Rosa

See the bias you supervise.

The independent bias-measurement instrument for the EU AI Act era.

Bias hides in your data, directly and through proxies.

It surfaces as unequal outcomes long after the data decisions were made.

ENCODED

Directly and by proxy

A protected attribute need not be present. A postcode stands in for race, a career gap for gender. Models recover what you removed.

INVISIBLE

Invisible until harm

Encoded bias surfaces as unequal decisions long after the dataset was built. Detection has to happen at the data level.

UNADDRESSED

Everyone measures, nobody removes

Toolkits and platforms assess, monitor, and document. The dataset leaves exactly as biased as it arrived.

Apple Card (2019) and COMPAS (ProPublica, 2016): the same failure, encoded in the data, surfacing as discriminatory decisions.

High-risk AI must be examined for dataset bias. With mitigation in place.

Article 10 applies from 2 December 2027.

Article 10 requires the datasets behind high-risk AI to be examined for possible biases, with appropriate measures to detect, prevent, and mitigate them.

Rosa produces exactly that examination, the mitigation, and the evidence, in one run.

2 DECEMBER 2027

EUR 35M / 7%

maximum penalties for the most serious breaches. Data-governance obligations for high-risk providers carry up to EUR 15M or 3% of worldwide turnover.

EU AI Act, Article 99

An independent measurement no supervised firm can dispute.

A reproducible bias score with a hash-verifiable manifest. When the supervisor can measure bias, the market follows.

TOOLKITS

Open-source fairness toolkits

Assess and mitigate at the data scientist's desk. A toolbox you assemble and run yourself. No evidence artefact, no service.

GRC PLATFORMS

Governance and GRC platforms

Measure, monitor, and document bias across the enterprise. Useful paperwork, but the dataset leaves exactly as biased as it arrived.

ROSA

Rosa

Removes the bias at source, preserves the data's statistics, and proves the change with an immutable, hash-verifiable record of every run.

Diagnose. Remove. Prove.

One run does all three. The evidence is a by-product of the work.

01 / DIAGNOSE

Find the bias

Point Rosa at a dataset and one protected attribute. It finds where bias is encoded, directly or by proxy, and scores it.

02 / REMOVE

Remove it at source

Rosa transforms the data so models no longer rely on the protected attribute, while preserving every column's statistics.

03 / PROVE

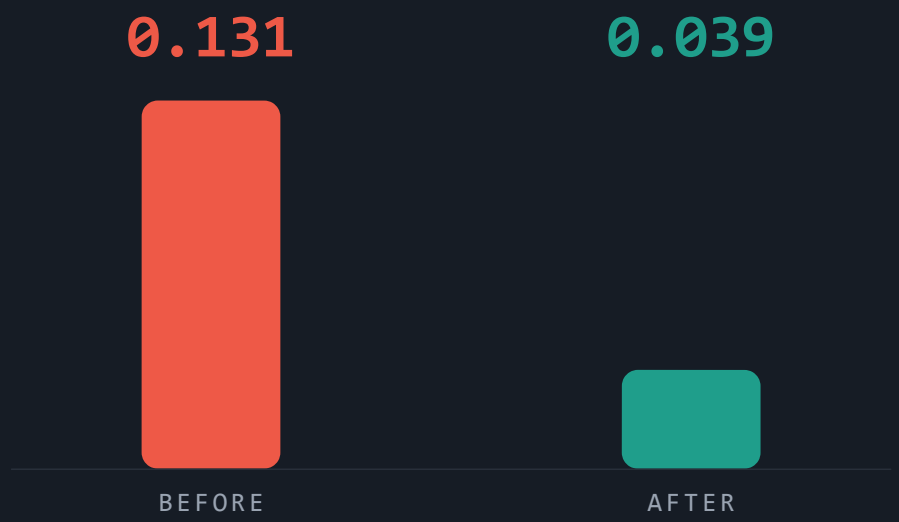
Keep the evidence

Every run emits an immutable Run Manifest. The audit evidence is generated by doing the work, not written up afterwards.

Univariate: one protected attribute per run (in this phase). Rosa declines to debias what it cannot statistically measure, and says so. Missing values are imputed (column mean or most-frequent).

The numbers do the persuading.

Measured on public and synthetic datasets. Reproducible with the shipped product.



-70% disparity ✓

ProPublica two-year cohort, held-out (82-run mean, 30-100%)

0.670 / 0.598 / 0.546

downstream R-squared: Rosa + a model, vs the model alone, vs a parity constraint.

Controlled credit-risk experiment

No bias reads ~0

held-out permutation test; ~0 on unbiased data of any width.

A bias score you cannot game

bit-for-bit

distribution preservation in both training and inference outputs, to float64 CSV precision.

Validation suite

Bias score: how far above a held-out permutation baseline a model can recover the protected attribute, on a 0 to 1 scale (higher = more bias).
About 70% lower race disparity (false-positive-rate gap, 82-run mean, 30.0-99.9%, Rosa v1.59.0) on ProPublica's two-year cohort — the honest hard case, at about a 14-point accuracy cost.
R-squared measures how well predictions fit reality; higher is better. Reproducible.

Every run leaves a record a firm cannot dispute.

Immutable, hash-verifiable, written once per job, kept indefinitely.

- **Every job emits:** an immutable Run Manifest, written once and kept indefinitely.
- **It carries the SHA-256 hash of the input:** anyone with the original CSV can recompute it and confirm what was examined.
- **Article 10 evidence as a by-product:** generated by doing the work, not assembled by hand before an audit.
- **REST API and an eight-tool MCP server:** so the run sits inside your own workflow.

RUN MANIFEST ✓ verified

job_id	550e8400-...440000
mode	remove_bias_training
job_status	complete
row_count	3,075
bias_columns	["race"]
input_hash	sha256:9f1c...e7a2
schema_hash	sha256:4b2a...01dd
config_hash	sha256:7e90...c3f1
container_digest	sha256:a17c...9b6e
bias (pre)	0.242
residual_bias	0.066
artifacts	report.pdf, compas_fair.csv

Immutable. Written once per job, kept indefinitely.
Illustrative values; field set mirrors the live schema.

Use cases with numbers, not adjectives.

The same engine across regulated domains.

FINANCIAL SERVICES · CREDIT

The Apple Card scenario

Proxy-encoded gender bias removed with no fairness/accuracy trade-off. Run it live as Customer Portal Test 1.

CRIMINAL JUSTICE · RECIDIVISM

COMPAS, the canonical biased dataset

Downstream race disparity cut about 70% (false-positive-rate gap, held-out) on the two-year cohort, at a ~14-point accuracy cost. Published model bias falls 0.85/0.92 to 0.00/0.17. ProPublica 2016.

HIRING AND HR

Proxies found and removed

Career gaps, postcodes, schools, names.
Published age bias 1.81 to 0.13 on the absenteeism dataset.
NYC Local Law 144 mandates audits.

HEALTHCARE · CLINICAL RISK

Estimates equalise after Rosa

Published heart-disease result: 0.44/0.44 healthy and 0.54/0.54 with disease, women and men.
The gender gap is effectively eliminated.

One instrument, three buyers.

Regulators measure it. Auditors evidence it. Enterprises fix it.

REGULATORS

See the bias you supervise

A standing, free Diagnose on the shared trial, with no end date. An independent, reproducible measurement with a manifest a firm cannot dispute.
When the supervisor can measure bias, the market follows.

AUDITORS

Article 10 evidence in one run

One standard, a hash-verifiable artefact, reusable across engagements, embeddable via REST and MCP.
Make bias audit a product, not a project.

ENTERPRISES

Fix it before the regulator finds it

Findings close instead of staying monitored. Audit prep collapses to a job run. The model keeps performing (R-squared 0.670 vs 0.598), and residency is yours to choose.

What sets Rosa apart.

Built to drop into a real pipeline, not a research project.

- **No proxy labelling required**

  Name one protected attribute; Rosa finds the proxies itself.

- **Any model or stack**

  Rosa outputs a clean CSV. No rewrite of your training code.

- **Data residency you choose**

  UK region for the trial; a dedicated instance in your own region for a PoC.

- **Works for human decisions too**

  Cleans the data analysts use directly, not just model training data.

- **Designed to protect predictive utility**

  Rosa transforms the data rather than constraining the model, with utility measured and reported alongside bias reduction.

- **Honest scope**

  Univariate: one protected attribute per run. Rosa says what it can and cannot detect.

Designed to support the frameworks enterprises answer to.

An evidence-producing instrument, not a certification.

EU AI Act | ISO/IEC 42001 | NIST AI RMF | SOC 2 | GDPR | EN 304 223

UK PROCESSING

Processed in the UK (eu-west-2, London); EU-UK adequacy.

ENCRYPTION

TLS in transit and EBS at rest; API keys stored as argon2id hashes.

ISOLATION

Per-job isolation; workspaces opaque to each other.

AUDIT LOGGING

Structured audit logging, no customer dataset content, 365-day retention.

RETENTION

Outputs expire in days; manifests are kept indefinitely.

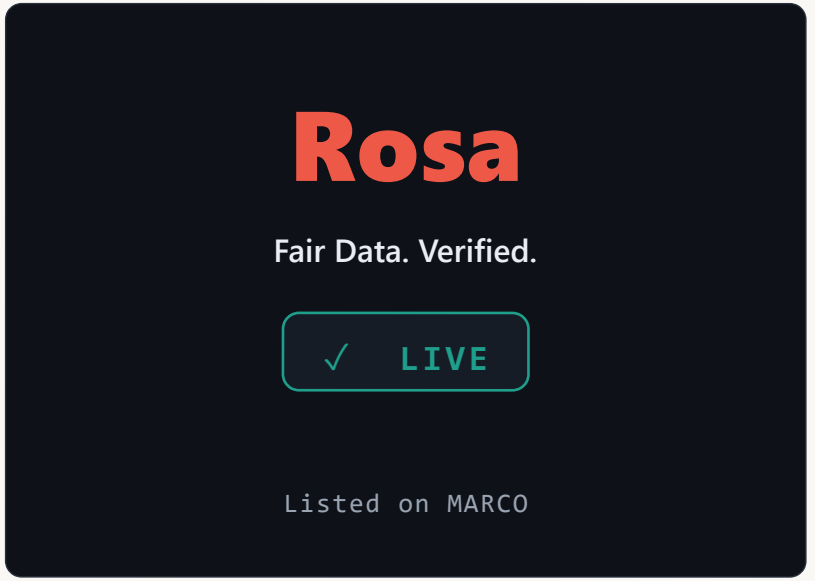
EVIDENCE, NOT CLAIMS

Rosa claims no certifications it does not hold.

Rosa is live and discoverable on MARCO.

Distribution through the marketplace where enterprises and AI agents find tools.

- **Listed and live on MARCO:** the AI-tool marketplace by AI Collaborator Inc, Los Angeles. Available now, not a roadmap promise.
- **A channel for buyers and for agents:** enterprises browsing for compliance tooling, and AI-agent toolchains via Rosa's MCP capability.
- **Free shared-instance trial on listing:** REST and MCP access, no procurement friction to evaluate.
- **10% discount for MARCO customers:** AI Collaborator offers MARCO customers a 10% discount when they sign up through the marketplace.



Start where the risk is zero.

Free-first: prove it on your data before any conversation about money.

FREE TRIAL · SHARED

Self-serve, REST and MCP

Via the Customer Portal, REST API, and MCP (use Claude/ChatGPT etc to test). For evaluation on small datasets: shared job queue, capped at 20 jobs per month.

Free for up to 3 months

FREE FOR REGULATORS · TIER 0

Diagnose, no end date

Free for supervisory authorities on the shared trial. For sensitive customer audits, regulators buy a dedicated instance.

Free, no end date

DEDICATED POC · YOUR REGION

Paid, production-scale

A dedicated Rosa Instance, in the region you choose, sized to your datasets, with all monthly-job caps lifted. A 12-week PoC on your production data.

Talk to us

Processed in the UK (eu-west-2, London) · No credit card for the free trial.

See the bias you supervise. Keep the evidence.

Measure it this week, free.

Start free → portal.rosadebias.com

Book a dedicated PoC → rosadebias.com/contact

Free to try. Free for regulators, with no end date. Processed in the UK.

enquiries@rosadebias.com · rosadebias.com