Blacklink Trust Standards

Version 2.0 — Extended Specification

Trust Standards v2.0 defines the formal methodology, constraints, and governance model behind Blacklink Trust. This document is intended for educators, platform administrators, developers, auditors, and partners who require a deep understanding of how trust decisions are generated, interpreted, and reviewed.

1. Philosophy of Trust

Blacklink Trust is built on the principle that trust is not binary. Rather than categorizing information as simply "trusted" or "untrusted," Trust provides a structured, explainable assessment that reflects uncertainty, context, and available evidence.

Trust prioritizes transparency over certainty. A lower score is not an accusation, and a high score is not an endorsement. Trust Scores exist to inform judgment, not replace it.

2. Trust Score Model

Every source evaluated under Trust Standards v2 receives a numerical Trust Score between 0 and 100. This score represents the system's confidence in the source's credibility based on measurable signals.

2.1 Score Bands

1. 90–100 — Highly Trusted: Strong signals across all dimensions.
2. 75–89 — Trusted: Reliable with minor limitations.
3. 60–74 — Caution: Mixed or incomplete signals.
4. 40–59 — Low Confidence: Significant gaps or ambiguity.
5. 0–39 — Not Trusted: High risk or misleading indicators.

3. Signal Categories

3.1 Identity

Assessment of authorship clarity, organizational ownership, and accountability structures.

3.2 Evidence

Evaluation of citations, data sources, methodologies, and factual support.

3.3 Recency

Measurement of timeliness relative to subject matter requirements.

3.4 Transparency

Disclosure of intent, funding, affiliations, limitations, and data usage.

3.5 Safety

Suitability for educational contexts and avoidance of harmful framing.

4. AI-Assisted Evaluation

Trust Standards v2 allows the use of AI systems such as Aero to assist in evaluating sources. AI outputs are treated as advisory signals, not authoritative judgments.

AI-generated Trust Scores must include explicit reasoning, surface uncertainty, and avoid overclaiming. Human review is encouraged for high-impact decisions.

5. First-Party & Verified Sources

First-party verification is applied to domains operated by Blacklink or verified partners. Verification confirms identity and accountability but does not exempt a source from scrutiny.

6. Governance & Review

Trust Standards are reviewed periodically to reflect changes in technology, education policy, and information ecosystems. Scores may be recalculated as standards evolve.

7. Limitations & Disclosure

Trust Scores reflect available information at the time of analysis. No Trust Score should be interpreted as an absolute guarantee of accuracy. Independent verification is always encouraged.