

Responsible AI Policy

How RiderScore models are designed, governed, explained and contested, including inputs that are explicitly excluded.

Version 1.0 | Effective 19 August 2026 | Last reviewed 19 August 2026

RiderScore Technologies Ltd · riderscore.co.ke

1. Principles

Explainability first: every score is published with the factors that produced it and the model version used.

Human accountability: RiderScore outputs support decisions; they do not replace the institution's decision maker.

Proportionality: we process the minimum data needed for the assessment purpose.

Contestability: riders can question a score and request review.

2. Excluded inputs

Our scoring models do not use sensitive or protected characteristics as inputs. Excluded categories include gender, ethnicity, tribe, religion, health status, disability, political affiliation, marital status and sexual orientation.

We also exclude fields whose only value would be as a close proxy for those characteristics.

3. Model design and versioning

Models are deterministic and versioned. Each score snapshot records the model version, the input signals, the weights applied and the resulting factor breakdown.

Historic snapshots are immutable so that a past assessment can always be reconstructed and explained.

4. Data quality and provenance

Every rider record carries its provenance, for example rider-provided, institution-provided or operator ledger.

Unverified data is labelled as unverified rather than presented as confirmed. We do not fabricate verification status.

5. Monitoring and review

Model changes go through documented review before release. We monitor score distributions and factor behaviour for unexpected drift.

Where a defect is found, affected snapshots are identified and impacted institutions are notified.

6. Human review and contestability

A rider may request an explanation of their score and may submit a correction or challenge through the rider portal or complaints@riderscore.co.ke.

Where a correction is accepted, the underlying data is updated and a new score snapshot is produced. The prior snapshot remains on record with the reason for the change.

7. Limits of the model

A RiderScore reflects the data available at the time of calculation. Thin or missing history reduces confidence, and confidence is reported alongside the score.

Scores are not predictions of individual future conduct and must not be used as the sole basis for an adverse decision.