable 37** at the Verified tier. Arithmetic: $4 + 19 + 6 + 4 + 4 = 37$. Read the two claim-honouring criteria. At SI-2 the four headline claims — repeatability, payload, reach and runtime — were measured on the rated version; after crediting each claim's declared tolerance and the measurement's expanded uncertainty, the mean per-claim result is $\bar{r} = 0.80$, which lands on the 0.75 rung (19 points) and is held at the Verified ceiling of 19 (the top rung of 26 is a Certified-tier witnessed check). At SI-3 the three secondary claims (cycle time, idle power and acoustic level) give $\bar{r} = 0.78$, earning the 6-point rung at the Verified ceiling of 6. SI-1 and SI-5 reach 4 on the version-tagged declaration, and SI-6 earns the documentary rung of 4 because no headline claim fell below $r = 0.6$ and a corrections practice is published — the 5-point rung would require the witnessed SI-2 results of the Certified tier. Because every criterion rests on Verified-tier evidence, the category tier is **Verified** (Step 6). The score, per-criterion points, counted-claim counts, and tier and ceiling reached for each criterion are recorded to the passport (Step 7).

11 10 Assessor checklist

- [] Model granularity, datasheet edition and firmware version fixed; no pooling across editions (4.2).
- [] Assessment tier fixed per criterion; not-applicable set recorded (Step 1); reachable maximum recorded.
- [] **SI-1** claim declaration quality scored on tolerance + standard + version coverage of the headline claims (6.1).
- [] **SI-2** headline claims (incl. runtime) scored on $\bar{r}$ over counted claims; ≥ 1 admissible measurement per counted claim (evidence bar); reduced ceiling applied with 1 or 2 counted claims; tolerance and uncertainty credited (7.4).
- [] **SI-3** secondary claims scored on $\bar{r}$; ceiling reduced below 3 counted claims (6.3).

- [] **SI-5** claims and their evidence bound to the rated hardware/firmware version; version-tagged (6.4).
- [] **SI-6** overclaim conduct scored on absence of material overclaim and published corrections practice; SI-2 has ≥ 2 counted claims (evidence bar).
- [] Each award capped at the assessed-tier ceiling (5.3); no-admissible-measurement and below-bar criteria scored 0, never null (5.4, 7.4).
- [] Measurement conflicts resolved by grade $\rightarrow$ recency $\rightarrow$ replication; discarded measurement retained for audit (7.5).
- [] Score, per-criterion points, counted-claim counts, tier and ceiling reached recorded to the passport; outputs mapped to the basis/parameter-risk data-quality driver and the SI-6 misrepresentation signal (Step 7, VRS-GEN-009).

12 Bibliography

- ISO 8373:2021, *Robotics — Vocabulary*.
- ISO 9283:1998, *Manipulating industrial robots — Performance criteria and related test methods*.
- ISO 18646-1:2016; ISO 18646-2:2024; ISO 18646-3:2021; ISO 18646-4:2021, *Robotics — Performance criteria and related test methods for service robots*.
- ISO/IEC 17050-1:2004 and ISO/IEC 17050-2:2004, *Conformity assessment — Supplier's declaration of conformity — Parts 1 and 2*.
- ISO/IEC 17025:2017, *General requirements for the competence of testing and calibration laboratories*.
- ISO/IEC Guide 98-3:2008, *Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*.
- Regulation (EU) 2023/1230 of 14 June 2023 on machinery (applicable 20 January 2027).
- Actuarial Standards Board, *ASOP No. 12, Risk Classification (for All Practice Areas)* (rev. 2005, effective 1 May 2006) — the signals-not-prices boundary (Clause 8, Step 7).

13 Change history

DATE	VERSION	STATUS	CHANGE	AUTHORITY
2026-09-15	1.0	Draft	VRS 2026 rewrite: replaced the rate-based scoring (input schema, claim-metric coding, directionality/materiality/gap/effective-shortfall formula, materiality-weighted mean, measurement window/recency/small-sample rules, the determination formula and the qualitative level descriptors with the overclaim cap) with five point-scored criteria SI-1, SI-2, SI-3, SI-5 and SI-6 (47 points; Clause 6) with per-tier ceilings; folded runtime into SI-2 (F13), added the reduced ceilings for 1–2 counted claims (F14) and the SI-2 documentary/overclaim rungs; added Clause 5 Scoring basis, carried the tolerance and measurement-uncertainty rules into the SI-2/SI-3 evidence bar (Clause 7), added Clause 8 point-sum determination and passport binding with the basis/parameter-risk (data-quality) loss-driver mapping, regenerated worked example (Verified tier, 37/37) and assessor checklist; retired the confidence band, provisional flag, renormalisation, class/dataset median, credibility and qualitative IR levels, and the cold-start/GEN-201 dependency; dropped the GEN-201/GEN-202 references and the ASOP No. 23/25 statistics bibliography that served the rate model; kept the ISO 8373:2021 / ISO 9283:1998 / ISO 18646 series / ISO/IEC 17050 / ISO/IEC 17025:2017 / GUM / EU 2023/1230 anchors; see GEN-005 §6. No constant value changed.	VRS 2026 rewrite Stage 4
2026-09-11	0.3	Draft	ACTUARIAL review: added §9.2 Passport binding + exposure denominator and §9.3 Loss-driver/decision map; added VRS-GEN-009 normative reference, ASOP No. 23 and ASOP No. 25 anchors + bibliography, and a §12 checklist line. No constant value changed. Net 2975→3572 words; shall 22→26	CEO direction 2026-09-11
2026-09-11	0.2	Draft	DEEPEN: expanded to publishable depth — input schema (Cl. 5), claim-metric vocabulary + directionality + materiality (Cl. 6), tolerance/uncertainty handling (Cl. 7), window/recency/small-sample rules (Cl. 8), 9-step determination procedure with formula, units and prior-stage weights (Cl. 9), overclaim cap (Cl. 10), worked example computing the sub-score with a recorded cap mismatch (Cl. 11), assessor checklist (Cl. 12); corrected the ISO/PAS 5672 mischaracterisation and added dated anchors (ISO 9283:1998, ISO 18646 series, ISO/IEC	CEO direction 2026-09-11

DATE	VERSION	STATUS	CHANGE	AUTHORITY
17050, ISO/IEC 17025, GUM, EU 2023/1230); CALIBRATION-PENDING replaced by GEN-202 stage labels				
2026-09-05	0.1	Draft	Split from VRS-GEN-005 §5.4; level descriptors added per the category-standard wireframe	CEO goal 2026-09-05

About Veyrum Research Institute

Veyrum Research Institute is the research arm of Veyrum — the independent intelligence and trust layer for the robotics industry (research, ratings and registry, advisory, data, and specialist insurance & finance introductions).

Disclaimer

The opinions in this document are independent and current as of the date of publication, based on information believed reliable; they are not guarantees or warranties of safety, fitness, or compliance, and are not engineering, legal, or financial advice. This document may be revised. © 2026 Veyrum.

Contact: hello@veyrum.com · VRS-GEN-104 · Version 1.0