TECHNICAL WHITE PAPER · ACR EVALUATOR SERIES

Accessibility and Governance Risk Scoring

How the ACR Evaluator quantifies product-level and governance risk across the ICT procurement lifecycle

AUTHORS

Mark Turner, Ph.D.
Chief Product Designer, Revelo Software

Jatan Rathod
Analytics Engineer, Revelo Software

PUBLISHED

June 2026 · v1

ABSTRACT

Determining which accessibility requirements apply to an ICT product is only the first step. Organizations must also quantify the risk those requirements represent — both the risk created by the product's own conformance shortfalls and the governance risk created by weak evidence, vendor accountability, and contract controls.

This white paper describes the two-score risk model at the core of the ACR Evaluator's Accessibility Risk Assessment capability: the Accessibility Risk Score (ARS) and the Governance Risk Score (GRS). It explains how each score is calculated, how they interact, and how together they produce a lifecycle-aware, defensible procurement recommendation.

Executive Summary

Accessibility conformance tells a buyer what a product can and cannot do for users with disabilities at a point in time. It does not tell the buyer how much risk that finding represents, how likely remediation is to succeed, or whether the governance controls surrounding the product are strong enough to protect accessible outcomes over the life of a contract. Quantifying those dimensions requires a separate and complementary layer of analysis.

The ACR Evaluator's Accessibility Risk Assessment capability provides that layer through two linked scores. The Accessibility Risk Score (ARS) measures the risk created by the product's own conformance shortfalls — aggregated from the product's evaluation findings and normalized across the applicable criteria set. The Governance Risk Score (GRS) measures something different: whether the agency, vendor, contract, evidence set, and operational controls are collectively strong enough to manage accessibility risk over time.

Two products can receive identical ARS values yet carry very different GRS values — a product with minor defects but no remediation plan, stale evidence, and no contractual enforcement provisions may pose far greater long-term risk than a product with more findings but strong governance controls in place. Keeping the scores separate surfaces that difference. Combining them into a single number would hide it.

Both scores are deterministic, auditable, and stage-aware. The questions asked, evidence requirements, and area weightings all shift depending on where the buyer stands in the ICT procurement lifecycle: selecting a vendor, accepting delivery, reviewing a major release, exercising a contract option, or conducting portfolio oversight. A separate Evidence Confidence Score (ECS) accompanies every result, signaling how trustworthy the inputs are — and a set of hard override floors prevent a favorable GRS from masking critical governance failures.

> **Core principle.** A product may score well on conformance while the surrounding governance is weak. A product may score poorly on conformance while remediation commitments, contractual controls, and evidence quality make the risk manageable. Both conditions are invisible to a conformance-only assessment.

Why Two Scores Are Necessary

A conformance grade alone can mislead a procurement decision. Consider two competing proposals:

- Vendor A delivers a product with several low-severity WCAG failures, a current ACR, an independent third-party test report, a formal remediation SLA, and contractual enforcement rights.
- Vendor B delivers a product with fewer reported failures, but a self-assessed ACR that is 18 months old, no independent validation, and no contractual obligation to remediate.

On conformance alone, Vendor B looks stronger. On governance, Vendor B is the higher-risk choice. Without a GRS to surface the governance dimension, a procurement team evaluating only the product grade would make the less-defensible decision.

The two-score model is designed to make exactly this pattern visible. The ARS captures conformance risk; the GRS captures governance risk. Together they produce a single recommendation, but neither alone is sufficient.

The Accessibility Risk Score (ARS)

The ARS aggregates risk across every accessibility criterion applicable to the product, then normalizes and adjusts the result for the technology category. It is computed directly from the ACR Evaluator assessment, so no separate data collection is needed — the same evaluation that produces the conformance record feeds the ARS calculation automatically.

Criterion-Level Risk Calculation

For each applicable criterion, the system computes a criterion risk value as the product of four organization-configurable factors:

Factor	What it measures	Set by
Criterion Criticality	How important this criterion is relative to others in the applicable standard	Organization profile
Population Impact	The breadth of the disability community affected by non-conformance with this criterion	Organization profile
Severity of Shortfall	How significantly the product fails the criterion (derived from the evaluation result: Supports, Partially Supports, Does Not Support)	Evaluation finding
User Impact	The degree to which non-conformance affects the user's ability to complete key tasks	Evaluation finding

Table 1 — The four organization-configurable factors multiplied together to produce each criterion's risk value.

These criterion scores are summed into a raw risk value, divided by the maximum possible risk for the applicable scope, scaled to 0–100, and then multiplied by an ICT Risk Multiplier that reflects the inherent risk profile of the technology category under evaluation. The result is the ARS, mapped to one of five risk bands.

ICT Risk Multiplier

Because different technology categories carry different compliance stakes — a public-facing web application delivering benefits carries more inherent risk than an internal desktop tool used by a small team — the ARS includes a multiplier that adjusts the score for the ICT type. The multiplier values are organization-configurable and are effective-dated so that historical scores remain reproducible.

Reproducibility

Because all factor values are organization-specific and effective-dated, the same product can score differently for different organizations, and any historical score can be exactly reproduced from the configuration that was in force on the evaluation date. This makes the ARS suitable for audit trails and procurement records.

The Governance Risk Score (GRS)

The GRS measures whether the governance context surrounding a product — the agency's accountability structures, the evidence quality, the vendor's track record, the contract's enforcement provisions, and the operational controls — are collectively strong enough to keep accessibility risk managed over time. It is computed from a structured questionnaire whose questions, scoring weights, and emphasis shift by lifecycle stage.

Six Scoring Areas

The GRS is built across six scoring areas. Five measure governance weakness; the sixth — Mission Exposure — represents the consequence if governance fails.

Scoring area	What it measures	Code
Governance Program & Accountability	Named ownership, accessibility policy, leadership commitment, workforce capability	GOV
Evidence & Validation Assurance	ACR/VPAT currency, independence of testing, methodology documentation, version match	EVA
Vendor Delivery Capability & Responsiveness	Vendor track record, remediation history, disclosure behavior, support processes	VCR
Contractual Controls & Enforceability	Standards requirements in contract, remediation SLAs, enforcement mechanisms, validation rights	CCE
Deployment, Change & Operational Control	Change management integration, monitoring processes, post-deployment reassessment obligations	DCO
Mission Exposure & User Impact	Public-facing vs. internal, user volume, rights/benefits delivery, known legal or complaint exposure	MUC

Table 2 — The six GRS scoring areas. Five measure governance weakness; Mission Exposure represents the consequence if governance fails.

Stage-Aware Weighting

The relative weight of each scoring area changes by lifecycle stage. Early in procurement, evidence quality and contract structure are the buyer's strongest levers, so those areas carry the most weight. After deployment, change management and continuous monitoring dominate. At the portfolio level, agency-wide governance program strength matters most.

This stage-awareness prevents a single weighting profile from being misapplied across the lifecycle. A pre-award assessment and a contract-renewal assessment are answering different questions, and the scoring model reflects that.

Answer-to-Number Scoring

Every questionnaire answer is converted to a risk value on a 0–5 scale:

Risk value	Meaning
0	Strong control — current, verified, complete, effective
1	Minor weakness — low residual risk
2	Limited weakness — manageable but notable gap
3	Material weakness — meaningful exposure
4	Major weakness — high probability of unmanaged risk
5	Absent, stale, contradicted, undocumented, or unknown

Table 3 — The 0–5 answer risk scale applied to every questionnaire response.

A governing principle of the scoring design is that unknown or missing data is treated as risk, not as neutral information. Every "information not provided," "don't know," and "not yet determined" answer

is scored at 5. A vendor who declines to provide evidence is scored as risky rather than given the benefit of the doubt. “Not applicable” answers are excluded from the calculation rather than scored, so they neither help nor hurt the result.

Score Calculation

Area risk percentages are computed by averaging the answer risk values within each area, weighted by applicability (1.0 for applicable questions, 0.0 for not-applicable questions). Area contributions are then weighted by the stage profile and summed into a preliminary GRS.

> **Area risk %** = sum(answer risk × applicability) ÷ sum(5 × applicability)
> **Area contribution** = stage weight × area risk %
> **Final GRS** = max(sum of contributions, highest triggered override floor)

Override Floors

A weighted average is not always sufficient. Federal guidance treats certain failures as decision blockers rather than ordinary risk deductions. The model applies hard floors — if specific critical conditions are met (for example, acceptance of a product with critical unresolved defects and no remediation plan, or a contract with no enforcement mechanisms), the final GRS is automatically elevated to the floor level regardless of the calculated score. This prevents a favorable average from concealing a single disqualifying condition.

The Evidence Confidence Score (ECS)

A risk score is only as reliable as the evidence behind it. The Evidence Confidence Score travels with every GRS result and signals how much the assessment inputs can be trusted. It is scored 0–100 — higher is better — and is computed across five weighted components:

ECS component	What it measures	Weight
Completeness	Whether all required documentation has been provided	25%
Recency & version match	How current the evidence is and whether it matches the exact version delivered	20%
Independence	Whether testing was conducted by an independent third party, internal experts, or vendor self-assessment	20%
Scope coverage	Whether testing covered the full applicable criteria set and critical user workflows, including assistive technology combinations	20%
Reproducibility & traceability	Whether methodology and findings are documented in enough detail to reproduce the result	15%

Table 4 — The five weighted components of the Evidence Confidence Score.

The ECS produces a confidence tier — High, Moderate, Low, or Very Low — that gates the recommendation. A Low ECS prevents an unconditional “proceed” or “renew” recommendation regardless of how favorable the GRS looks. An evaluator confidence override can further cap the tier: a reviewer who assesses the evidence as Low confidence prevents the result from being reported as High confidence irrespective of the computed ECS.

Five Lifecycle Assessment Stages

Accessibility risk is not a one-time determination. The model defines five distinct assessment stages, each anchored to a specific procurement decision point. Selecting a stage determines the questionnaire shown, the evidence requirements applied, and the GRS weighting profile used.

Stage	Decision being made	Dominant GRS areas
Pre-award	Which vendor to select; what contract terms to require	EVA, CCE
Before acceptance	Whether to accept delivery; what conditions to apply	EVA, VCR, CCE
After major releases	Whether a post-deployment change has preserved accessible outcomes	DCO, GOV, EVA
At option exercise / renewal	Whether the vendor has earned renewal on a governance basis	VCR, CCE, DCO, EVA
Periodic portfolio review	How accessibility risk is distributed across the agency's ICT portfolio	GOV, DCO, MUC

Table 5 — Each stage anchors to a procurement decision point and shifts the dominant GRS scoring areas.

The portfolio review stage is structurally different from the others. Rather than a per-product questionnaire, it combines a short agency-level program questionnaire — covering governance program maturity, lifecycle gate enforcement, and complaint processes — with an automatic roll-up of individual product scores, weighted by each product's criticality and user reach. This produces an exposure-adjusted portfolio GRS and a set of distribution metrics — percentage of assets in High or Critical band, percentage with current version-matched evidence, percentage with overdue remediation — that surface concentration risk that averages alone would hide.

Worked Example: A Multi-Vendor Comparison

Consider a federal agency evaluating three competing proposals for a new collaboration platform. The ACR Evaluator has produced an evaluation report for each vendor's product. The risk assessment then runs each through the pre-award questionnaire and produces the following results:

	Vendor A	Vendor B	Vendor C
ARS (product risk)	38 — Low	22 — Low	61 — Moderate
GRS (governance risk)	29 — Low	74 — High	41 — Moderate
ECS (evidence confidence)	82 — High	34 — Very Low	58 — Low
Recommendation	**Proceed**	**No-go (floor triggered)**	**Proceed with conditions**

Table 6 — Three competing proposals scored at the pre-award stage. Values are illustrative of the determination, not a complete result.

Vendor B is the instructive case. Its product score is the strongest of the three — the fewest reported conformance failures. But its evidence is 22 months old and self-assessed, the contract proposal includes no accessibility SLA, and the vendor's response to evidence requests was incomplete. The

ECS of 34 falls in the Very Low tier, triggering a confidence floor that elevates the GRS to the High band. A procurement decision based on the product grade alone would have selected the highest-governance-risk vendor.

Vendor A is the recommended award. Its product has more reported findings than Vendor B, but the findings are low-severity, the evidence is current and independently validated, and the proposal includes specific remediation SLAs and contractual enforcement rights. The governance controls make the product risk manageable; the GRS reflects that.

Risk Bands and Recommendations

Both the ARS and GRS use the same five risk bands, scored 0–100, so they are legible side by side. The GRS band drives the base recommendation; the ECS tier and any override floors can tighten or override it.

Score	Band	Base recommendation (GRS)
0–19	**Very Low**	Proceed / routine monitoring
20–39	**Low**	Proceed with standard controls
40–59	**Moderate**	Proceed with conditions and targeted validation
60–79	**High**	Escalate; enhanced validation and contractual cure plan required
80–100	**Critical**	No-go unless executive risk acceptance is explicitly documented

Table 7 — The five shared risk bands and the GRS base recommendation each one drives.

The recommendation is not the final word — it is a structured input to the procurement official's decision. The audit trail records the GRS, the ECS tier, which override floors fired and why, the full questionnaire answers, and the configuration version in effect at the time. That trail is what makes the assessment defensible in procurement, compliance, and audit contexts.

Conclusion

Conformance assessment answers the question of what a product does or does not support for users with disabilities. Risk assessment answers the complementary question of what that finding means for a procurement decision — how much risk it represents, how likely it is to be remediated, and how strong the governance controls are that surround it.

By separating product risk (ARS) from governance risk (GRS), attaching an evidence confidence signal (ECS), and calibrating all three to the lifecycle stage at which the decision is being made, the ACR Evaluator produces assessments that are consistent across reviewers, transparent in their derivation, and traceable to specific answers and configurations. That foundation is what makes the resulting risk profile suitable for procurement, compliance, and portfolio risk management.

References

U.S. Access Board. Information and Communication Technology (ICT) Standards and Guidelines — Revised Section 508 Standards and Section 255 Guidelines. 36 CFR Part 1194, 2017.

World Wide Web Consortium (W3C). Web Content Accessibility Guidelines (WCAG) 2.2. W3C Recommendation, 5 October 2023.

ETSI. EN 301 549 — Accessibility requirements for ICT products and services. V3.2.1, 2021.

Information Technology Industry Council (ITI). Voluntary Product Accessibility Template (VPAT) / Accessibility Conformance Report (ACR), Version 2.5.

Turner, M., Rathod, J. An Automated Framework for Determining ICT Accessibility Requirements. Revelo Software ACR Evaluator Technical Series, June 2026. [ACR Evaluator Technical White Paper No. 1]